Université de Paris
Ecole doctorale de Chimie Physique et de Chimie Analytique de Paris Centre (ED 388)

Laboratoire de Biochimie Théorique UPR 9080
Institut de Biologie Physico-Chimique

Multi-scale simulation methods for biological membrane systems applied to confined environment and membrane fusion

Par Astrid BRANDNER

Thèse de doctorat de Chimie Théorique et Modélisation Moléculaire

Dirigée par Marc BAADEN et Fabio STERPONE

Présentée et soutenue publiquement le 10.12.2019

Devant un jury composé de :

Président du jury:	Samuela PASQUALI	PR, Univ. Paris Diderot, Paris, France
Rapporteur:	Syma KHALID	PR, Univ. Southampton, Royaume-Uni
Rapporteur:	Nicolas FLOQUET	CR, CNRS, Montpellier, France
Examinateur:	Marco LAURICELLA	CR, CNR, Roma, Italie
Directeur de thèse:	Marc BAADEN	DR, CNRS, Paris, France
Co-directeur de thèse:	Fabio STERPONE	DR, CNRS, Paris, France
Membre invité:	Sergio PANTANO	DR, ANII, Uruguay

Except where otherwise noted, this is work licensed under
https://creativecommons.org/licenses/by-nc-nd/3.0/fr/

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text.

Astrid Brandner
October 2019

Acknowledgements

Here I would like to acknowledge many people that were involved in different moments during the production of this thesis. Firstly I would like to thank my advisors Fabio Sterpone and Marc Baaden for their scientific guidance, encouragement and constant support during this thesis.

Secondly, I would like to thank all the members of the jury: Samuela Pasquali, Syma Khalid, Nicolas Floquet, Marco Lauricella and Sergio Pantano. Special thanks go to Nicolas Floquet and Syma Khalid for accepting to revise this manuscript and give useful feedback and to Sergio for accepting to be a guest jury and meet again since my days as an undergraduate and MSc student in Uruguay.

I would also like to thank Philippe Derreumaux for the help derived from his invaluable experience with the OPEP force field. To Simone Melchionna for his involvement in MuPhy debugging and for implementing modifications to the code that made it possible to carry out this work. Many thanks go also to Stepan Timr for his collaboration and helpful discussions with the lipid simulations (and his travel tips during the BPS meeting).

I would also like to mention Antoine Taly and Mickael Cohen that were key participants in the work related to mitochondrial fusion. To Mickael in particular for lending an experienced hand to interpret experimental information and willingness to share his knowledge related to mitofusin dynamics. Also to Dario De Vecchis, the "OG" PhD student to study Fzo1 for all the work we did together.

Of course, I have also to thank all the lab members with whom we crossed paths during these last 3 years and making the LBT a motivating place to work. In particular the people from our cozy "tropical climate" office: Samuel, Dario, Anne-Elisabeth, Karen, Marina, Olivier (merci pour ton aide avec le resume!). Also to Nicolas with whom we are constantly deciphering the bureaucratic ways of the ED388 and sharing the experience with the younger PhD generations (you are welcome Matt and Alejandro!). Special thanks also go to LABEX Dynamo for funding my thesis, to Geoffrey Letessier, for the internal cluster maintenance and to Victoria Terziyan for her on-going help since my very (troubled!) first day in Paris.

Finally I would like to thank my family and friends from (far) away for still being part of this journey regarding of the distance and time-zones. Particularly to my parents, who overcame tough times but were even able to visit Paris during my stay here.

Abstract

In this thesis we have used and developed computational approaches to study biological membrane systems through molecular dynamics simulations. In particular the focus was on coarse-grained models as well as multi-scale simulations. By coarse-graining (CG) we reduce the degrees of freedom in a system in comparison to a traditional all-atom model. This simplification enables us to study bigger systems as well as extend the time-scale of our simulations with less computational cost.

Firstly, we have addressed some structural aspects in the early stages of mitochondrial membrane fusion. Mitochondrial dynamics is based on a delicate balance between fusion and fission that allows a normal cell metabolism. To approach the study of mitochondrial fusion, our current work focuses on the scaffolding and tethering of the mitochondrial outer membranes. Here we used an homology model of Fzo1, an outer mitochondrial membrane protein from yeast, to study its contribution to membrane fusion and tethering. For this purpose we have used coarse-grained approaches to understand the conformational dynamics and assembly of this system at different oligomerisation states and levels of detail. Our results contributed to share with the community structurally rigorous models, that were not available before. Moreover, we have identified a possible tetramer complex that is in agreement with experimental information available related to a so-called docking ring (one of the first steps in fusion and tethering) observed via cryo-EM.

Besides the study of Fzo1 tethering, we have also implemented a new methodology to couple an implicit solvent coarse-grained force field for lipids (Dry Martini) with a mesoscopic representation of the fluid via the Lattice Boltzmann approach (LB). This development opened up the possibility to study systems where the hydrodynamic interactions are a key part. On top of that, we have continued the development of an implicit-solvent CG protein force field (OPEP) previously developed in the Lab, to also include lipids and lipid-protein interactions. This choice was justified by the fact that OPEP does not restrain the secondary structure of the proteins and has a detailed representation of the backbone. A first version of this coupling between lipid and protein models has been created and tested. Our results indicate, in general terms, that transmembrane peptides are able to be stable in a

membrane environment. However, further modifications will be needed in particular to refine the tryptophane parameters.

All in all, we hope that the new methodologies presented in this work will aid to improve the modelling of membrane and membrane-protein systems at a pseudo-atomistic coarse-grain level, as well as including mult-scale approaches.

Keywords: coarse-grained, molecular dynamics, membrane, hydrodynamics, multi-scale

Résumé

Dans cette thèse, nous avons utilisé et développé des approches informatiques pour étudier les systèmes de membranes biologiques. En particulier, l'accent a été mis sur les modèles gros grain ainsi que sur les simulations multi-échelles. Cette stratégie gros grain permets de réduire le nombre de degrés de liberté dans un système par rapport à un modèle traditionnel tout-atome. Cela permet d'étudier des systèmes plus grands ainsi que d'étendre la durée des simulations avec un coût de calcul moindre.

Tout d'abord, nous avons abordé certains aspects structuraux des premières étapes de la fusion membranaire mitochondriale. La dynamique mitochondriale repose sur un équilibre délicat entre fusion et fission, et qui garantit un métabolisme cellulaire normal. Pour l'étude de la fusion mitochondriale, nos travaux se concentrent sur l'approche et l'ancrage des membranes externes mitochondriales. Nous avons utilisé ici un modèle d'homologie de Fzo1, une protéine de la membrane mitochondriale externe des levures, pour étudier sa contribution à la fusion et à l'ancrage membranaire. Pour ce faire, nous avons utilisé des approches gros grain (CG) pour comprendre la dynamique conformationnelle et l'assemblage de ce système à différents niveaux d'oligomérisation et de détail. Nos résultats ont contribué à partager avec la communauté des modèles structurellement rigoureux, qui n'étaient pas disponibles auparavant. De plus, nous avons identifié un complexe tétramère possible en accord avec les informations expérimentales disponibles concernant l'anneau d'amarrage (l'une des premières étapes de la fusion et de la fixation).

Outre l'étude de l'ancrage de Fzo1, nous avons également mis en œuvre une nouvelle méthodologie pour coupler un champ de force gros grain à solvent implicite pour les lipides (Dry Martini) avec une représentation mésoscopique du fluide via l'approche Lattice Boltzmann (LB). Cela a ouvert la possibilité d'étudier des systèmes où les interactions hydrodynamiques sont un élément clé. De plus, nous avons poursuivi le développement d'un champ de force protéique CG (OPEP) en solvant explicite, développé précédemment dans notre laboratoire, pour inclure également les lipides.

Grâce à ce couplage, nous pourrons modéliser les changements conformationnels des protéines membranaires. Une première version de ce couplage entre modèles lipidique et protéique a été paramétrée et testée. Nos résultats indiquent que les peptides transmem-

branaires modélisés sont stables dans un environnement membranaire. Toutefois, d'autres modifications seront nécessaires, en particulier pour affiner les paramètres du tryptophane.

Dans l'ensemble, nous espérons que les nouvelles méthodologies présentées dans ces travaux contribueront à améliorer la modélisation des systèmes membranaires et de systèmes membrane-protéines à un niveau gros grain pseudo-atomique, ainsi qu'à inclure des approches à plusieurs échelles.

Mots clefs : gros grain, multi-échelles, dynamique moléculaire, membrane, hydrodynamique

Table of contents

List of figures **xv**

List of tables **xxix**

Nomenclature **xxix**

1 Introduction **1**

1.1 Biological membranes 1

1.1.1 Eukaryotic cell membrane composition 1

1.1.2 Mitochondrial lipid composition 2

1.1.3 Structural overview of a cell membrane bilayer 3

1.2 Membrane proteins and peptides 3

1.2.1 Lipid-protein interactions 6

1.3 GTPases 7

1.3.1 Dynamins and Dynamin-like-proteins (DLPs) 7

1.3.2 The role of DLPs in mitochondrial membrane fusion 9

1.4 Molecular dynamics simulations of membrane protein systems 12

1.4.1 Membrane protein structures databases: The structures we have so far 13

1.5 Main objectives 14

2 Theoretical background and methods **15**

2.1 Molecular Dynamics simulations 15

2.1.1 Fundamentation and examples 15

2.1.2 Generalities 16

2.1.3 Force Fields 16

2.1.4 Technical details for a simulation setup: Periodic boundary conditions and long range electrostatics treatment 20

2.1.5 Theory behind Molecular Dynamics (MD) 22

2.1.6 Algorithms in MD 22

2.2 Coarse-grained methods . . . 25
2.2.1 Generalities . . . 25
2.2.2 Lipid Models . . . 26
2.2.3 Protein Models . . . 28
2.3 Including Hydrodynamics Interactions . . . 32
2.3.1 Lattice Boltzmann technique . . . 33
2.3.2 Lattice Boltzmann Molecular Dynamics . . . 35

3 Computational study of mitofusin assembly 37
3.1 Introduction . . . 37
3.2 Methods . . . 38
3.2.1 Modelling the Fzo1 GTPase dimer construct . . . 39
3.2.2 Modelling cis-dimer configurations . . . 40
3.2.3 Modelling the trans-dimer configurations . . . 41
3.2.4 Molecular Dynamics . . . 41
3.2.5 Analysis . . . 42
3.3 Results and discussion . . . 43
3.3.1 Putative models for Fzo1 cis-dimers . . . 43
3.3.2 Proposed models for Fzo1 trans-tethered oligomers . . . 46
3.3.3 Trans-dimer dynamics . . . 48
3.4 Conclusion . . . 52
3.5 Supplementary Material associated to this publication . . . 55

4 Adding Hydrodynamics to Lipids 65
4.1 Introduction . . . 65
4.2 Methods . . . 67
4.2.1 Dry Martini force field . . . 67
4.2.2 Lattice Boltzmann MD . . . 68
4.2.3 Systems and simulation setups . . . 69
4.3 Results . . . 71
4.3.1 The implementation of Dry Martini . . . 71
4.3.2 Coupling with HI . . . 72
4.3.3 Lipid lateral diffusion . . . 72
4.3.4 Vesicles in fluid flows . . . 75
4.3.5 Nanoreactor . . . 80
4.4 Conclusions . . . 81
4.5 Supplementary Material associated to this publication . . . 85

5 Expanding OPEP force field to study protein-lipid systems: fat OPEP 91
5.1 Generalities and Implementation 91
5.1.1 Initial energy comparisons between the different force fields 92
5.2 Parametrization of lipid-protein interactions 95
5.2.1 Tryptophan (TRP) 96
5.2.2 Other hydrophobic cyclic amino acids: tyrosine (TYR) and phenylalanine (PHE) 98
5.2.3 Histidine (HIS) 100
5.2.4 Charged amino acids: Lysine (LYS) and Arginine (ARG) 101
5.2.5 Remarks and final considerations about re-parametrisation 101
5.3 Validation: Testing small transmembrane peptides 102
5.3.1 System preparations and simulation setup and analysis 103
5.3.2 The case of polyleucine: a stable polypeptide inside the bilayer . . 104
5.3.3 Flanked peptides 106
5.4 Conclusion 110
5.5 Supplementary Material associated to this chapter 111

6 Conclusions and perspectives 113

7 Résumé de la thèse 117
7.1 Introduction générale au résumé 117
7.2 Partie 1: Complexes de fusion de membranes mitochondriales externes . . . 117
7.3 Partie 2: Ajouter l'hydrodynamique aux lipides 122
7.4 Partie 3: Couplage lipide-protéine 125
7.5 Conclusion finale et perspectives 129

References 133

Appendix A Peer-reviewed publications derived from this thesis 155

List of figures

1.1 Proportional digram of a typical cell composition by type of molecules *Taken from [169].* 2

1.2 Lipid composition of a mitochondrion belonging to yeast with its associated 2D representation(*Saccharomyces cerevisiae*). The values are reported for both membranes together and for the inner and outer membrane separately. Values taken from [115] 3

1.3 Fluid mosaic membrane representations.(a) Historical representation of possible membrane protein insertion for the major erythrocyte membrane proteins. The carbohydrates of the glycoproteins are shown as green rods exposed to the extracellular space[238]. (b) More detailed representation of a fluid mosaic model colouring the lipids according to their type. Different regions can be appreciated. *Adapted from [75]* 4

1.4 DLP Mgm1 forms a regular protein coat on the outer surface of lipid tubes. (a) Four filaments of Mgm1 dimers wrap around a membrane tube in a left-handed surface lattice (b) cryo-electro tomography showing the protein lattice on the membrane surface. *Adapted from [79]* 5

1.5 Typical adaptations to counteract the effect of either positive or negative hydrophobic mismatch of a transmembrane (TM) α-helix. To be noted, that some modifications affect the membrane (change of the lipid ordering, disruption of bilayer character) whereas the majority affect the peptide's structure or its relative positioning towards other peptides or in the bilayer. *Adapted from [68]* 6

1.6 Example DLPs localisation in a cell *Taken from [124].* 8

1.7 Schematic representation of the modular domain architecture of selected proteins from the dynamin superfamily. (a) Cartoon summarizing the modular domain architecture for selected DLPs. The classical nomenclature is indicated. The helical bundles HB1 and HB2 of the mitofusins/Fzo1 are aligned to the stalk/neck domains of other DLPs and dynamins showing a similar architecture and they might serve as BSE (Bundle Signaling Element). MTS: Mitochondrial Targeting Sequence, TM:transmembrane domain, PAD: paddle, PH: Pleckstrin Homology, PRD: Proline-Rich Domain. (b) The BDLP conformational change upon GTP and lipid binding. Hinges 1a, 1b, 2a and 2b are indicated. The structure is closed in its GDP-bound state (not shown) and in its apo-form (left) (PDB: 2J68 [152]). Lipid binding concomitant with GTP hydrolysis, reveals a conformational change detected using cryo-electron microscopy (right) (PDB: 2W6D [153]). (c) (Left) The Fzo1 structural model used in this thesis [70] with HB1 and HB2. (Center) Crystal structure of the human Mfn1 fragment-construct obtained upon addition of GDP/AlF4- (not shown) (PDB: 5GOM [46]) and observed in presence of GDP/BeF3- (not shown) (PDB: 5YEW [271] (Right). (d) Cryo-electron microscopy of the human dynamin 1 (Dyn1) in complex with GMPPCP (not shown) (PDB: 6DLU [134]). The PRD was not present in the reconstruction. *Figure taken from [69].* 10

1.8 Growth in membrane protein entries as a function of the year of the mpstruc database. 13

2.1 Table with all the interaction types and potential energy expressions of a typical force field. The potentials are plotted below to show the different shapes. 17

2.2 POPC lipid representations for a classical all atom force field, charmm36m (a), and for Martini 2.1 coarse-grained force field (b). 26

2.3 Martini topology for amino acids (including backbone bead). *Taken from [171]* 29

2.4 Schematic representation of the coupling between the molecular system (CG, MD) and the fluid (Lattice Boltzmann). The blue arrows represent the fluid velocity population that evolves on the lattice. We also show a depiction of interpolation and extrapolation strategies used to comunicate forces between particle and fluid dynamics. Finally, we show the 3D lattice topology used in this work: D3Q19. *Adapted from [239, 246]* 36

3.1 Key steps of the modelling workflow to build a model of Fzo1 in open conformation, based on the BDLP structure. a. The Fzo1 structural blocks delimited according to the putative hinges proposed for BDLP are highlighted by different colours. The blocks are coloured in rainbow from N-terminal (red) to C-terminal (blue). b. Each block in Fzo1 (coloured parts) was superposed onto its corresponding fragment in 2W6D (grey ribbon). Note the backbone interruptions in the structure. c. The final Fzo1 model in open conformation. Note that each GTPase domain (orange) from both chains is in close contact with each other. The GDP nucleotide is shown in space-filling representation. A cartoon scheme showing the colouring and sequence positioning of the structural domains is shown at the bottom. 39

3.2 Putative Fzo1 models for interaction in cis configuration. a. and b. Wide and narrow head-to-head complexes, respectively, in which two Fzo1 molecules interact via their GTPase domains, consistent with interactions observed for the bacterial BDLP [152, 153] and human mitofusin [46]. c. Back-to-back complex for the closed model where two Fzo1 molecules interact through their HR domains. The domains highlighted in colour are: violet, HRN (N-terminal heptad-repeat); green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (blue) from lipid bilayer headgroups and GDP nucleotides are depicted in the space-filling representation. A cartoon scheme showing the colouring and sequence positioning of the functional domains is shown at the bottom. 44

3.3 Lipid-protein and protein-protein interface evolution during CG MD simulations for the dimeric complexes. a. Lipid-protein interfaces for the three different studied cis-dimers. b. Protein-protein interfaces belonging to the three different cis-dimers. The colours in the curves match the outline of the corresponding image in the inset (black: head-to-head narrow, red: head-to-head wide, cyan: back-to-back). The domains in the inset schemes are coloured in the same way as Fig. 2. 45

3.4 Exposed Ugo1-interacting regions in Fzo1 putative cis-dimers. a., b. and c. Final snapshots from the simulations showing the Ugo1-interacting surfaces (yellow) for each of the cis-dimers. Each Fzo1 monomer is represented with a different colour (white or grey) and by a letter (A or B); the membrane bilayer is depicted as a silver transparent surface. The solvent accessible surface areas for the Ugo1-interacting regions corresponding to each monomer are reported at the bottom. 47

3.5 Putative model complexes for Fzo1 trans-tethering interactions. a. The Fzo1 models in open conformations, as suggested from analogy to the BDLP system, interacting through their GTPase domains. b. The Fzo1 trans-tetramer in which the HRs are oriented in a parallel fashion and the trans interaction occurs towards the GTPase domain. c. The Fzo1 trans-tetramer in closed conformation. Here the interaction occurs towards the GTPase domain as well as through the respective HR domains oriented in an antiparallel fashion. Outer membranes corresponding to two different mitochondria are labelled m1 and m2. The domains highlighted in colour are: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (blue) from lipid bilayer headgroups and GDP nucleotide are depicted in the space-filling representation. 48

3.6 Lipid-protein and protein-protein interface evolution during CG MD simulations for the tetrameric complexes. a. Lipid-protein interfaces for the three different studied tetrameric complexes (black: head-to-head trans open, red: head-to-head trans parallel, cyan: back-to-back trans antiparallel). b., c. and d. depict the evolution of the characteristic monomer-monomer interfaces belonging to the trans-tetramer in open conformation, the trans-tetramer in which the HRs are oriented in parallel and the trans-tetramer with HRs oriented in an antiparallel fashion. Legends A, B, C and D represent each of the monomers in the complex, following the colouring of the image inset. . 49

3.7 Final configurations of the putative model complexes for Fzo1 trans-tethering interactions after relaxation through molecular dynamics simulations. a., the Fzo1 models in open conformation, as suggested from the BDLP system analogy, interacting through their GTPase domains. b., the Fzo1 trans-tetramer in which the HRs are oriented in a parallel fashion and the trans interaction occurs towards the GTPase domain. c., the Fzo1 trans-tetramer in closed conformation. Here, the interaction occurs towards the GTPase domain as well as through the respective HR domains oriented in an antiparallel fashion. In all figures the domains are coloured: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (dark blue) from the lipid headgroups of two mitochondria outer membranes (m1, m2) are depicted as spheres, whereas the tails are highlighted as a blue shadow. Intermembrane distances between the outer phosphate layer of the membranes are indicated. 50

3.8 Exposed Ugo1-interacting regions for Fzo1 trans-tethering complexes. a., b. and c. Final snapshots from the simulations showing the Ugo1-interacting surfaces (yellow) for each of the trans-tethering complexes. Each Fzo1 monomer is represented with alternate colours (white or grey) and by a letter (A, B, C or D); the membrane bilayer is depicted as a silver transparent surface. The solvent accessible surface area for the Ugo1-interacting regions corresponding to each monomer are reported at the bottom. 51

3.9 Putative membrane-inserted Fzo1 models interacting in cis configuration after molecular dynamics relaxation. a and b. Wide and narrow head-to-head complexes, respectively. c. Back-to-back complex of the closed model. The domains are coloured: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (dark blue) from the lipid bilayer headgroups are depicted as spheres, whereas the bilayer is highlighted as a blue shadow. 55

3.10 a. The Fzo1 GTPase dimer construct in which the canonical G-interface commonly observed in the dynamin superfamily [66] is reconstituted. b. The back-to-back HR-parallel trans-tetramer during the modeling procedure (see Method). The Figure shows how the two GTPase domains that interact in trans share an apparent incorrect orientation (compared with a.). Without an extensive conformational rearrangement that would involve both hinges 1a and 1b, the formation of the canonical G-interface observed in the dynamin superfamily [66] would be impeded. In particular, the dotted line should be located on the same side. The domains are: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (blue) from lipid bilayer headgroups and GDP nucleotide are depicted in the space-filled representation. 56

3.11 Inter-bilayer difference distance matrix (Å). Left panel shows the final difference in distances compared to the initial structure, red regions correspond to regions that got closer in space compared to the initial structure of the production run whereas blue regions represent areas that moved further apart. The right panel shows a rotation of 90 degrees around the y-axis to visualize the position of the complex. 57

3.12 Interface details from the cis-dimers. a. Narrow head-to-head complex, b. Wide head-to-head complex. c. Back-to-back complex. The domains are coloured in: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Darker shade of the same colours are used to differentiate between monomers when needed. The residues in contact between both monomers are highlighted with a transparent surface coloured by occupancy with a green gradient (grey: minimum value, green: maximum value). For clarity, only the regions corresponding to the interface are shown. 58

3.13 Molecular dynamics simulation details for each system. 59

3.14 RMSD values averaged over the last 100 ns of simulation for each dimer (Fig. 2a, b, c) and tetramers (Fig. 3a, b, c). Reported values include a full RMSD calculated over all the backbone particles (BB) as well as by every chain and structural fragment separately. 60

3.15 Interface residues for the narrow head-to-head cis-dimer by chain (ch) residue number (resid) according to UNIPROT as shown in previous figures. Occupancy (occup) values are calculated on the last 100 ns of the simulations. . . 61

3.16 Interface residues for the wide head-to-head cis-dimer by chain (ch) residue number according to UNIPROT (resid) as shown in previous figures. Occupancy (occup) values are calculated on the last 100 ns of the simulations. . . 62

3.17 Interface residues for the back-to-back cis-dimer by chain (ch) residue number according to UNIPROT (resid) as shown in previous figures. Occupancy (occup) values are calculated on the last 100 ns of the simulations. 63

4.1 Left panel. Attractive and repulsive Coulomb potential with relative dielectric constant $\varepsilon = 15$ compared to the switched electrostatic potentials used in the Dry Martini force field to describe the electrostatic interactions among charged groups. Right panel. Tabulated potentials that incorporate the electrostatic switched interactions and the short-range LJ interactions for the phosphate (P) and choline (C) charged groups of the lipids. 68

4.2 Probability distributions of bonds, angles and charged beads distances. Black lines and gray surfaces represent results obtained from a simulation carried out using Gromacs, and the orange lines represent the results from our implementation in Muphy. Panels A and B present the probability distributions of each of the unique bond or angle types present in Dry Martini respectively. The panel C depicts the probability distributions of the distances between charged beads (NC3: choline bead, PO4: phosphate bead). 71

4.3 Spontaneous bilayer assembly of 128 POPC lipids. The panels A-D show the time evolution of the assembly process. For the sake of clarity, the phosphate beads are represented as van der Waals spheres, whereas the rest of the lipid is shown as lines. The system is represented with periodic images, and the orange-coloured particles belong to the simulation box. 73

4.4 **Box size effect on lipid diffusion.** Diffusion coefficient for POPC lipid extracted from 800 ns simulation of a planar membrane in different boxes of size $L_x = L_y = L$ and $L_z = 240$ Å(cyan circles). The diffusion coefficient D was obtained from a linear fit of the MSD computed during the last 800 ns of the simulation and using a time window of 40-70 ns for the fit. In the figure (black circles) we also report the data from Ref. [274] based on STRD simulations of the POPC lipid planar membrane with inter-membrane separation of 200 Å. Inset: Log-linear plot of the diffusion coefficients. . . . 74

4.5 Nanotube adsorption. Time evolution of the minimal distance of the nanotube centre of mass with respect to the external surface of the lipid bilayer for one of the simulated systems. Starting from the same initial condition two simulations were performed with (LBMD, cyan) and without (Langevin, orange) HI. On the top part of the figure we represent the streamline of the fluid surrounding the nanotube during its adsorption process. 76

4.6 Vesicle in shear flow. Time evolution of the three eigenvalues of the vesicle gyration tensor for shear rates $\dot{\gamma} = 3,7,20,30 \times 10^9 s^{-1}$. 78

4.7 Vesicle in shear flow. In the top layer of the figure (Panel A) we represent the short time rotational motion of the vesicle under shear flow. Two lipid beads corresponding to the outer layer are highlighted (yellow and green spheres) to help visualising the membrane rotation. The yellow and green beads remain aligned at a short time, but at a longer time, as we show in SI Figure S4, the different lipid mobility in the bilayer causes their mis-alignment. In panel B and C we represent the time evolution of the normalised projection on the Z axis of the system of two characteristic vectors of the vesicle: the eigenvector $\Lambda^{(3)}$ of the gyration tensor and associated to the longest axis of the prolate (Panel B), and the vector $\mathbf{v}$ connecting one polar bead of the lipids in the membrane (yellow sphere) to the centre of mass of the vesicle (Panel C). For the projection of the eigenvector we considered the absolute value. 79

4.8 Vesicle in shear flow. Time evolution of the shape of a larger vesicle ($R \sim 100\,\text{Å}$) under the action of the shear flow. Panel A, B and C refer to simulations at different shear rate, $\dot{\gamma} = 3.7 \times 10^7 s^{-1}$, 1.9 and $3.7 \times 10^9 s^{-1}$, respectively. Two lipid beads corresponding to the outer and inner layer are highlighted (yellow and green spheres, respectively) to help visualising the membrane rotation and the lipid motion inside the membrane. 79

4.9 Substrate release in shear flow. Time evolution of the single substrate distance with respect to the vesicle centre of mass. In the top part of the figure we represent the molecular detail of the substrate release. Clusters of substrate molecules are highlighted by dashed lines. Arrows indicate the tumbling motion of the vesicle. 80

4.10 Nanoreactor. Panel A. Pictorial representation of the nanoreactor and the internal and external fluid. Panel B. Probability distribution of proteins and substrate localisation inside the vesicle. Panel C. Proteins and substrate mean square displacement in dilute condition (dashed lines) and inside the vesicle (solid lines). 82

4.11 Left. Time correlation function of the membrane thickness fluctuation for a membrane system simulated with and without hydrodynamics, and using two different value of the friction coefficient γ. Right. Radial distribution function of hydrophobic beads for a simulation of a POPC bilayer with (HI, dotted cyan line) and without (no HI, black line) hydrodynamic interactions. 85

4.12 Lateral mean square displacement calculated (MSD) for the lipids in a planar membrane of size $L_x = L_y = L$ and fixed $L_z = 243$ Å. For each system the MSD is calculated using independent blocks of the trajectory, each of 100 ns. The diffusion coefficient was derived by a linear fit in the time window 20-40 ns (gray zone). 86

4.13 Vesicle sphericity computed as the ratio of the eigenvalue of the vesicle gyration tensor as a function of the shear rate. 87

4.14 Time evolution of the angle formed between the polar heads of two lipids in the vesicle under the action of shear flow with $\dot{\gamma} = 20 \times 10^9 s^{-1}$. 88

4.15 Vesicle under shear flow. At very high shear rate, $\dot{\gamma} = 30 \times 10^9 s^{-1}$, the vesicle breaks in smaller entities that encounter, fuse and disconnect as effect of the solvent shearing. The polar heads of two set of lipids, initially separated in the vesicle, were coloured in red and blue. The time evolution of the system shows their mixing. 89

4.16 Vesicle under shear flow. In panel A we represent the time evolution of the normalised projection on the Z axis of the system of the eigenvector $\Lambda^{(3)}$ of the gyration tensor associated to the longest axis of the vesicle. The absolute value of the projection is reported in the figure for the three shear rates $\dot{\gamma} = 1.9$ (violet) and 3.7 (cyan) $\times 10^9 s^{-1}$ and $\dot{\gamma} = 3.7 \times 10^7 s^{-1}$ (black). Panel B-D: projection of the the vector **v** connecting one polar bead of the lipids in the membrane to the centre of mass of the vesicle. In each panel the green and yellow curves refer to two different beads in the membrane. Panel B refers to the lowest shear rate while Panel D to the highest. 90

5.1 Martini topology for amino acids colored by polarity (including backbone bead). *Taken from [171]* 92

5.2 Dry Martini v2.1 POPC mapping highlighting the different beadtypes by colours: C1 (cyan, representing the hydrophobic beads from the acyl chains), C3 (grey, representing the unsaturation in the acyl chain), Na (pink, representing the glycerol moiety), Qa (brown, representing the phosphate) and Q0 (blue, representing the choline moiety) 92

5.3 Non-bonded interaction potentials between the different POPC lipid beads in Dry Martini (v2.1) (orange).The non-bonded interaction potential fo their analogous beads in OPEP protein force field (v4) are also plotted to be compared (black). The cyan line shows the attractive ion-pair potential for charged-charged interactions as implemented in the OPEP force field (v5, OPEP IP). The asterisk highlights the case with a marked energetical difference: the interaction between the hydrophobic bead for the lipid (C1) and its counterpart in the protein model (LEU). *For an extended comparison of small polar and apolar interactions see figure 5.15 in the Appendix* . . . 93

5.4 Second strategy to obtain an effective potential between a side-chain and each of the lipid beads. A set of concentric spheres is located around the side-chain of the studied amino acid of interest. Lipid beads (in this example C1) are located along each sphere surface of radius r. For each sphere, an average effective potential is calculated in order to plot $E_{LJ}(r)$ (exemplified in red for a value of r). Finally, the potential is fitted using a LJ functional form to obtain a first set of parameters, σ and ε. 95

5.5 Example of the whole parametrisation process in the case of tyrosine (TYR) as a representative amino acid. Panel a: Setup for PMF calculations showing the initial coordinates of the window corresponding to z = 0 nm (zoom on the region surrounding the amino acid side-chain). Panel b: Example of different PMFs obtained from the reference simulation and different LJ guesses either with strategy 1 (1 bead SC4) or strategy 2 (1 bead guess). For the first strategy several tests are shown with different ε (eps) and σ (sig) values. Panel c: Radial distribution functions between the side-chain and hydrophobic core beads present in the lipid, for the reference system (black), best PMF result after fitting when modifying ε values only (turquoise), and final result (red) 97

5.6 PMFs obtained from switching off each of the beadtypes in TRP, HIS, ARG and LYS amino acids. The PMF obtained with the original parameters is plotted in black as a reference. TYR and PHE side-chain's PMF are omitted in these graphs as TYR is composed by the exact same bead types as TRP (and has an analogous curve) and PHE's side-chain beads belong to the same bead type. 98

5.7 Final PMF curves for the six treated amino acids. Black lines correspond to the reference simulation (PMF^{ref}) whereas red lines correspond to the simulation with the final parameters as shown in table 5.2 99

5.8 Radial distribution functions between the side-chains of all the six multi-beaded side-chain amino acids and characteristic beads from the hydrophobic core of the membrane. From left to right the lipid beads that are being plotted are: C1, C3 and Na. Panels a-f correspond to the amino acids: tryptophan, tyrosine, phenylalanine, histidine, arginine, lysine 100

5.9 a: Position in Z-axis of the center of mass of L10 during the simulation with its associated histogram. The average position of the center of mass of the phosphates in the upper and lower monolayer are also plotted to have as a reference. b: L10 tilt angle with respect to the normal of the bilayer. c: Secondary structure assignment for simulations in All Atom and with OPEP/Dry Martini coupling. d: Initial and final snapshots showing the peptide inside the bilayer for each of the force fields used. The protein backbone is shown as pink and side-chain beads in yellow for all the coarse-grained models. The first backbone atom for the peptide is coloured in dark pink to recognise the orientation of the helix. To help visualisation phosphate beads are shown as vdW spheres in tan and the membrane is depicted as a transparent surface. 105

5.10 Example of amino acid distributions in the transmembrane region of a potassium channel (*Streptomyces lividans KcsA potassium channel*). Charged amino acids are shown in red whereas hydrophobic amino acids are depicted in green and aromatic in yellow. *Taken from [68]* 106

5.11 Left:End-to-end distance of WALP helix: distance between first and last alpha carbon in the peptide (All atom, OPEP) or between the first and last backbone beads (Dry Martini, Martini). Right: Secondary structure assignment plot for the WALP23 run in All atom and with OPEP/Dry Martini coupling; magenta: α-helix, blue: 3-10 helix, white: coil 107

5.12 Left:Position in Z-axis of the center of mass of WALP23 during the simulation. The average position of the center of mass of the phosphates in the upper and lower monolayer are also plotted to have as a reference. Right: WALP23 tilt angle with respect to the normal of the bilayer with the associated histogram from the angle population. 108

5.13 End-to-end distance of KALP helix: distance between first and last alpha carbon in the peptide (All atom, OPEP) or between the first and last backbone beads (DryMartini, Martini). Right: Secondary structure assignment plot for the KALP23 run in All atom and with OPEP/Dry Martini coupling; magenta: α-helix, blue: 3-10 helix, white: coil. 109

5.14 Left: Position in Z-axis of the center of mass of KALP23 during the simulation. The average position of the center of mass of the phosphates in the upper and lower monolayer are also plotted as a reference. Right: KALP23 tilt angle with respect to the normal of the bilayer with its corresponding histogram. 109

5.15 Plot with all the small polar and apolar amino acids with POPC lipid beads. 111

7.1 Configurations finales des complexes de modèles putatifs pour les interactions trans théoriques Fzo1 après relaxation par simulations de dynamique moléculaire. a. Les modèles Fzo1 en conformation ouverte, comme suggéré par l'analogie du système BDLP, interagissent à travers leurs domaines GTPase. b. Le transtétramer Fzo1 dans lequel les HRs sont orientés en parallèle et l'interaction trans se produit vers le domaine GTPase. c. Le trans-tétramer Fzo1 en conformation fermée. Ici, l'interaction se produit vers le domaine GTPase ainsi qu'à travers les domaines HR respectifs orientés de manière antiparallèle. Dans toutes les figures, les domaines sont colorés : violet, HRN ; vert, HR1 ; orange, HR2 ; rouge, GTPase et jaune, transmembrane. Les atomes de phosphore (bleu foncé) des deux membranes mitochondriales externes (m1, m2) sont représentés sous forme de sphères, tandis que le restant des membranes est souligné par une ombre bleue. Les distances intramembranaires entre la couche externe de phosphate des membranes sont indiquées. 119

7.2 a:Représentation schématique du couplage entre le système moléculaire (CG, MD) et le fluide (Lattice Boltzmann). Les flèches bleues représentent les populations de vitesses du fluide évoluant sur le réseau. Nous montrons également une description des stratégies d'interpolation et d'extrapolation utilisées pour communiquer les forces entre la dynamique des particules et celle du fluide. *Modifié de [239, 246]* b: Effet de la taille de la boîte sur la diffusion lipidique. Coefficient de diffusion du lipide POPC extrait d'une simulation de 800 ns d'une membrane plane dans différentes boîtes de taille $L_x = L_y = L$ et $L_z = 240$ Å(cercle cyan). Le coefficient de diffusion D a été obtenu à partir d'un ajustement linéaire du MSD calculé pendant les 800 dernières ns de la simulation et en utilisant une fenêtre temporelle de 40 à 70 ns pour l'ajustement. Dans la figure (cercles noirs), nous présentons également les données de la Réf. [274] basées sur des simulations STRD de la membrane planaire lipidique POPC avec une séparation inter-membrane de 200 Å. Encart : Diagramme log-linéaire des coefficients de diffusion. . . 124

7.3 Exemple de l'ensemble du processus de paramétrisation dans le cas de la tyrosine (TYR) comme un acide aminé représentatif. a : Configuration pour les calculs de PMF montrant les coordonnées initiales de la fenêtre correspondant à z = 0 nm (zoom sur la région entourant la chaîne latérale des acides aminés). b : Exemple de différents PMF obtenus à partir de la simulation de référence et de différentes suppositions LJ soit avec la stratégie 1 (un pseudo-atome SC4) ou la stratégie 2 (un pseudo-atome deviné). Pour la première stratégie, plusieurs tests sont présentés avec différentes valeurs ε (eps) et σ (sig). c : Fonctions de distribution radiale entre la chaîne latérale et les pseudo-atomes hydrophobes présents dans le lipide, pour le système de référence (noir), meilleur PMF après ajustement lors de la modification des valeurs ε seulement (turquoise), et résultat final (rouge) 128

List of tables

4.1 Nanotube adsorption times obtained in 17 different simulations with and without hydrodynamics. 87

5.0 Table with final coupling parameters for all the six remapped amino acids with POPC beads 102

5.1 Table assigning the side-chain beadtype for each amino acid that has only one side-chain bead in Dry Martini. The asterisk marks the special case of amino acids with no side-chain bead that have a particular beadtype assigned to their backbone bead. In all the other cases the default backbone bead is P5. 112

5.2 Table with all the already existing cross-interactions between POPC lipid beads (C1, C3, Na, Q0, Qa) and all the different beadtypes associated to the one-beaded amino acids. 112

Chapter 1

Introduction

In this chapter the reader will be introduced to the main biological components from this thesis. First we give a brief description of biological membranes, their structures and composition, and their interactions with peptides and membrane proteins. In the final part we will focus on two specific membrane-protein types that we have studied in this thesis, the GTPase and a subgroup of them: dynamins.

1.1 Biological membranes

Many cellular processes occur in a highly compartmentalised manner to ensure a successful function. In eukaryotic cells this partitioning is assured by the presence of organelles that create separated environments. This spacial separation is physically assured by lipid membranes, which makes them a key component for life. On average in a *E. coli* cell, lipid molecules account for about 9% of the total dry cell mass [169] being the third largest component after Protein and RNA. Moreover, the estimate of the number of lipid molecules in an average *E. coli* cell is about 20 million, being the most represented molecule in the cell. However, cell membranes are not only formed by lipids but by a high amount of proteins and some minor amount of carbohydrates. According to the cell type, the percentage of protein in a cell membrane composition can vary from as little as 18% (in myelin) up to 76% (in the inner mitochondrial membrane) [104].

1.1.1 Eukaryotic cell membrane composition

Eukaryotic cells membranes are composed by glycerophospholipids. The most common ones are phosphatidylcholine (POPC), phosphatidylethanolamine (POPE), phosphatidylserine (POPS), phosphatidylinositol (PIP) and phosphatidic acid (PA). In all of these lipids the

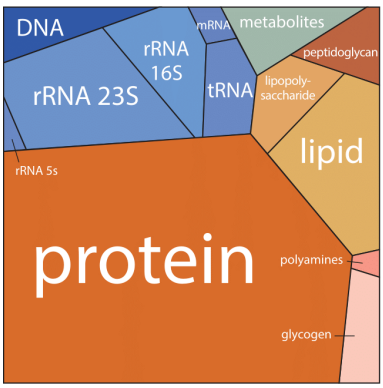

Fig. 1.1 Proportional digram of a typical cell composition by type of molecules *Taken from [169].*

hydrophobic core is a diacylglycerol (DAG), which contains saturated or unsaturated fatty acyl chains of different lengths. POPC is the most frequent phospholipid in this type of cell, representing roughly 50% of the total lipids. It organises spontaneously into a planar bilayer where the hydrophobic tails are in contact which each other from the opposite monolayers. Depending on the specific composition the mechanical properties of a membranes, and the preferential structure of the self-assembling, can vary. For example the presence of POPE add stress to the assembled bilayer, and its presence is correlated to membranes fusion and fission processes. On the contrary other lipid, like glycosphingolypids (GSLs) induce rigidity to the membrane environment.

1.1.2 Mitochondrial lipid composition

Since in this thesis we have worked with mitochondrial membrane proteins, it is worth to present this particular membrane's lipid composition. The mitochondrion has a particular membrane organisation respect to other organelles, in that that it is delimited by two membranes: the outer (OMM) and the inner mitochondrial membrane (IMM). Their lipid composition is the same but the lipid relative composition is different (see Table in figure 1.2). The main component is POPC followed by POPE and phosphatidylinositol (PI). Cardiolipin (CL) even if it is represented in small amounts, is a characteristic lipid from mitochondria and has been related to numerous functions such as protein translocation to the mitochondria [96, 139] and conservation of the mitochondrial morphology (intervening in fusion[13] and regulating fission[43]). It is particularly important when located in the inner mitochondrial membrane.

Lipid	Mitochondrion (%)	OMM (%)	IMM (%)
POPC	40	46	38
POPE	27	33	24
PI	15	10	16
POPS	3	1	4
CL	13	6	16
PA	2	4	2

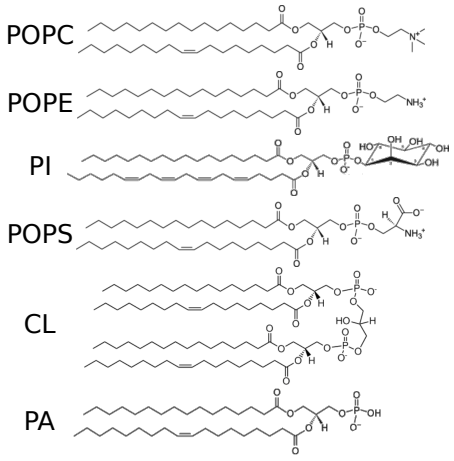

Fig. 1.2 Lipid composition of a mitochondrion belonging to yeast with its associated 2D representation(*Saccharomyces cerevisiae*). The values are reported for both membranes together and for the inner and outer membrane separately. Values taken from [115]

1.1.3 Structural overview of a cell membrane bilayer

A membrane bilayer is stabilised by hydrophobic interactions that keep the amphipathic lipids together. A planar bilayer is formed when the hydrophobic tails of the lipids (usually glycerophospholipids) are in contact with each other from the opposite monolayers.

Along the years, several models have been proposed to describe the membrane environment. For instance, in the fluid mosaic model, cell membranes were modeled as a mainly fluid lipid bilayer in which proteins are embedded [232]. This model also derived from experimental results of proteins extracted from various membranes that led to the key distinction between integral and peripheral membrane proteins. In figure 1.3 we show two representations of such a bilayer model, where the integral and peripheral proteins can be seen.

Overall, the main concept to keep in mind is that membranes behave like a two-dimensional fluid of oriented lipids with proteins, where the lateral diffusion is faster compared to the transition of a lipid molecule between leaflets (also known as “flip-flop” movement) [232]. This membrane fluidity is controlled by the fatty acid composition, temperature and cholesterol content.

1.2 Membrane proteins and peptides

As we said before, proteins are a large component of a cell membrane being key players in diverse functionalities such as recognition, signalling, transport or structural properties

just to mention a few examples. From a purely structural point of view, membrane proteins can be classified in peripheral or integral. Peripheral proteins interact only in the outside of the membrane and not with the inner hydrophobic core, whereas integral proteins have hydrophobic interactions. A particular case of an integral protein is a transmembrane protein, where the protein spans the whole membrane at least once. In nature, we can also find many transmembrane proteins that span the membrane more than once like the G-protein-coupled

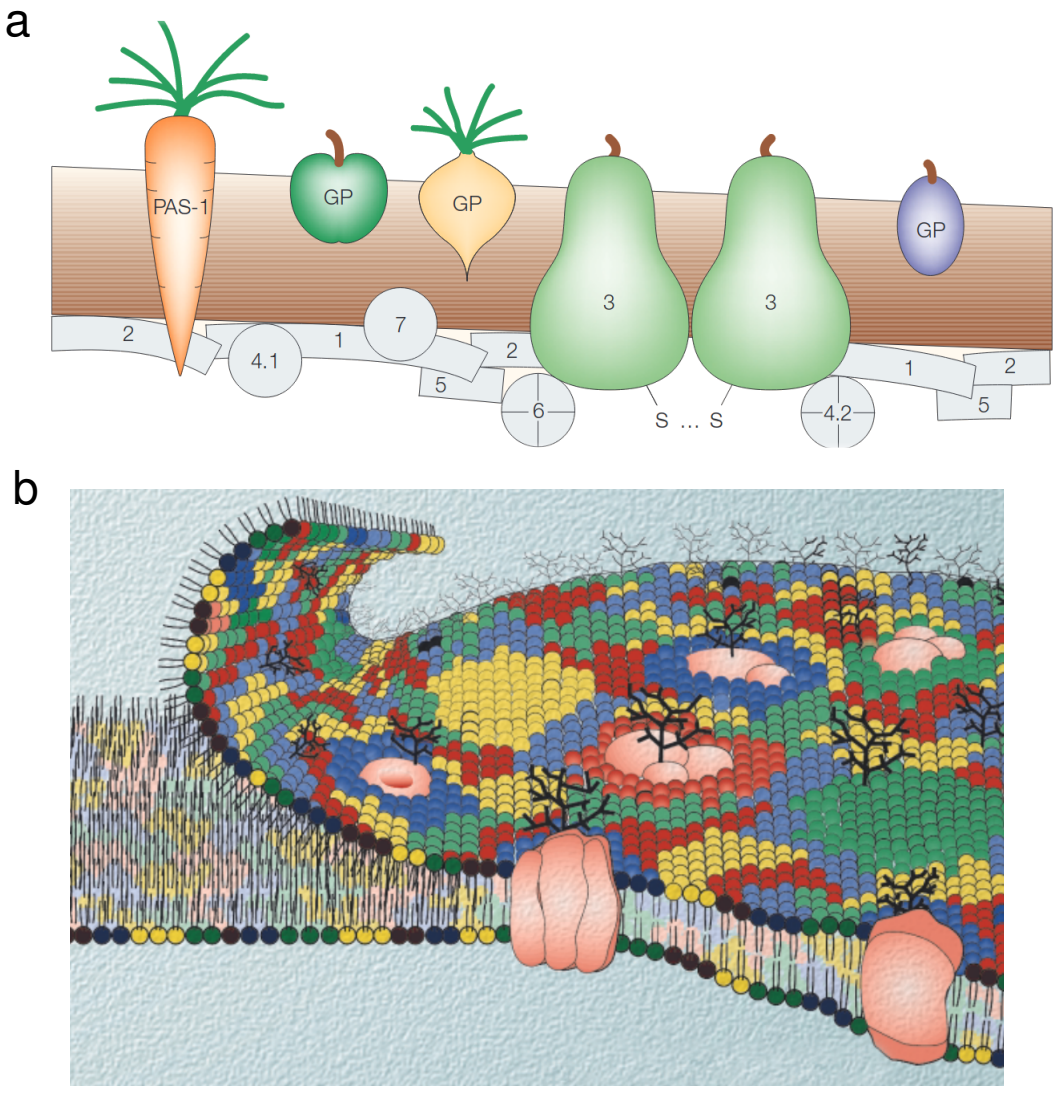

Fig. 1.3 Fluid mosaic membrane representations.(a) Historical representation of possible membrane protein insertion for the major erythrocyte membrane proteins. The carbohydrates of the glycoproteins are shown as green rods exposed to the extracellular space[238]. (b) More detailed representation of a fluid mosaic model colouring the lipids according to their type. Different regions can be appreciated. *Adapted from [75]*

receptors (GPCRs), membrane proteins with seven-transmembrane helices or ion channels. Nonetheless, there is also another type of transmembrane protein which is called monotopic and are embedded only in one layer of the membrane, interacting with the hydrophobic core but not spanning the totality of the bilayer membrane.

The same principle applies to membrane peptides, where they could be separated into two categories: those that interact on the surface of the membrane only and those that are embedded in the membrane. In the last years a great effort has been spent to study β-amyloid peptides, in the context of Alzheimer's disease. These peptides have been shown to adhere to the membrane surface and start to oligomerise into harmful structures that disrupt the membrane[263]. Another interesting point of study are the antimicrobial peptides (AMPs), or the membrane-disrupting peptides in general, that destabilise the membrane and in some cases even create a pore after inserting many protein molecules that span the membrane [217].

On a higher organisational level, membrane proteins are critical in the creation of functional multi-complexes and scaffolding systems in general. They can be related to different function from signalling pathways[205] to respiration [247], membrane fusion [108] or membrane transport [18] and vesicle coating [16], just to mention a few. Here, in figure 1.4 we show an example of a dynamin-like protein supramolecular organisation. We depict the structure solved by cryo-electro tomography for the case of helical assemblies generated by Mgm1 protein from fungi (required for inner mitochondrial membrane fusion)[79].

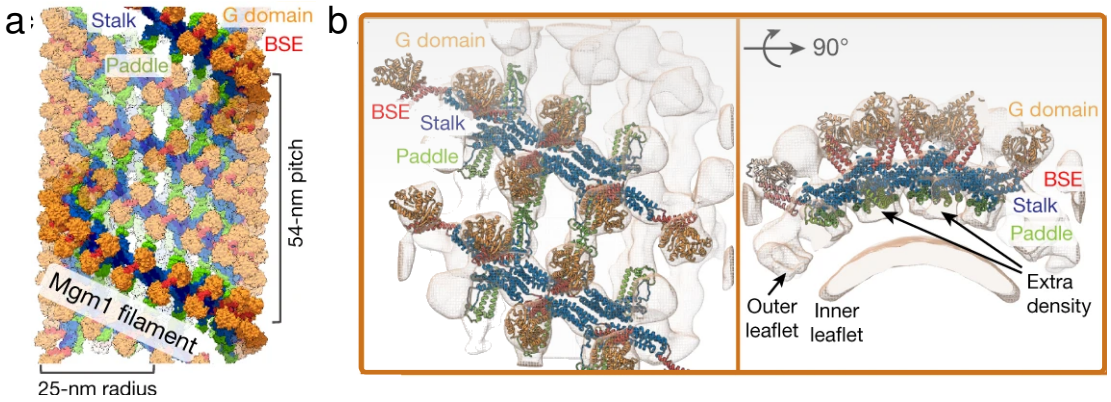

Fig. 1.4 DLP Mgm1 forms a regular protein coat on the outer surface of lipid tubes. (a) Four filaments of Mgm1 dimers wrap around a membrane tube in a left-handed surface lattice (b) cryo-electro tomography showing the protein lattice on the membrane surface. *Adapted from [79]*

1.2.1 Lipid-protein interactions

Different types of binding modes can be distinguished for lipid interactions with membrane proteins. They can first be divided into integral lipid binding or not integral. The integral protein lipids, are those that bind into a particular place of the protein in an unusual position regarding to their typical membrane localisation e.g: non-perpendicular to the bilayer plane or with the headgroup below the membrane plane. They could be involved in processes like folding or assembly of membrane proteins. The rest of non-integral lipids that interact with the protein, can be categorised by their relative position regarding to the protein. They can surround the protein resembling a bilayer structure (annular lipids) or be immersed in cavities or in an inter-protein surface typically seen in oligomeric structures.

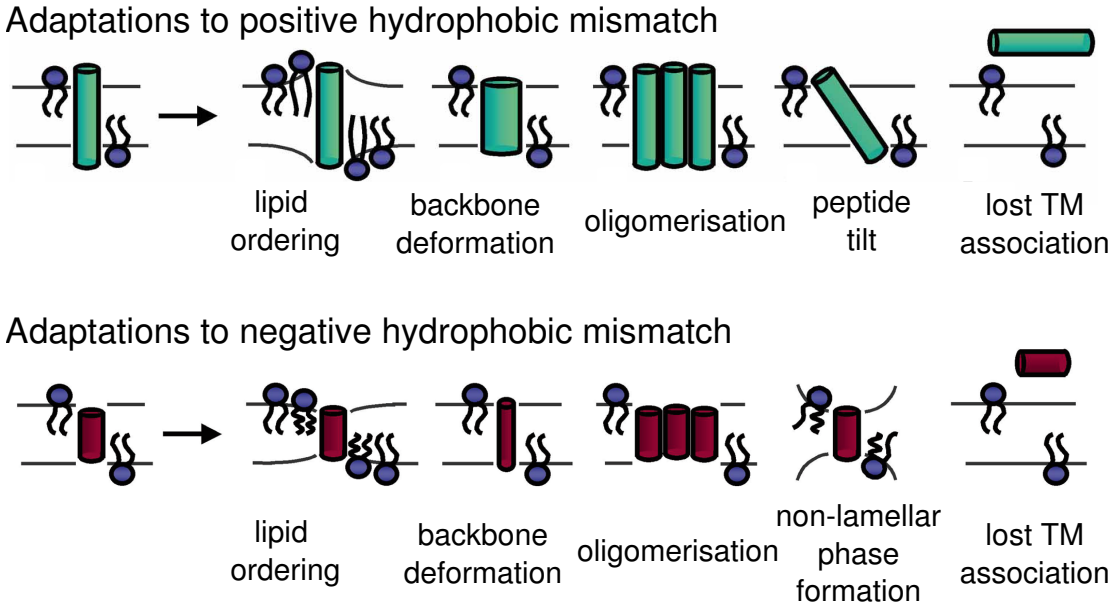

Fig. 1.5 Typical adaptations to counteract the effect of either positive or negative hydrophobic mismatch of a transmembrane (TM) α-helix. To be noted, that some modifications affect the membrane (change of the lipid ordering, disruption of bilayer character) whereas the majority affect the peptide's structure or its relative positioning towards other peptides or in the bilayer. *Adapted from [68]*

Transmembrane protein segments usually have a secondary structure that corresponds to α-helix or β-barrels in order to reduce the polar backbone interactions with the inner hydrophobic membrane core. In the case of helical transmembrane segments, they typically consist of a sequence of approximately 20 residues that are in contact with the lipid acyl chains of the bilayer. When the transmembrane protein region spans the hydrophobic core perfectly in a stable manner, it is said that there is hydrophobic match. Since it is

energetically more favourable for a system to match the hydrophobic length of the peptide to the hydrophobic thickness of the bilayer, there is a tendency to avoid hydrophobic mismatch that could lead to membrane disruptions[173]. Indeed, on that note, it has been reported that membrane thickness can alter protein activity, micro-domain formation and protein arrangements [74, 127, 174].

1.3 GTPases

Now, that we have discussed the main aspects of protein-lipid interactions, it is worth to spend some time introducing a specific type of membrane protein that has been studied in this thesis. In Chapter 3 we have addressed the role of Fzo1 (a yeast membrane protein located in the outer mitochondrial membrane) in mitochondrial membrane fusion. This protein belongs to the GTPase superfamily, and is in particular a dynamin-like protein.

The superfamily of GTPases, corresponds to a group of proteins that act as molecular switches by binding and hydrolysing GTP. They have a conserved structural design and intervene in a plethora of functions such as regulation of ribosomal protein synthesis, transmembrane signalling by hormones and light, translocation of nascent proteins into the endoplasmic reticulum, controlling differentiation and cell proliferation and guiding vesicular traffic within cells. Acquired knowledge of individual GTPases—crystal structures, biochemical properties, or results of molecular genetic experiments initially supported and generated hypotheses relating structure to function in other members of the diverse family of GTPases, that have been proved useful later to advance the study of GTPases [33].

From a catalytic point of view, the GTPase cycle includes three conformational states of the protein, defined by the stage of the GTP hydrolysis. In the cytoplasm, GTP is more likely to be bound in the nucleotide-binding pocket than GDP, and upon binding of GTP the protein takes its *active* state. This state can revert back to *inactive* when GTP is hydrolysed to GDP but is still bound to the protein.

1.3.1 Dynamins and Dynamin-like-proteins (DLPs)

Dynamin-like proteins are a particular type of large mechano-chemical GTPase, that are involved in membrane-related processes and owe their name because of the similarity with the protein "dynamin" that was discovered as a microtubule-associated protein. These proteins are involved in processes that include membrane fusion, fission, membrane protection and membrane stabilisation, in cellular processes such as endocytosis and organelle division.

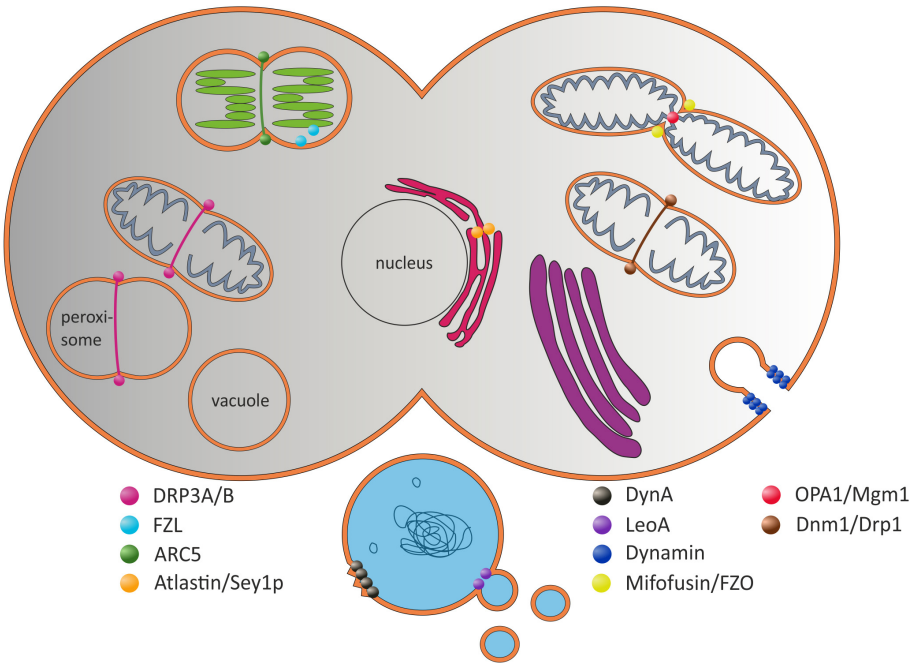

Fig. 1.6 Example DLPs localisation in a cell *Taken from [124].*

DLPs use the energy of GTP hydrolysis to perform a mechanical work that is used to perform membrane remodelling. Other members of this superfamily have been identified with time. Some popular examples include: Dynamin-1-like protein (DNM1L, or also called Drp1), that is involved in scission of mitochondria; OPA1 (Optic atrophy type 1) or Mgm1p in yeast (mitochondrial genome maintenance protein 1), that mediate the fusion of the internal mitochondrial membrane and are involved in cristae formation; and mitofusins (Mfn1, Mfn2) or fuzzy onions (Fzo1) are involved in the fusion of the outer mitochondrial membrane. DLPs have also been found in bacteria, where they are proposed to mediate bacterial remodelling events, such as thylakoid divisions. All in all, these proteins have two main characteristics: (1) their conserved structural features and (2) the mechanical effect after hydrolysis of bound GTP (i.e. hydrolysis-driven conformational change).

DLPs have a global characteristic domain organisation with 3 main domains: GTPase domain, a middle domain and a GTP effector domain (GED). A structural comparison between human dynamin and the bacterial dynamin-like protein (BDLP) showed that the middle domain and GED are part of a helical bundle (stalk) in bacterial dynamin and in both cases they are connected to the GTPase domain via a linker (BSE, bundle signaling element)[152]. The GTPase domain binds GTP with low affinity and hydrolyses it without the need of guanine exchange factor (GEF) which differentiates them from most G proteins. [72, 188, 259]. It has internally 5 different conserved motifs (called G1-G5) with associated

consensus sequences that have been reported from the dynamin superfamily. The G1 motif (or P-loop), wraps around the β-phosphate of the GTP and intervenes in binding a Mg^{2+} ion by its conserved terminal Ser/Thr [49, 216]. In the G2 and G3 motives, there is also a conserved Thr and Asp respectively that mediate Mg^{2+} binding. It has been reported that in human, the dimerisation of the G-domain and GTP hydrolysis induce conformational changes in G2 and G3 and are thus refered sometimes as a switch region [49]. Regarding G4 an Asp is reported to mediate specific binding to the nucleotide base. G5 is not well conserved between different dynamin families and interacts in various ways either with the guanine base and/or ribose.

The large GTPase domain (LG domain) is frequently followed by helical bundles that may include elements for the N-terminal region of the GTPase domain. An example would be the bundle signaling element (BSE) of dynamin (see figure 1.7). The second helical element is called *stalk* and it was shown to mediate the assembly of dynamin, MxA and DNM1l into a helical filament via different interfaces[80]. The stalk has different regulatory interaction sites for the BSE and for the PH domain, which brings the idea that alternative functions can be addressed to this interaction site for different members of the superfamily. In many dynamin superfamily proteins the site that interacts with the membrane is located at the tip of the helical domain, as is the case of the PH domain in dynamin that binds to phosphatidyl-inositol phosphates contained in the membrane, or the *paddle* region in BDLP that inserts into the membrane (pink regions shown in figure 1.7).

1.3.2 The role of DLPs in mitochondrial membrane fusion

Mitochondria are central vital organelles widely involved in many important cellular functions and are the key to the energy source in the cell. They have a characteristic double membrane which is inherently connected to their ultrastructural organization. Our current understanding about their morphology is beyond the initial beliefs that depicted them as simple rod-shaped bodies, they are integrally conected and highly dynamic during cell life. Mitochondria continuously move on cytoskeletal tracks, divide and fuse, forming shapes that range from isolated corpuscles to filamentous networks [38, 48, 58, 252]. This precisely controlled process is known as mitochondrial dynamics for which they evolved a dedicated set of protein machineries [140, 148, 233]. When particular cell locations are in need of high ATP production, mitochondria are fragmented and mobilized to those sites [140, 148, 233]. Similarly, an efficient organellar fragmentation also ensures propagation of mitochondrial genome from mother to daughter cell [5, 122, 249]. Isolation of damaged or senescent mitochondria from the network followed by degradation through a specific form of autophagy (i.e. mitophagy), ensures proper quality control of the mitochondrial network [99, 126]. Neverthe-

less, too much fission leads to mitochondrial heterogeneity and dysfunction in yeast as well as in mammalian cells [52, 166]. Because of that, a concerted fusion activity counterbalances potential excessive fragmentation. Healthy mitochondria fuse back to the existing network which homogenizes their proteomic content. Fusion also contributes to face the gradual accumulation of mutations in the nucleoid (i.e. mtDNA) through genetic complementation [12, 156, 183]. Therefore, fusion and fission processes are critical steps of cellular quality

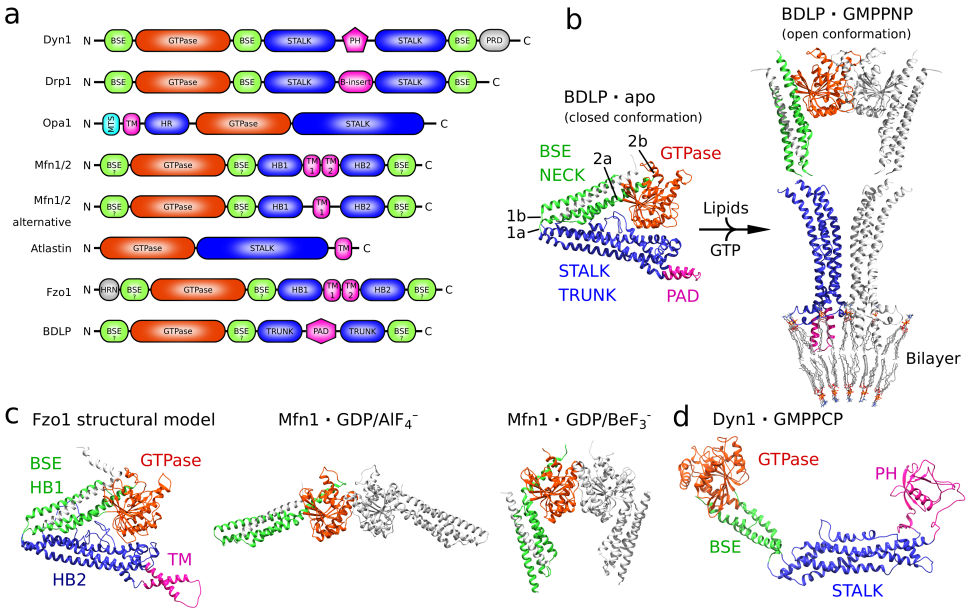

Fig. 1.7 Schematic representation of the modular domain architecture of selected proteins from the dynamin superfamily. (a) Cartoon summarizing the modular domain architecture for selected DLPs. The classical nomenclature is indicated. The helical bundles HB1 and HB2 of the mitofusins/Fzo1 are aligned to the stalk/neck domains of other DLPs and dynamins showing a similar architecture and they might serve as BSE (Bundle Signaling Element). MTS: Mitochondrial Targeting Sequence, TM:transmembrane domain, PAD: paddle, PH: Pleckstrin Homology, PRD: Proline-Rich Domain. (b) The BDLP conformational change upon GTP and lipid binding. Hinges 1a, 1b, 2a and 2b are indicated. The structure is closed in its GDP-bound state (not shown) and in its apo-form (left) (PDB: 2J68 [152]). Lipid binding concomitant with GTP hydrolysis, reveals a conformational change detected using cryo-electron microscopy (right) (PDB: 2W6D [153]). (c) (Left) The Fzo1 structural model used in this thesis [70] with HB1 and HB2. (Center) Crystal structure of the human Mfn1 fragment-construct obtained upon addition of GDP/AlF4- (not shown) (PDB: 5GOM [46]) and observed in presence of GDP/BeF3- (not shown) (PDB: 5YEW [271] (Right). (d) Cryo-electron microscopy of the human dynamin 1 (Dyn1) in complex with GMPPCP (not shown) (PDB: 6DLU [134]). The PRD was not present in the reconstruction. *Figure taken from [69].*

control through which mitochondria maintain their homeostasis [123, 149]. Perturbations of this delicate equilibrium are associated with several diseases which are mostly neurological disorders [27]. This is due to the fact that neuronal cells are energetically demanding and thus particularly sensitive to mitochondrial dynamics [131]. Therefore, understanding the mechanistic details of mitochondrial dynamics is physiologically and medically relevant. The main factors mediating mitochondrial dynamics are multi-domain mechano-chemical GTPases that belong to the dynamin superfamily [66, 81, 197]. These dynamin-like proteins, as previously discussed are mostly implicated in membrane-remodeling processes [66, 203].

Human Drp1 (Dnm1 in yeast) mediates the scission of mitochondria [136, 234]. The role of Drp1 appears to be highly conserved in eukaryotes [4, 141, 151, 219, 234] and cells lacking Drp1 (or Dnm1) contain highly interconnected mitochondria networks [187, 235]. Drp1 is recruited to the mitochondrial outer membrane (OM) by protein adaptors: the mitochondria fission factor (Mff) and the two mitochondrial dynamics proteins MiD49 and MiD51 [186, 189]. Investigation into individual and combined loss of each adapter showed a combined prominent role of MiD51/Mff [185]. In yeast, Dnm1 localization relies on the integral OM protein Fis1 and two adaptors Mdv1 and Caf4 [102, 175, 182, 251] which recruit Dnm1 to mitochondria. Fis1 is conserved in human [73] but its function evolutionary diverged to act as an inhibitor of the mitochondrial fusion machinery [272]. After Drp1/Dnm1 is recruited to the mitochondrial OM [219, 234] its innate capacity to assemble into multimeric complexes, leads to membrane constriction upon GTP hydrolysis [94, 165]. Rings and spiral-like structures of Drp1 assemble around the organelle at the region of contact with the endoplasmic reticulum (ER), which is a requisite for the fission event [88]. ER tubules wrap around the organelle to decrease its average diameter from 300-500 nm to 150 nm [88]. Another Dynamin, Dyn2, was suggested to collaborate with Drp1 [145], but the involvement of this second DRP in mitochondrial division is controversial [83].

OPA1(optic atrophy 1) and its yeast homologue Mgm1 (mitochondrial genome maintenance 1) mediate the fusion of the mitochondrial inner membrane (IM) and participate in cristae organization [87, 268]. The fusion of the mitochondrial OM is triggered by the mitofusins Mfn1 and Mfn2 in mammals and by the fuzzy onion Fzo1 protein in yeast [58, 107]. Notably, the phylogenetically related group of atlastin (and its yeast homologue Sey1p), functions in fusion of ER tubules [116, 184]. Several structural observations on recombinant atlastins and Sey1 have provided deep insights into fusion of ER tubules [128, 146, 270]. In contrast, full-length mitofusins have not been purified which has limited the study of their assembly and oligomerization, which is an essential step in the OM fusion process [37, 93, 120, 121, 135, 204, 214].

Current developments in molecular modeling have made it possible to model and study homologous proteins whose structural details have not been elucidated yet. Moreover, advances in molecular dynamics simulations allowed to study membrane proteins in a native-like lipid environment and overcome current limitations in their isolation and purification. Clear examples are widely discussed in recent review works [62, 90, 158]. The involvement of mitofusins in mitochondrial fusion was also recently reviewed [58]. In Chapter 3 we will present recent observations on mitofusin oligomerisation and dynamics in their native membrane environment based on computational approaches such as molecular modeling and molecular dynamics (MD) simulation.

1.4 Molecular dynamics simulations of membrane protein systems

Since the main scope of this thesis is to create computational tools to model and study membrane protein systems, it is relevant to introduce the reader to what has already been done in this area. A particular emphasis will be put on other systems treated via coarse-grained molecular dynamics simulations. This type of simulations has been carried out to identify lipid-binding sites on membrane proteins [218, 223], study scaffolding on nanodisks [213], membrane fusion [206, 236], protein-mediated vesicle fusion simulations [14], signalling [31], study clustering of multiple proteins (such as e.g. GPCRs) in the context of a membrane [50], simulation of complex plasma membrane model for ten different membrane proteins [61], just to mention a few of the latest examples.

Molecular dynamics simulations of biological bilayers were run for the first time in the 1980s by Berendsen and coworkers [255] whereas the first explicit simulations of peptides embedded in a membrane were done by Roux and coworkers during the 1990s [211]. One of the most widely studied type of membrane proteins are membrane channels, proteins that facilitate permeation of different types of chemical compounds according to their electrochemical gradient across the membrane. These types of systems have been simulated for many decades and in particular computational methods have been shown helpful to study the binding of ligands and the allosteric effects in the channel as well as studying the binding free energy [150, 181, 229, 231, 276]. Also these simulation approaches allowed to study the conductivity, opening and gating mechanisms that could be compared to experimental macroscopic data and give an insight into the structural aspect of these interactions.

Another example of biological system that profit from the computational approach is Photosystem II (PSII) that was studied by coarse-grained molecular dynamics simulations.

PSII is a protein complex located in thylakoid membranes and intervenes in photosynthesis) The PSII dimer was embedded in a thylakoid-like membrane and run for a full microsecond telling that there is flexibility at the dimer interface [257]

1.4.1 Membrane protein structures databases: The structures we have so far

Membrane protein structure is hard to solve experimentally because of the difficulty to find a proper way to obtain a crystal. For that reason it is not surprising that they only represent 0.35% of the total PDB entries at the moment of the writing of this thesis (only 546 structures were annotated as membrane proteins out of 154478). However, if we compare that value to the number of experimentally resolved structures that are reported in a specific Membrane Protein Database (mpstruc) currently they have only 949 unique proteins (non-redundant).

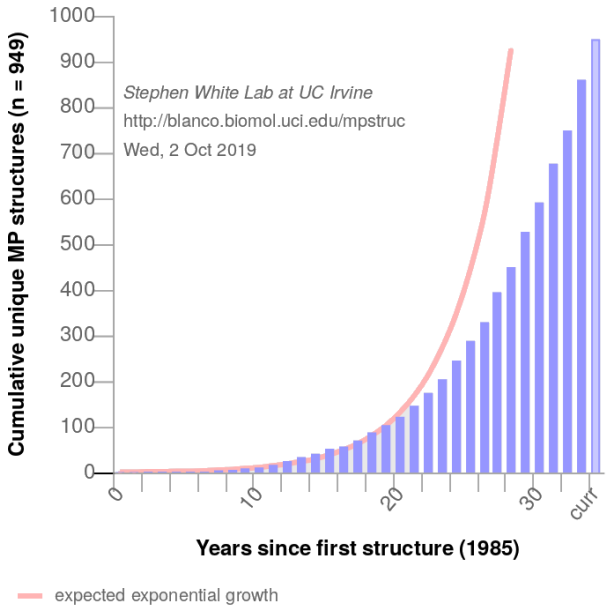

Fig. 1.8 Growth in membrane protein entries as a function of the year of the mpstruc database.

Nonetheless, these structures are missing a critical part of their environment that is the bilayer. For that reason, other databases were created that report structures with the bilayers. One of the most popular ones is OPM (***O**rientations of **P**roteins in **M**embranes database*), that includes all structures of transmembrane protein complexes and selected monotopic, peripheral proteins and membrane-bound peptides from the PDB (currently a total of 4572 distinct membrane proteins or peptides). This database also provides a link to a webserver (PPM $https://opm.phar.umich.edu/ppm_server$) so that the user can upload a structure

and it will be processed in the same way as all the entries in their database. Even more interestingly, there is another database that specialises in running coarse-grained simulations and letting the bilayer form spontaneously around the protein. It uses as input, structures from the PDB that are systematically treated to run with coarse-grained molecular dynamics simulations. They also provide a structural analysis with each entry, that is coming from the last section of their molecular dynamics production run as well as the protocols and the final coordinates in either all atom resolution or coarse-grained. Currently they have 3500 structures available. More information about the fundaments of coarse-grained models and simulations in general can be found in the following chapter. Given the difficulty of access to experimentally defined membrane protein structures, it can be straightforward to rationalise the collaboration that can come from computational approaches.

1.5 Main objectives

This thesis has two main purposes:

- Extend the work on understanding mitochondrial membrane fusion, in particular at the tethering stage, from a structural and dynamic point of view.
- Advance the methodology for studying lipid-protein systems via molecular dynamics simulation. In particular with a multi-scale approach where the lipid and protein models are treated with relative high-resolution whereas the solvent is treated via a mesoscopic approach with a Lattice-Boltzmann technique.

For the study of mitochondrial fusion, the main objective is to propose and study the different oligomerisation states of Fzo1 in a membrane context (Chapter 3).

Concerning the methodological development, the goal of our effort was two-fold. First, extend the operative capability of a well known and effective model for lipids, the Dry Martini coarse-grained model, by coupling it to hydrodynamic interactions. This coupling could enable the study of biological processes where the perturbative action of a surrounding fluid could play a major role in membrane-related processes (e.g. fusion, protein adsorption, shear induced deformation). The second one, was to couple the Dry Martini lipid model with the OPEP coarse-grained force field for proteins that is developed in the laboratory. The success of this coupling would enable the study of processes where secondary structure changes of proteins play a major role. This is not possible for instance with the present protein models developed for the Dry Martini and Martini lipids force fields.

Chapter 2

Theoretical background and methods

2.1 Molecular Dynamics simulations

2.1.1 Fundamentation and examples

Since the 1960s Molecular Dynamics (MD) simulations have been used and have grown until becoming a fairly common technique amongst computational biophysicists. Nowadays, the synergy that exists between experimental information and simulations opens up the possibility to study and deepen the knowledge of different phenomena, in particular those requiring a structural detail at atomistic or molecular level. In biological systems like, e.g peptide or proteins the scales of interest involve from hundreds of atoms to thousands and millions (for full protein in solvation or in complexes). Moreover, the timescales that can currently be reached and are of interest for these systems are comprised between picoseconds (to study e.g solvation effects) up to microseconds (e.g. to study protein conformational changes). Already, there have been advances in the conjoint work between experimental information and simulations that yielded structural and dynamical information on the molecules under study. A renowned example would be the study of DNA strain separation with atomistic detail via MD [193]. In that case, MD allowed to obtain detailed structural information about a process that could not be assessed experimentally at the time. Other very popular systems that have been studied via computational approaches are membrane protein systems, such as ion channels, molecular motors or proteins embedded in the membranes that undergo conformational changes [105, 181, 254]. Several examples of those systems can be mentioned, such as the ATP synthase (also known as F_0-F_1-ATP synthase), where structural effects in the complex were analysed as a consequence of rotational motions of subunits [125]. A more recent example is the dynamical study of the complete HIV capsid from structures obtained via cryo-electronmicroscopy [275]. This system, formed by approximately 64 million atoms

allowed to study how the way capsid proteins assembled affects the curvature of the capsid, adding knowledge to the mechanism in viral assembly. All of these examples show the power of computational strategies to complement the knowledge that can be obtained by experimental methodologies, and moreover, help understand biological system assemblies and behaviour at an atomistic or near-atomistic level. In the following sections we will provide a basic description of the molecular dynamics simulation technique.

2.1.2 Generalities

Biologically relevant systems (usually macromolecules) consist of several thousand atoms and the temporal evolution of their positions is also of high importance, for these reasons purely classical approaches are used to model existing interactions. The classical treatment is implemented as a consequence of some simplifications. The first of them is the Born-Oppenheimer approximation. This approximation allows to dissociate the motion of nuclei and electrons since the nucleic mass is much higher than the electronic one and the energies associated to them is in a clearly distinct scale. Because of this relationship, variations in the nuclear positions occur at much slower scales than the electronic ones, giving the electrons the ability to practically instantaneously adapt to the nuclear motion. The fact that it is formally correct to separate the movement description of electrons and nuclei determines that the behaviour of nuclei can be treated in a classical manner (i.e. they follow the laws of classical physics). This gives rise to the second vital approximation; the use of force fields to obtain the energies in classical mechanics. These force fields could be empiric or with information derived from quantum mechanics where electronic effects are explicitly considered.

2.1.3 Force Fields

In computational chemistry a force field is a set of functions and parameters used to obtain the potential energy of a system. These parameters can be obtained by quantum mechanics calculations or based on experimental information (derived from spectra analysis for example). Generally, a set of parameters is obtained from data regarding small molecules or functional groups that are later transferred to molecules, given thus the transferability property of the force fields. Even if there are different ways to model the potential energy of an atomistic system, a minimal force field implementation that describes the interatomic interactions has converged over the years. The modelled interactions as well as the typical mathematical expression for their potential energy are described in the table of figure 2.1.

Interaction type	Expression for the potential energy
Bond length	$\sum_{bond} = \frac{1}{2}k_r(r-r_0)^2$
Angle (Angle bending)	$\sum_{angle} = \frac{1}{2}k_\theta(\theta-\theta_0)^2$
Dihedral angle (torsion)	$\sum_{torsion} = \frac{U_n}{2}(1+cos(n\psi-\delta))$
Electrostatic interaction (Coulomb)	$\sum_{Coulomb} = k_{el}\frac{q_iq_j}{r_{ij}}; k_{el}\frac{1}{4\pi\varepsilon}$
vdW interaction: LJ potential	$\sum_{LJ} = (\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}); A = 4\varepsilon\sigma^{12}; B = 4\varepsilon\sigma^{6}$
Total potential energy: $E_{pot} = \sum_{bond} + \sum_{angle} + \sum_{torsion} + \sum_{Coulomb} + \sum_{LJ}$	

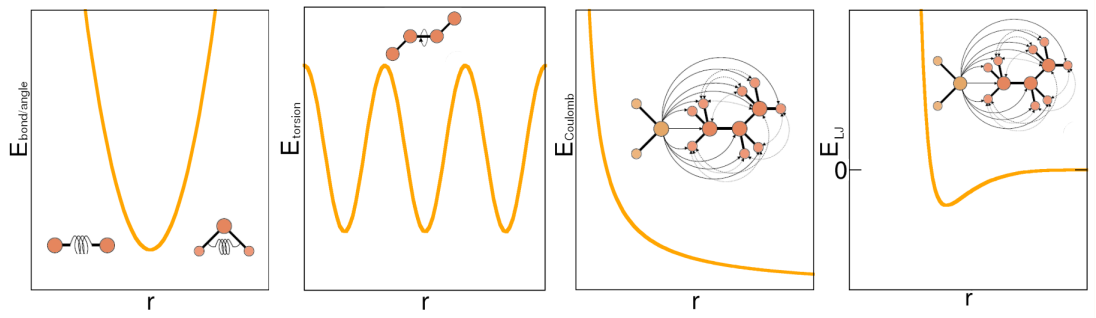

Fig. 2.1 Table with all the interaction types and potential energy expressions of a typical force field. The potentials are plotted below to show the different shapes.

These terms are the minimal expressions in any standard atomistic force field, they can be divided into two main groups: bonded and non-bonded interactions. The first group is composed by all of the interactions that involved covalently bonded atoms. There are different expressions that describe the energies coming from bond vibrations (2 atoms involved), between the angle variation (3 atoms involved) and the variation of the torsional angle (4 atoms involved). It is important to keep in mind, that some variation of this minimal set of potential expressions can be found depending on the implementation of a particular force field. For example the form of a potential could vary or extra terms could be added, as for instance a term for hydrogen bonding, for accounting for solvation effects or metal-ligand interactions are relatively common.

Bonded interactions: "bond stretching":

This term comes from the energy variation as a consequence of the variation in the bond length (stretching or shrinking) with respect to a reference value. This reference value is the value adopted when all the other energy terms in the force field are zero and does not necessarily correspond to the equilibrium bond length. In molecular mechanics, it is frequently modelled via an harmonic potential between two masses (atoms) as shown in the table of figure 2.1. In this expression of a Hooke-type harmonic potential k_{el} is the elastic constant spring, r the distance between the atoms that take part in the bond, and r_0 the equilibrium distance between both atoms.

Bonded interactions: "angle bending"

This potential is frequently also expressed mathematically as a Hooke-type harmonic potential, but in this case the potential is applied between 2 atoms that are covalently connected via a third atom. As shown in table in the table of figure 2.1, k_θ is the elastic constant, θ is the angle defined by the three atoms and θ_0 the equilibrium angle of the three atoms.

Bonded interactions: "torsional dihedrals":

This potential describes the relative orientation of two atoms that are separated by three covalent bonds. As seen in the table of figure 2.1, this is done by a cosine series where E_n is the height of the barrier, ψ is the dihedral angle defined by the 4 atoms, n the multiplicity and δ the phase. The multiplicity allows to correctly represent the different energy minima related to the the staggering or eclipsing orientation of the atoms i, i+4.

Non-bonded interactions: van der Waals interactions:

The van der Waals interactions are one of the main types of non-bonded interactions that stabilise the structure of molecules. They are calculated in a pairwise manner and include interactions of type dipole-dipole, dipole-induced dipole and London dispersion (instantaneous induced dipole- induced dipole). A very common way to express the energy potential of them is via a Lennard-Jones potential, where it is possible to consider both characteristics of the interaction: the repulsive and the attractive part. At short distance they are repulsive, and attractive at longer ranges. For instance, in the table of figure 2.1, the term r^{-12} models the Pauli repulsion whereas r^{-6} models the attractions coming from van der Waals (vdW) forces. The key parameters to model this interaction are ε, which is the depth of the well from the potential and σ, which is the interatomic separation such that the potential energy

is zero. Usually in force fields a set of ε and σ parameters are listed for each distinct atom type. However, in order to compute the interactions between different atom types different combination rules are used. One of the most widely used ones is called Lorentz-Berthelot, and it is the one used in this work. With this combination rule, the σ of an interaction is obtained as the arithmetic mean from the σ values from the two atoms involved, whereas the ε of the interaction is obtained as the geometric mean. That means that for a vdW interaction between an atom i and j, the parameters σ and ε are obtained as:

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \qquad \varepsilon_{ij} = \sqrt{\varepsilon_i \cdot \varepsilon_j} \tag{2.1}$$

It is worth to note, that these interactions are computationally expensive to treat and principally effective in a short range. Because of the main energetical contribution happens at shorter inter-particle distances, some simplifications are used that only consider these interactions up to a certain distance. This gives rise to the concept of $cut-off$, that represents the maximal distance to define a non-bonded pair-wise interaction. From a computational point of view, given a particle i, a list of neighbouring particles that are within the cut-off is generated to compute the vdW interactions.

Non-bonded interactions: "Coulomb potential":

One key component of the potential energy is the contribution due to electrostatic interactions. This type of non-bonding interaction can be modelled generally with a pairwise Coulomb potential. According to the expression shown in the table of figure 2.1, k_{el} is the Coulomb constant, q_i and q_j the charges of atoms i and j respectively and r_{ij} the distance between them. Notably, the calculation of these interactions can be extremely computationally expensive. Some techniques are used to introduce a distance limit for considering these interactions (i.e by using a cut-off), and thus reduce the computational cost of this calculation. A special treatment to avoid neglecting the long-range effect of electrostatic interactions and treat them explicitely is called Ewald summation and usually it is implemented in its later and more computationally efficient form as Particle Mesh Ewald, that will be explained in detail in the next section.

2.1.4 Technical details for a simulation setup: Periodic boundary conditions and long range electrostatics treatment

Periodic boundary conditions

Setting up the system for molecular dynamics correctly is paramount to have a good model representation. In particular one should be aware of the border artefacts that could arise when setting up a molecule in a simulation box. If no extra treatment is added one would have the system bordering with vacuum, which logically gives rise to unrealistic representation. Furthermore, in order to simulate a system that mimics the realistic environment, it is critical to have boundary conditions so that the system is not treated in isolation. In particular, periodic boundary conditions (PBC) are usually being used in simulations of biological systems. In simple terms, a system that is being treated with PBC consists of a unit cell (i.e. simulation box) that is replicated infinitely in all the dimensions in space. So as if a particle "leaves" the box from e.g the right side of the box, it will re-enter the box from the left-side again. It follows then, that, when using a cut-off simplification to optimise the computational cost, the size of the box needs to be big enough to completely contain the solute inside the box initially without it interacting with its periodic image through short range interactions. For that reason, for short-range interactions, a computational simplification arises that defines a minimum image convention so that the cut-off(R_c) used to truncate all the short-range non-bonded interactions should not exceed the shortest box vector (as shown in equation 2.2).

$$R_c < \frac{1}{2} min(\|L_x\|, \|L_y\|, \|L_z\|) \tag{2.2}$$

In particular, to avoid a single solvent molecule to interact in short-range with two different images in one box direction, the length of that box vector should be larger than the length of the solute in that direction plus two times the cut-off radius R_c.

Long-range electrostatics

Since electrostatic interactions decay as r^-1, they are particularly problematic in molecular simulations. One approach to solve this issue could be to use a large simulation cell, however that will be computationally too expensive. Nonetheless, a couple of methods have been developed to improve the treatment of electrostatics at longer ranges such as Ewald summation or Particle Mesh Ewald (PME).

The Ewald summation method was introduced in 1921 by Ewald [78] to study energetics of crystals. It defines the interactions of a particle with all the other particles in the simulation box (or unit cell) and also with their infinite periodic images.

The total electrostatic energy of N particles and their respective periodic images is given by:

$$V = \frac{1}{2}\sum_{n_x}\sum_{n_y}\sum_{n_z}\sum_{i}^{N}\sum_{j}^{N\,*}\frac{q_i q_j}{4\pi\varepsilon_0 \boldsymbol{r}_{ij,\boldsymbol{n}}} \tag{2.3}$$

where $(n_x, n_y, n_z) = \boldsymbol{n}$ is the box index vector, and the the asterisk refers to exclude the term j=i, if and only if $(n_x, n_y, n_z) = (0,0,0)$. The distance $\boldsymbol{r}_{ij,\boldsymbol{n}}$ is the real distance between the charges and the minimum image. This sum is conditionally convergent, however it converges slowly, so Ewald introduced the conversion of equation 2.3 into two faster converging terms (one in direct space and the other in the reciprocal space) and a constant term:

$$V = V_{dir} + V_{rec} + V_0 \tag{2.4}$$

$$V_{dir} = \frac{1}{2}\sum_{ij}^{N\,*}\sum_{n_x}\sum_{n_y}\sum_{n_z}\frac{q_i q_j}{4\pi\varepsilon_0}\frac{erfc(\beta r_{ij,\boldsymbol{n}})}{r_{ij,\boldsymbol{n}}} \tag{2.5}$$

$$V_{rec} = \frac{1}{2\pi V}\sum_{ij}^{N\,*}\frac{q_i q_j}{4\pi\varepsilon_0}\sum_{m_x}\sum_{m_y}\sum_{m_z}\frac{exp(-(\pi\boldsymbol{m}/\beta)^2 + 2\pi i\boldsymbol{m}\cdot(\boldsymbol{r}_i - \boldsymbol{r}_j))}{\boldsymbol{m}^2} \tag{2.6}$$

$$V_0 = -\frac{\beta}{4\pi\varepsilon_0\sqrt{\pi}}\sum_{i}^{N} q_i^2 \tag{2.7}$$

where β is a parameter that determines the relative weight of the direct and reciprocal sums and $\boldsymbol{m} = (m_x, m_y, m_z)$ is the reciprocal vector. By splitting the reciprocal and direct terms up, it is possible to use a short cut-off in the direct space and another cut-off in the reciprocal space. However since the computational cost of the reciprocal part of the sum increases as N^2 it is very computationally expensive for a large system. For this reason another approach has been developed: The Particle Mesh Ewald method.

The Particle Mesh Ewald method is inspired by the Ewald summation to improve the performance of the sumation in the reciprocal space. Here the charges are interpolated on a grid instead of directly summing wave vectors. The grid is usually Fourier transformed and the reciprocal energy term obtained by a sum over the grid in k-space. This algorithm scales as $Nlog(N)$, which shows an improvement in the performance with respect to the Ewald summation, particularly for large systems, where the setting up of the grid and transforms do not slow down the calculations.

2.1.5 Theory behind Molecular Dynamics (MD)

Molecular dynamics is a computational methodology based on numerically integrating classical equations of motions for the atoms in the system at each time step. In other words, the classical equation of movement needs to be solved. In the case of one atom it has the form of:

$$\boldsymbol{F_i} = m_i \cdot \boldsymbol{a_i} = m_i \cdot \ddot{\boldsymbol{r}}_{\boldsymbol{i}} \tag{2.8}$$

Where $\boldsymbol{F_i}$ corresponds to the acting force on an atom i as a consequence of the interactions with the rest of the atoms, m_i to the mass of atom i and $\boldsymbol{a_i}$ to the acceleration (or the second time derivative of the position $\ddot{\boldsymbol{r}}_{\boldsymbol{i}}$. The equation 2.8 is a second order differential equation that can be rewritten as two separate first order differential equations; one expressing the momentum derivative and the other one the derivative of the atomic positions:

$$\dot{\boldsymbol{p}}_{\boldsymbol{i}} = \frac{\partial p_{\boldsymbol{i}}}{\partial t} = \boldsymbol{F_i} \tag{2.9}$$

$$\ddot{\boldsymbol{r}}_{\boldsymbol{i}} = \frac{\partial r_{\boldsymbol{i}}}{\partial t} = \frac{\boldsymbol{p_i}}{m} \tag{2.10}$$

Where $\boldsymbol{p_i}$ is the momentum for an atom i ($\boldsymbol{p_i} = m_i \cdot \boldsymbol{v_i}$), $\boldsymbol{r_i}$ its position and $\boldsymbol{F_i}$ the total force that acts on atom i. F_i can also be written as the sum all its components: the forces between bonded atoms and non-bonded atoms and the external restriction forces that could eventually exist. These forces are obtained as the gradient of the potential energy obtained by the movement of atoms:

$$\boldsymbol{F_i}(\boldsymbol{r_i}, ..., \boldsymbol{r_N}) = -\nabla_{\boldsymbol{r}_i} E_{pot}(\boldsymbol{r_i}, ..., \boldsymbol{r_N}) \tag{2.11}$$

Finally, it is possible to combine 2.8 with 2.11 and obtain the following expression from the equation of motion for a system with N atoms:

$$-\nabla E_{pot}(\boldsymbol{r_i}, ..., \boldsymbol{r_N}) = m_i \frac{d^2 \boldsymbol{r_i}}{dt^2} \tag{2.12}$$

Now, the calculations are based on expressions of potential energy $E_{pot}(\boldsymbol{r})$ that are obtained through the force field of choice.

2.1.6 Algorithms in MD

There are several methods to integrate Newton's equation of motion. In all of them a sufficiently small time step has to be chosen to achieve energy conservation for an isolated

system. Because of this requirement, the movements with the highest frequency are the ones that will determine the minimal time step that can be used in a molecular dynamics simulation. In biological systems, those correspond to the vibrational motions from the bonds containing hydrogen, since hydrogen atoms have the smallest mass. This condition delimits the time step to at the most 1 fs. However, usually, these bonds are constrained (using e.g the LINCS algorithm [111]) by choice in molecular dynamics simulations so to expand the time step to 2 fs. Another simplification that can be made to speed-up the computational time is the use of multiple time steps. For instance, the time derivative of pair interactions decreases with the inter-particle distance, meaning that the fluctuations at long-range will be smaller than at short-range. It follows then that the long-range forces can be computed less frequently than the short-range ones.

In the following sections, we will describe two of the most common algorithms to propagate movement in molecular dynamics that were also used in this thesis: Verlet and leapfrog algorithms.

Verlet

The Verlet algorithm is one of the simplest to comprehend. It is based on obtaining the positions at time t, time $t - dt$ and the acceleration at time t. In the Verlet algorithm the atom positions are obtained via a Taylor expansion such that:

$$\boldsymbol{r}(t+\delta t) = \boldsymbol{r}(t) + \delta t \boldsymbol{v}(t) + \frac{1}{2}\delta t^2 a(t) + O\delta t^4 \tag{2.13}$$

$$\boldsymbol{r}(t-\delta t) = \boldsymbol{r}(t) - \delta t \boldsymbol{v}(t) + \frac{1}{2}\delta t^2 \boldsymbol{a}(t) + O\delta t^4 \tag{2.14}$$

By summing 2.13 and 2.14 we obtain the expression of the Verlet algorithm:

$$\boldsymbol{r}(t+\delta t) = 2\boldsymbol{r}(t) - \boldsymbol{r}(t-\delta t) + \delta t^2 \boldsymbol{a}(t) \tag{2.15}$$

The error of the methodology is defined by the truncation of the Taylor series and is noted by $O\delta t^4$

The acceleration is obtained directly via the evaluation of the force on a particle i, following the equation:

$$\boldsymbol{a}_i(t) = \ddot{\boldsymbol{r}}(t) = \frac{d^2\boldsymbol{r_i}}{dt^2} = \frac{\boldsymbol{f_i}(t)}{m_i} \tag{2.16}$$

This algorithm has a popular variant named Velocity-Verlet, where the accelerations, positions and velocities are computed all at the same time instead if at subsequent time steps.

Leap-frog

In the leap-frog algorithm, positions and velocities are actualised considering the acceleration in the same way as in the Verlet algorithm (2.15).

$$\boldsymbol{v}(t+\frac{\delta t}{2}) = \boldsymbol{v}(t-\frac{\delta t}{2}) + \boldsymbol{a}(t)\delta t \tag{2.17}$$

$$\boldsymbol{r}(t+\delta t) = \boldsymbol{r}(t) + \boldsymbol{v}(t+\frac{\delta t}{2})\delta t \tag{2.18}$$

When the leap-frog algorithm is implemented the velocities at time $t+\frac{t}{2}$ are calculated initially from the velocities at time $t-\frac{\delta t}{2}$ and accelerations at time t.

The name of this algorithm derives from the kid's game $leapfrog$ because the velocities are computed every half time step ($\frac{\delta t}{2}$) whereas the positions are computed every time step (δt), in an alternate fashion.

Ensembles, pressure and thermic control

A molecular dynamics simulation can be done in different thermodynamic ensembles like the canonical (NVT), microcanonical(NVE) and the isobaric-isothermal ensemble (NPT), the most relevant one for biological systems. In order to assure that the simulation allows to sample a particular ensemble, some control methods should be used.
Here, we will introduce an example for pressure and temperature coupling methods relevant to molecular dynamics. There are different methods to maintain a constant pressure, the one by Berendsen being one of the oldest that is generally used only for initial steps of an equilibration. In this method, the system is softly coupled to an external bath using the principle of minimal local perturbation. This is achieved adding an extra term into the equations of motion that makes the necessary changes to achieve constant pressure:

$$\left(\frac{\partial p}{\partial t}\right)_{bath} = \frac{p_0 - p}{\tau_p} \tag{2.19}$$

Where τ_p is the coupling time constant. Analogously, to assure that the average temperature of the system is correct, an extra term is added to the equation of motion. This can be done also with the Berendsen thermostat, where all the fluctuations in the kinetic energy are eliminated to assure that the temperature remains stable [23]. The velocities are scaled at every step of a simulation, so that the rate of temperature variation is proportional to the

difference between the temperature of the coupled bath and the system in the following way:

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_{bath} - T(t)) \tag{2.20}$$

Where τ represents the coupling constant of the bath and the system. Thus, higher values of τ relate to stronger coupling. However it is worth mentioning that the Berendsen thermostat has a major defect since it does not sample correctly the canonical ensemble, so other strategies should be used and have been developed to correct this issue such as the velocity-rescaling thermostat [42]. The velocity rescaling thermostat is essentially a Berendsen thermostat with an additional stochastic term that ensures a correct kinetic energy distribution by modifying it according to equation 2.21

$$dK = (\bar{K} - K)\frac{dt}{\tau} + 2\sqrt{\frac{\bar{K}K}{N_f}}\frac{dW}{\sqrt{\tau}} \tag{2.21}$$

where $\bar{K} = N_f/2\beta$, β is the inverse temperature and N_f the number of degrees of freedom, and dW a Wiener process (stochastic component).

2.2 Coarse-grained methods

2.2.1 Generalities

Coarse-grained models have gained popularity in the past years among computational biophysicists due to the high computational cost of simulating biologically relevant systems. This cost is mainly due to the temporal and size scales that have to be reached in order to solve relevant questions. Coarse-grained models, are one way to overcome these barriers by reducing the degrees of freedom from the system and introducing a simplified molecular representation. There are several different approaches and strategies that can be used in order to choose the best approximation to model the properties of interest correctly. The interested reader can refer to an excellent review[180]. For instance, if one wants to model internal motions of biomolecules such as a protein, the interactions could be modelled via elastic networks, where e.g. the protein is modelled as a set of beads connected via elastic springs (with its associated elastic constant and a cutoff), allowing for a change in conformation around the energy minimum of this network. This approach allows for instance to recognise structural changes in dynamics where the protein does not change dramatically from the reference structure [194]. Following the same idea, but with a different approach are the so-called particle-based coarse-grained models. For them, the simplifaction consists in

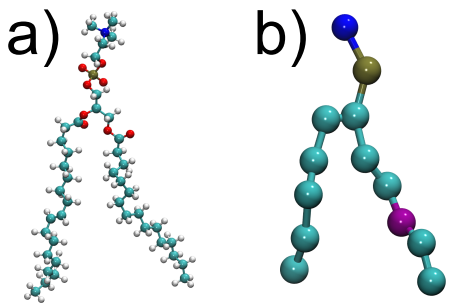

Fig. 2.2 POPC lipid representations for a classical all atom force field, charmm36m (a), and for Martini 2.1 coarse-grained force field (b).

representing the system via a set of particles (pseudoatoms or centroids), that still maintain a level of detail according to the user's needs. For instance, in many protein coarse-grained models, a set of atoms that have similar chemical or physical properties is represented via a new pseudoatom (bead) [64, 170, 191, 241]. In the following section we will describe the coarse-grained models that were used in this thesis to model lipids and proteins.

2.2.2 Lipid Models

Here, we describe the two different coarse-grained models that we have used in this project to model lipids: Martini and Dry Martini model.

Martini

The Martini coarse-grained force field is one of the most widely used coarse-grained force fields for biomolecules. In particular, their lipid parameters were the first to be published alongside a coarse-grained water model. This force field was originally designed expecting high transferability, which is why the CG beads are mapped from different chemical moieties. It has an approximate mapping ratio of 4:1 with respect to its all atom equivalent. The potential energy function as shown in equation 2.22 is composed from the same terms that we have introduced before as a general force field expression (equation 2.11), except for the particular expressions for the angle and torsional potentials (equation 2.23). The bonded interactions are parametrised to match thermodynamic and structural data from experiments as well as atomistic molecular dynamics simulations. The non-bonded interactions are modelled via a Lennard-Jones potential for the van der Waals interactions and via a Coulomb potential for the electrostatic part. The van der Waals radii of the beads is the same for every interaction except for charged beads and beads that form rings. For the case of particles that participate in rings they do not have intra-molecular LJ interactions between

them, but have an improper dihedral defined to maintain the ring planarity. The rest of the intramolecular interactions are excluded only between 2 directly bonded beads. Both non-bonded interactions (van der Waals and electrostatic) have a cut-off of 1.2 nm. The electrostatic interactions are screened with a dielectric constant of $\varepsilon = 15$ and effectively treated only as short-range.

and between 0.9-1.2 nm for the van der Waals interactions. Because of this shifting strategy, both, van der Waals and electrostatics interactions are effectively treated short range only. In its versions 2.0 and 2.1 it is available for numerous lipids, amongst them POPC and POPE which are the ones that we have used in this thesis. For a complete list of available lipids we refer to the latest publication[266].

$$E_{pot} = \sum_{bond} + \sum_{angle} + \sum_{\substack{improper \\ dihedrals}} + \sum_{Coulomb} + \sum_{LJ} \tag{2.22}$$

Or fully expressed by equation (2.23):

$$\begin{aligned} &\sum_{bond} \frac{1}{2} k_r (r - r_0)^2 + \sum_{angle} \frac{1}{2} k_\theta (cos(\theta) - cos(\theta_0))^2 + \sum_{\substack{improper \\ dihedrals}} K_{id} (\theta - \theta_0)^2 + \\ &\sum_{Coulomb} \frac{1}{4\pi\varepsilon_0\varepsilon_r} \frac{q_i q_j}{r_{ij}} + \sum_{LJ} 4\varepsilon_{ij} [(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}})] \end{aligned} \tag{2.23}$$

Dry Martini

This model is a derivation of the Martini model for lipids described in the previous paragraph. The main difference lies in the way both force fields consider solvation. In the case of Dry Martini, it is done implicitly, while standard Martini has explicit water particles. Hence, in the Dry Martini model, the beads are parametrised to already consider solvation implicitly. The Dry Martini force field uses the exact same potential energy expression as the Martini force field (equation 2.23). The non-bonded interactions are modelled again via a Lennard-Jones potential and a Coulomb potential that is from a practical point of view reduced to be only short-range in nature. The main difference with the Martini forcefield comes from the non-bonded parameters that are adjusted to fit the lack of explicit water in Dry Martini, and in some cases the bonded parameters have been slightly modified to fit the standard Martini distribution as well. Because of the lack of explicit solvent, the simulations are run with stochastic dynamics (Langevin dynamics) instead of molecular dynamics. Even though there are parameters available for the ions, the developers encourage users to only use them to neutralise the system. Logically a simulation without explicit solvent or an extra model

for implicit solvent will speed up the simulation, merely as a consequence of the reduction of number of particles in the system, which is an advantage of this model. However, some limitations arise as a consequence of this simplification, such as an over-stabilization of tightly packed beads as well as a somehow inaccurate representation of aggregation and dimerisation processes have been reported when using their protein parameters with lipids.

2.2.3 Protein Models

In the following two sub-sections we describe the two different protein models that were used during this thesis.

Martini

The Martini coarse-grained model for proteins was first published in 2008 to extend the usability of their already popular force field for lipids [170]. It follows the same concept for the mapping, where the reduction on the particle size dimension comes from a 4:1 ratio in mapping (i.e. every approximate 4 heavy atoms in all atom corresponds to 1 bead in CG). The backbone is represented only by one bead located at the center of mass of all the non-hydrogen atoms that correspond to the backbone in the fully atomistic resolution. It is particularly important to notice that its beadtype will relate to the secondary structure that is assigned at the initial mapping. So clearly, the properties of the backbone will depend on the initial reference structure that will be used to map into CG. Side-chain beadtypes, however, only depend on the amino acid. The amount of beads present in the side-chain of the amino acid will vary from 0 (glycine and alanine) to up to 4 (tryptophan) as shown in figure 2.3.

The bonded terms were obtained from a dataset of protein structures with backbone beads corresponding to the center of mass of real backbone atoms. All the bonded terms have the same expressions as presented in equation 2.23 with the exception of an added term for a torsional potential that is model by: $\sum_{torsions} K_\phi(1 + cos(n\phi - \phi_0))$. On top of the bonded interactions, an elastic network based on Cα is used to stabilise the secondary structure. The non-bonded interactions are defined by two terms: a Lennard-Jones type potential for the van der Waals interactions and a screened Coulomb potential for the charge-charge interactions. These interactions were parametrised from the partitioning behaviour of side-chain analogues between water and oil phases, and and with free energy calculations of the insertion of a side-chain in a DOPC bilayer. Both non-bonded interactions are shifted. Electrostatic interactions are switched between an inter-particle distance of 0 and the cut-off, whereas the Lennard-Jones interaction is switched in the region 0.9-1.2 nm, whereas the Coulomb interaction is switched in the region 0-1.2 nm, which makes again all non-bonded

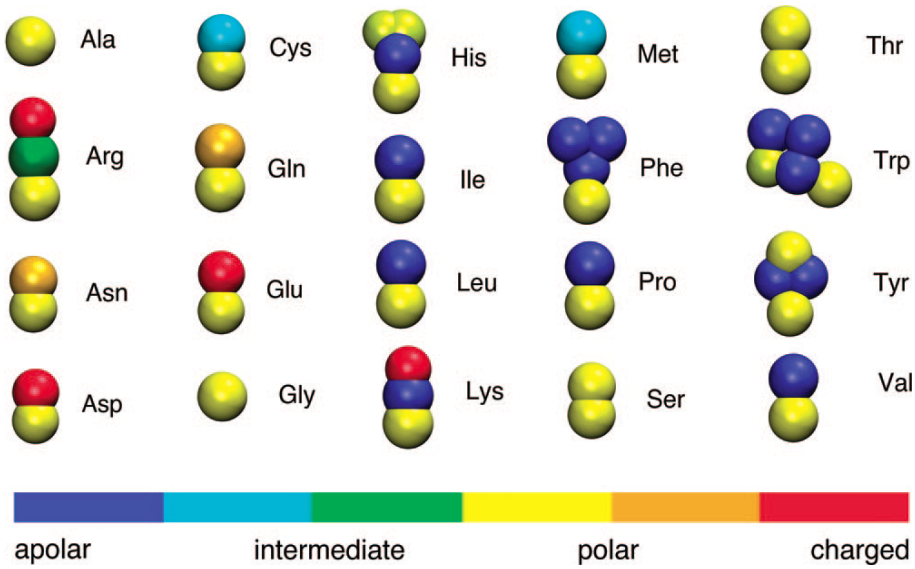

Fig. 2.3 Martini topology for amino acids (including backbone bead). *Taken from [171]*

interactions to be treated only short-range. This force field, although it is extensively used, has a main drawback in its alpha available versions (v.2.0, v.2.1), because of the restraints on the secondary structure that cannot change dynamically during the simulation. However, when being aware of that main limitation, simulating with this force field can be a useful approach to study tertiary conformational changes, supramolecular organization and the behaviour of proteins in a membrane environment [158]. This model has been tested and extensively used with membrane and solvent models from the same force field. To maintain the structural stability of the proteins, an elastic network is routinely used on top of the normal force field parameters.

OPEP

The OPEP force field (***O**ptimized **P**otential for **E**fficient **P**rotein structure prediction*) is an implicit solvent coarse-grained model for protein and peptides. The backbone is represented in full atomistic detail, i.e. N, H, Cα, C and O are CG beads. On the other hand, the side-chain is represented only by one bead. Being proline the only exception that has 3 beads. The side-chain position is defined by the backbone conformation with a a discrete representation. In the version 4.0 of this forcefield, which is the one that we are using in this work, the Hamiltonian is represented by a combination of generic bonded terms (inherited from AMBER all atom force field) and non-bonded interactions (van der Waals and H-bonds), as shown in equation 2.25. OPEP force field was optimised to reproduce the experimental lowest free-energy conformation and the melting temperatures of a small

set of peptides. On top of that, the interactions between charged side-chains have been optimised to fit all-atom PMF simulations. Even though there is no solvent in the model the solvent effects are implicitly taken into account within the non-bonded terms. Since it does not restraint the secondary structure, it can be particularly useful to study protein folding, structure prediction but also aggregation events [51, 54, 162]. Some drawbacks are however the reduced side-chain specificity, because of the low resolution of it, and the small time step that has to be used because of the full-detailed backbone (1.5 fs, that could be pushed up to 3.5 fs when rescaling the mass of the hydrogen).

Here we describe in detail the functional form of the OPEP force field as presented in [201]:

$$E_{pot} = E_{local} + E_{non-local} + E_{hydrogen\ bonds} \tag{2.24}$$

Where the local component (E_{local}) comprises all the intramolecular interactions that exist as a consequence of the intramolecular connectivity expressed, the non-local one includes all the non-bonded interactions and the last term ($E_{hydrogenbonds}$) accounts for the hydrogen-bond propensity of the backbone atoms.

In more detail, the local contributions are expressed by:

$$\begin{aligned} E_{local} = w_b \sum_{bonds} k_b(r-r_0)^2 \ + \ w_a \sum_{angles} k_\alpha(\alpha-\alpha_0)^2 \ + \\ w_t \sum_{torsions} K_\phi(1+cos(n\phi-\phi_0)) \ + \ w_{\phi,\psi}(\sum_{\phi} E_\phi + \sum_{\psi} E_\psi) \end{aligned} \tag{2.25}$$

where E_ϕ and E_ψ are the terms used to constrain the backbone dihedral angles to be confined in the Ramachandran plots via harmonic barriers as follows:

$$\begin{aligned} E_\phi &= k_{\phi,\psi}(\phi-\phi_0)^2 \\ E_\psi &= k_{\phi,\psi}(\psi-\psi_0)^2 \end{aligned} \tag{2.26}$$

where $\phi_0 = \phi$ if the torsion is found in the interval $[\phi_{lower}, \phi_{upper}]$, otherwise $\phi_0 = min(\phi - \phi_{lower}, \phi - \phi_{upper})$; the confined region is bound by $\phi_{lower} = -160^\circ$ and $\phi_{upper} = -60^\circ$. For the torsion ψ the angles are confined by $\psi_{lower} = -60^\circ$ and $\psi_{upper} = 160^\circ$.

The non-bonded interactions, previously referred to as non-local, are separated by interactions involving side-chain, Cα, or the rest of the backbone atoms (N,C,O and H) that will be

notated as M' in the following equations. The form of the non-bonded potential corresponds then to:

$$E_{non-local} = w_{1,4} \sum_{1,4} E^1_{vdW} + w_{1>4} \left(\sum_{M',M'} E^1_{vdW} + \sum_{M',C\alpha} E^1_{vdW} + \sum_{M',Sc} E^1_{vdW} + \sum_{C\alpha,Sc} E^1_{vdW} \right) + w_{C\alpha,C\alpha} \sum_{C\alpha,C\alpha} E^2_{vdW} + w_{Sc,Sc} \sum_{Sc,Sc} E^2_{vdW} \tag{2.27}$$

In OPEP, the vdW potential has two different expressions depending on the beads that are involved (E^1_{vdW} or E^2_{vdW}). E^1_{vdW} is a standard 12-6 potential that has the following functional form:

$$E^1_{vdW} = \varepsilon_{i,j} \left(\left(\frac{r^0_{ij}}{r_{ij}} \right)^{12} - \left(\frac{r^0_{ij}}{r_{ij}} \right)^{6} \right) \tag{2.28}$$

where r_{ij} is the distance between the particles i and j, $r^0_{ij} = (r^0_i + r^0_j)/2$ with r^0_i and r^0_j are the van der Waals radius of a particle i or j respectively and ε_{ij} the coupling constant.

The other potential form for the van der Waals interactions is:

$$E^2_{vdW} = E_{AR} - \varepsilon_{i,j} \left(\frac{r^0_{ij}}{r_{ij}} \right)^{8} H(-\varepsilon_{i,j}) \tag{2.29}$$

where the final repulsive term holds for $\varepsilon_{ij} < 0$ while the the attractive/repulsive interactions (E_{AR}) have the following potential form:

$$E_{AR} = \varepsilon_{i,j} \left(\left(\frac{G(r^0_{ij})}{r_{ij}} \right)^{6} e^{-2r_{ij}} + A_0 \, tanh[2(r_{ij} - r^0_{ij} - 0.5) - 1] \right) H(\varepsilon_{i,j}) \tag{2.30}$$

where $G(r^0_{ij})$ is determined by imposing the value of the potential at a specific location r^0_{ij}. $H(\varepsilon)$ is a heavyside function that equals to 1 for $\varepsilon \geq 0$ and equals to 0 otherwise. A_0 is a constant (0,6563701).

Finally, the last term of the OPEP Hamiltonian relates to the hydrogen-bond propensities. It is expressed as a sum of a two-body and a four-body term ($E_{HB} = E^{2-body}_{HB} + E^{4-body}_{HB}$). The

2-body term is decomposed in short and long range hydrogen bonds with respect to sequence position of the residues involved, such as:

$$E_{HB}^{2-body} = w_{1,4}^{HB} \sum_{i,j,i=j+4} \varepsilon_{1,4}^{HB} \mu(r_{ij}) \nu(\alpha_{ij}) + w_{1>4}^{HB} \sum_{i,j,j>i+4} \varepsilon_{1>4}^{HB} \mu(r_{ij}) \nu(\alpha_{ij}) \tag{2.31}$$

where

$$\begin{aligned} \mu(r_{ij}) &= 5\left(\frac{\sigma}{r_{ij}}\right)^{12} - 6\left(\frac{\sigma}{r_{ij}}\right)^{10} \\ \nu(\alpha_{ij}) &= \begin{cases} cos^2(\alpha_{ij}) & \text{for } \alpha_{ij} > 90 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{2.32}$$

where r_{ij} is the O-H distance between the carbonyl oxygen and the amide hydrogen of the backbone, α_{ij} the angle defined by the N-H-O atoms forming the hydrogen bond and σ is the equilibrium distance of the hydrogen bond.

The 4-body term has the form of weighted products of Gaussian functions that each monitor the existence of an H-bond based on a distance criteria. The Gaussian functions are calculated for each possible H-bond pair and they give a contribution to the 4-body term only if certain conditions are satisfied. This allows the stabilisation of α-helices and β-sheets independently of the (ϕ, ψ) dihedrals. For a full description of this potential we remit the reader to reference [201].

2.3 Including Hydrodynamics Interactions

The development of implicit solvent CG model for biological systems stimulated the study of very large systems as well as long time-scales. However, when removing the solvent degrees of freedom two problems must be considered. The first is the lack of explicit solvation contribution to the thermodynamics of the system. Depending on the nature of the CG these effects could be, in principle, incorporate at some extend in the potential energy function of the simplified model. The second, that is of particular relevance here, is a dynamical effect, namely the lack of solvent mediated correlations, or hydrodynamics, since there is no exchange of momenta with the solvent. A very simple approach to model the effect of a solvent bath on the dynamics of a particle is to use Langevin dynamics where the dissipative term relates to the viscosity of the medium, and the random fluctuations mimic collisional events. However, Langevin dynamics does not allow to introduce memory effects or spatial solvent-mediated correlation. For this reason, in order to properly consider hydrodynamics in

the simulation of particles it is necessary to use specialized techniques. In the context of the biosimulations, the most common approach is to use Brownian dynamics with hydrodynamic interactions (HI), developed in the 70's, by Ermak and McCammon [76]. However, this technique suffers, from the computational point of view, of bad scalability when the system is very big. Thanks to the increasing computational power, the technique has been applied to investigate several processes, like biomolecule diffusion [86] and folding kinetics [109]. A second approach, developped by R. Kapral, is the stochastic rotational dynamics [157], that has been recently applied to investigate how HI affect protein motion during enzymatic turnover. Finally, we mention the Lattice Boltzmann molecular dynamics (LBMD) approach that we have used in this work. The LBMD combines the Lattice Boltzmann technique from fluid-dynamics with particle molecular dynamics. This scheme was first applied to study polymer systems [2], but recently was extended to treat biological systems, namely to study protein systems such as crowded protein solutions, shear-induced unfolding and amyloid aggregation [54, 55, 143, 239, 240].

2.3.1 Lattice Boltzmann technique

The Lattice Boltzmann method is a discrete formulation of the Boltzmann equation on a lattice, leading to the Navier–Stokes equations in the incompressible limit. It has been rigorously funded and widely used in fluid dynamics [91, 200]. The technique is especially appealing in high perfomance computing since it exploits very well parallel computers, and therefore it allows to study very large systems [246].

Historical predecessor: Lattice Gas Cell Automata

The Lattice Boltzmann methodology has been developed in the late 1980's as an improvement from the lattice gas cellular automata (LGCA) model for fluid dynamics simulations[245]. The principal improvement relies on the avoided noise by avoiding a binary velocity treatment assignation in each node of the lattice. In fact, the continuum macroscopic representation of the fluid becomes discretised by a set of particles that move between consecutive nodes on a fixed regular lattice. The evolution of those particles can be done with two different types of movement: free-streaming and collision. Free streaming consists of the displacement of particles from site to site according to a set of discrete speeds. Consequently, a particle sitting at site $\boldsymbol{x}$ at time t with speed $\boldsymbol{c}$ will move to site $\boldsymbol{x}+c_p$ at time $t+\Delta t$. A set of boolean variables $n_p(\boldsymbol{x},t)$ with $(p=0,1,..,M)$ describes the occupation at each node of the lattice, where M corresponds to the number of directions of the particle velocities ($\boldsymbol{c}_p$) at each node.

The formal equation to describe the dynamics of LGCA is as follows:

$$n_p(\boldsymbol{x}+\boldsymbol{c}_p\Delta x, t+\Delta t) = n_p(\boldsymbol{x},t)+\Omega_p(n_p(\boldsymbol{x},t))(p=0,1,..M) \tag{2.33}$$

In equation 2.33 the first term of the right hand describes the free streaming dynamics and the second the collision ($\Omega_p(n_p(\boldsymbol{x},t))$, collisional operator)

It is important to notice that the collisions obey two principal rules that relate them with a real molecular collision, namely that they conserve the particle number and the total momentum. This is a necessary (but non-sufficient) condition that enables the model to achieve hydrodynamic behaviour when considering a large group of particles.

Non-linear Lattice-Boltzmann equation

In order to improve the main drawback of LGCA, namely, the issues with statistical noise derived from a boolean representation of the particle occupation, McNamara and Zanetti proposed the Lattice Boltzmann equation for the first time in 1988 [164]. The improvement, came from simply considering an ensemble-averaged population to occupy the nodes in the lattice instead of just a boolean occupation ($f_p = \langle n_p \rangle$). It is straightforward to rewrite equation 2.33 considering now population states, as:

$$f_p(\boldsymbol{x}+\boldsymbol{c}_p\Delta t, t+\Delta t) = f_p(\boldsymbol{x},t)+\Omega_p(f_p(\boldsymbol{x},t)) \tag{2.34}$$

where $f_p(\boldsymbol{x},t)$ denotes the particle velocity distribution function (i.e the probability distribution of finding a particle in the site x of the lattice at time t and moving in lattice space with velocity $\boldsymbol{c}_p$) and Δx the spacing of the lattice grid, and Ω_p the collision operator. Again we have a streaming and collisional contribution to the dynamics. A simple form of the collisional operator, introduced by Bhatnagar-Gross-Krook (BGK), mimics the local fluctuation around equilibrium of the velocity field. It reads as follows:

The left-hand side of the equation 2.34 corresponds to the free-streaming motion previously described, whereas the right-hand part of the equation describes the interaction part that drives the system towards a local equilibrium f_p^{eq}. It follows then, that the collisional operator can be written as:

$$\Omega_p(f_p) = -\frac{1}{\tau}(f_p - f_p^{eq}) \tag{2.35}$$

where f_p^{eq} is the discretised equilibrium velocity distribution, τ is the single relaxation time towards a local equilibrium that relates to the system viscosity. This BGK collisional operator allows for the conservation of mass, momenta and energy.

Another important factor to consider is the type of lattice that needs to be used in order to assure the conservation of mass and momenta laws of hydrodynamics but also to assure rotational isotropy. For instance, one of the 3D lattices that is widely used and compatible with these requisites is the D3Q19, which has a set of 19 discrete velocities, that are distributed as shown in figure 7.2

Finally, we can express the updating scheme of the Lattice Boltzmann fluid representation in the particular case of this thesis work. In this case, in the context of an hybrid simulation that has particle-fluid communication equation 2.34 can we written as:

$$f_p(\boldsymbol{x}+\boldsymbol{c}_p\Delta t, t+\Delta t) = f_p(\boldsymbol{x},t) - \frac{1}{\tau}\Delta t(f_p - f_p^{eq})(\boldsymbol{x},t) + g_p(\boldsymbol{x},t) \tag{2.36}$$

where the last term, $g_p(\boldsymbol{x},t)$ relates to the coupling between solvent and atoms [239].

2.3.2 Lattice Boltzmann Molecular Dynamics

In this joined approach, the particle-based dynamics is coupled to a kinetic representation of the solvent, that is simulated via the Lattice Boltzmann (LB) technique. In this implementation of LBMD, the coupling between the particles and the solvent is treated by adding a Stokes-like drag force that acts on a per-particle basis:

$$\boldsymbol{F}_i^D(R_i) = -\gamma(\boldsymbol{V}_i - \widetilde{\boldsymbol{u}}(R_i)) \tag{2.37}$$

where $\boldsymbol{V}_i$ is the particle velocities, and $\widetilde{\boldsymbol{u}}$ indicates the fluid velocity field distributed over the region occupied by the atom. γ represents the translational coupling coefficients that regulates the exchange of momentum between particles and fluid. This drag force is added to the forces deriving from the Hamiltonian of the particle systems, $\vec{F}_i^C = -\vec{\nabla}_i U(\{r\})$. The dynamics of the particle system is then controlled by its Hamiltonian, hydrodynamics interactions emerging from the coupling with the fluid, and a random white noise, $\vec{F}_i^R$, that represents thermal fluctuations. Because of the action-reaction principle the drag force enters in the evolution of the fluid, as mentioned before, via the term $g_p(\boldsymbol{x},t)$ [26, 239].

As a final note, we will coment on the choice of grid spacing. When coupling different LB and MD it is important to have a grid resolution that is sufficiently high in order to resolve local hydrodynamic interactions for the particle system resolution that is used as well as reproduce water kinematic viscosity [54, 239]. For instance in the case of this work that uses CG particles with a near-atomistic resolution, the grid spacing corresponded to 3 Å.

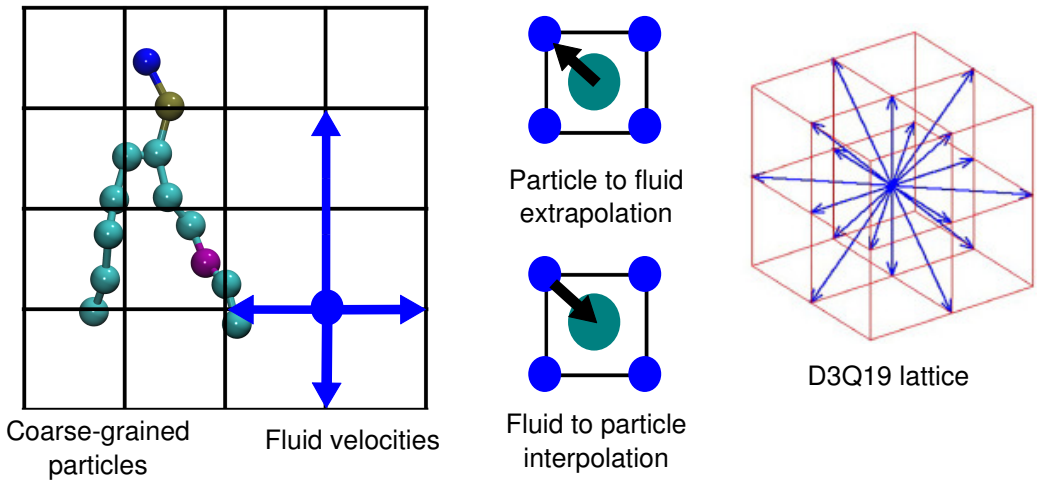

Fig. 2.4 Schematic representation of the coupling between the molecular system (CG, MD) and the fluid (Lattice Boltzmann). The blue arrows represent the fluid velocity population that evolves on the lattice. We also show a depiction of interpolation and extrapolation strategies used to comunicate forces between particle and fluid dynamics. Finally, we show the 3D lattice topology used in this work: D3Q19. *Adapted from [239, 246]*

Chapter 3

Computational study of mitofusin assembly

This chapter presents the work related to the study of mitofusin oligomerisation as published in [34, 35]. This work was done in close collaboration with Dario De Vecchis, where he elaborated the initial homology model of Fzo1 and proposed the different oligomers that we have studied. My contribution was to curate the models when needed, embed them into a membrane environment (prepare the set-up), run and analyse all the coarse-grain simulations.

3.1 Introduction

Mitochondria are dynamic organelles organized as a cytoplasmic reticulum. Mitochondria fuse their outer and inner membranes to form tubules. These mitochondrial tubules can interconnect through fusion but can also fragment through fission of their membranes to yield a network with remarkable plasticity. Together, fusion and fission regulate the entire morphology and dynamics of the mitochondrial network, which makes these processes essential for the maintenance of mitochondrial integrity and consequently all mitochondrial functions [89, 230, 252, 267]. From a mechanistic point of view, mitochondrial fission as well as fusion of outer and inner membranes are all mediated by members of the dynamin-related proteins (DRPs) family [203]. These large GTPases act as remodelers of intracellular lipid bilayers through two properties: their capacity to bind biological membranes and their propensity to oligomerize into high order macromolecular structures. Consistent with these features, fission DRPs are recruited to mitochondrial outer membranes by specific adaptor proteins, where they auto-oligomerize upon binding of GTP to assemble into macromolecular spirals that wrap around mitochondrial tubules [252]. Subsequent GTP hydrolysis induces

conformational rearrangements of the dynamins, which result in reduced diameter of the spirals and constriction of mitochondrial tubules, followed by their separation. Fusion-DRPs are transmembrane proteins that promote homotypic merging of the lipid bilayers in which they are inserted [58]. The mitofusins MFN1 and MFN2 (Fzo1 in yeast) fuse mitochondrial outer membranes, whereas fusion between inner membranes is mediated by OPA1 (Mgm1 in yeast). Another key member of fusion DRPs is the atlastin ATL-1 (Sey1 in yeast), which merges membranes from ER tubules. Mitofusins and atlastins have in common the ability to auto-oligomerize in trans (from opposing lipid bilayers), in a GTP binding and hydrolysis-dependent manner, to tether opposing lipid bilayers and promote their fusion [58]. Recent crystal structures of protein portions lacking the transmembrane regions of ATL-1, Sey1 and MFN1 further allow depicting the assembly of DRPs during membrane fusion as protein dimers interacting in trans through their respective GTPase domains [30, 44, 46, 172, 270, 271] However, the yeast mitofusin Fzo1 also forms high molecular weight complexes [204]. In particular, Fzo1 was shown to trigger the formation of a ring-shaped macromolecular complex during mitochondrial docking [37], suggesting that DRPs assembly during membrane fusion might involve the formation of high order complexes rather than solely trans homodimers. Consistent with this possibility, mitofusins form oligomers in both cis (on the same lipid bilayer) and trans to mediate membrane attachment and fusion [8, 46, 101, 120, 135, 214, 230]. In addition to the GTPase domain, mitofusins also include two heptad-repeat domains (HR1 and HR2), which may become involved in homotypic interactions during mitochondrial tethering [70, 101, 135]. Whether GTPase and HR domain interactions take place in cis or trans before or during mitochondrial tethering remains to be investigated or confirmed [84, 135]. Here, we rely on our recent Fzo1 model [70] to investigate and discuss distinct hypotheses for the formation of cis and trans mitofusin oligomers. In addition to being the only near full-length model as of today, this model contains a membrane domain, which is a requisite for investigating mitofusin orientations in a bilayer. We employed a modelling procedure that has been guided by available experimental data from the literature. The models proposed in this study were further assessed for stability and dynamics in a membrane environment using a coarse-grained molecular dynamics (MD) simulation approach. This study opens new possibilities to start exploring cis and trans conformations for Fzo1 and mitofusins in general.

3.2 Methods

The modelling work presented here builds upon an experimentally validated model of the monomeric unit, namely our previously published model of Fzo1 in a closed conformation

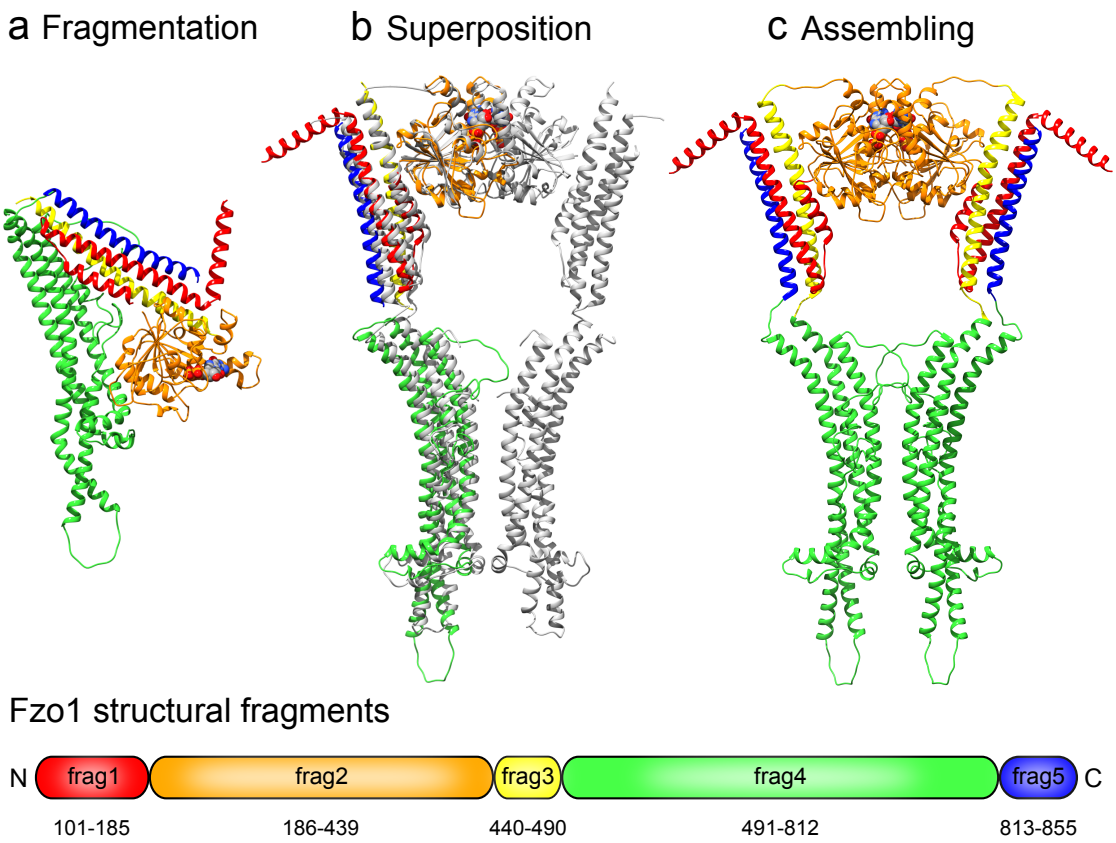

Fig. 3.1 Key steps of the modelling workflow to build a model of Fzo1 in open conformation, based on the BDLP structure. a. The Fzo1 structural blocks delimited according to the putative hinges proposed for BDLP are highlighted by different colours. The blocks are coloured in rainbow from N-terminal (red) to C-terminal (blue). b. Each block in Fzo1 (coloured parts) was superposed onto its corresponding fragment in 2W6D (grey ribbon). Note the backbone interruptions in the structure. c. The final Fzo1 model in open conformation. Note that each GTPase domain (orange) from both chains is in close contact with each other. The GDP nucleotide is shown in space-filling representation. A cartoon scheme showing the colouring and sequence positioning of the structural domains is shown at the bottom.

[70]. It is first used to generate a monomeric model of the open conformation. Those two models are the basis for the construction of Fzo1 dimer models in cis and then tetramers via the dimerization in trans of Fzo1 cis-dimers.

3.2.1 Modelling the Fzo1 GTPase dimer construct

Two chains of the Fzo1 model in closed conformation [70] were placed by superimposing their GTPase domains onto that of BDLP in open conformation (PDB-id 2W6D, [153]). Then, only the coordinates of the two fragments that comprise the Fzo1 GTPase domain (res

188-461) were retained to form the final GTPase domain dimer model. The loop refinement tool implemented in MODELLER [82] was used to remove a clash in both chains involving an unresolved loop in the template 2J68 [152] (res 215-219). Models were ranked according to the discrete optimized protein energy (DOPE) method [224], selecting the best-scoring loop out of 20 candidates.

3.2.2 Modelling cis-dimer configurations

The Fzo1 head-to-head cis-dimer: Two Fzo1 chains in closed conformation [70] were oriented facing each other within a compatible distance to accommodate two interacting GTPase domains. Subsequently, the coordinates of residues 188-440 enclosed between hinges 2a and 2b (i.e., comprising the GTPase domain) were removed from both chains and replaced with the GTPase dimer construct described above. The latter was manually positioned between the two deleted chains, resulting in the head-to-head interaction dimer (3.2b). The backbone interruptions were connected using the loop refinement tool implemented in MODELLER [82] using positions 185-188 and 436-445 as anchors. Solutions were ranked according to the DOPE method [224], selecting the best-scoring loop out of 20 candidates.

The Fzo1 back-to-back cis-dimer: Two chains of the Fzo1 model in closed conformation [70] were manually oriented with respect to each other to generate the back-to-back interaction (discussed in the text). In the resulting model system (figure 3.2c), the HR domains face each other in a parallel fashion.

The Fzo1 open cis-dimer: The coordinates from BDLP in open conformation (PDB-id 2W6D, [153]) derived from the electron density map of native BDLP lipid tubes (accession code: EMD-1589) were used as a template to model Fzo1 in open conformation. Starting from our previous Fzo1-BDLP target-template alignment [70] and using an analogous approach to the one described in [153], we introduced homologous chain breaks on the Fzo1 model [70], resulting in five rigid blocks (figure 3.1a). Each fragment was superposed to its corresponding fragment in 2W6D to reconstitute the orientation found in BDLP (figure 3.1). The MatchMaker tool from the UCSF Chimera software [196] was used during this procedure. The loop refinement tool implemented in MODELLER [82] enabled us to complete the model in the resulting backbone interruptions and to remove a clash in both chains between the side chain of the Lys271 and the backbone of the Ala401 residues, using positions 268 and 273 as anchors. Solutions were ranked according to the discrete optimized protein energy (DOPE) method [224], selecting the best-scoring loop out of 10 models.

3.2.3 Modelling the trans-dimer configurations

The Fzo1 head-to-head trans tethered tetramer in open conformation: Two Fzo1 open cis-dimer models obtained as described above were manually oriented to mimic the interactions in trans towards their respective GTPase domains. In the resulting model system, the two transmembrane segments are located at opposite ends (figure 3.5a).

The Fzo1 head-to-head trans tethered tetramer (antiparallel): The transmembrane segment of two Fzo1 head-to-head cis-dimers described above were manually oriented at opposite ends to optimize the interaction between their respective HR domains oriented in an antiparallel fashion (figure 3.5c). Note that this system, although antiparallel, could also be considered as back-to-back.

The Fzo1 head-to-head trans tethered tetramer (parallel): Two Fzo1 back-to-back cis-dimers obtained as described above were initially positioned with the respective transmembrane segments at opposite ends to mimic the supposed tethering process. Subsequently, the coordinates of residues 101-491 and 816-855 enclosed between hinges 1a and 1b (i.e., comprising the GTPase domain and the 3-helix bundle), were removed from the two resulting juxtaposing chains. Then, in a similar manner, the GTPase dimer construct was built (see above). We superposed the GTPase domain alpha carbons of two Fzo1 chains in closed conformation [70] with the human mitofusin dimer (PDB-Id: 5GOM, [46]). This choice was motivated and directly inspired by the work from Gao and collaborators, who proposed a possible Mfn1 trans cross oligomer [46]. From the resulting Fzo1 dimer, only the GTPase domain and the 3-helix bundle were selected and used to replace the aforementioned deleted portions, thus generating the trans head-to-head interaction (figure 3.5b). A clash in one chain (res 215-218) was removed from the resulting Fzo1 dimer using the loop refinement tool implemented in MODELLER [82]. Models were ranked according to the discrete optimized protein energy (DOPE) method [224], selecting the best-scoring loop out of 10 candidates. The same tool was used to reconstitute the backbone interruptions. Positions 491-495 and 812-816 were selected as anchors. The best-scoring loop was selected out of 10 candidates.

3.2.4 Molecular Dynamics

System setup and parameters. Topologies to run coarse-grained (CG) simulations were generated with the martinize tool choosing the Martini v.2.1 force field with an elastic network [67, 170]. The force bond constant was set to $500 kJmol^{-1}nm^{-2}$ with lower and upper elastic bond cutoffs of 0.5 and 0.9 nm respectively. Firstly, 5000 steps of steepest descent with position restraints for the protein were run followed by 5000 steps without restraints. The obtained coordinates were inserted in a POPC:POPE (1:1) membrane via the

insane tool [266], where the membrane position was manually set up to match the reported transmembrane regions corresponding to residues 706-726 and 737-757 according to UniProt numbering [70]. All systems were fully solvated to mimic an environment of 150 mM of NaCl solution. (See Supplementary Fig. 3.13 for more details about each simulation setup.) All the final systems followed the same simulation protocol using the Gromacs 5.0.4 software [1]. A further 5000 steps of steepest descent minimisation with position restraints of $1000 kJmol^{-1}nm^{-2}$ in protein and lipids were followed by 5000 steps without position restraints. Equilibration was performed in three stages, with timesteps of $20fs$. Firstly, 25000 steps of equilibration were run at 310 K using the V-rescale thermostat [42] and semi-isotropic pressure coupling via Berendsen barostat [23] with position restraints of $1000 kJmol^{-1}nm^{-2}$ for protein and lipids, followed by the same setup without position restraints. Finally, the last equilibration step was run for 50000 steps with the V-rescale thermostat (coupling constant $tau_t = 1ps$) and semi-isotropic coupling with Parinello-Rahman barostat [190] (coupling constant $tau_p = 12ps$). Production runs were 1 μs long for all six systems, following the same parameters as those used in the last equilibration setup.

3.2.5 Analysis

Root mean square deviations were calculated in Gromacs 5.0.4 via the gmx rms tool considering only backbone beads. Interfaces were obtained based on differences between solvent accessible surface areas obtained with the gmx sas tool using a probe radius of 0.2638 nm in agreement with the CG water size. An interface between, e.g., protein A and B (IA/B), was calculated as the sum of the surface areas of protein A and protein B in the interface, following the equation: $I_{A/B} = I_A + I_B = (SASA^A - SASA^A_{complex}) + (SASA^B - SASA^B_{complex})$ where $SASA^A$ is the solvent accessible surface area calculated considering the protein A as isolated and $SASA^A_{complex}$ is the solvent accessible surface area calculated for the protein A considering its solvent accessibility in the complex. Interface contacts between dimers were calculated for the last 100 ns of the simulation. A contact was defined when a backbone bead (BB) corresponding to a monomer was up to 8 Å away from another BB bead from the other monomer. The total contact counts were divided by the number of frames to report an occupancy value. The distance in the membrane-normal Z axis between the phosphate beads from both bilayers (oriented in the x, y plane) was measured in blocks (grid in x, y of size 20 Å). Z coordinates of atoms corresponding to each block were averaged and, subsequently, the distance between the average values of each block from the upper bilayer was subtracted from the lower one. Resulting difference inter-membrane distances between the final and initial structure of the production run are shown as a matrix.

3.3 Results and discussion

3.3.1 Putative models for Fzo1 cis-dimers

Mitochondrial tethering requires oligomerization of mitofusins in cis and then in trans [8, 93, 120, 121, 135, 204, 214]. We thus reasoned that focusing on cis-oligomerization aspects in the first instance might facilitate investigating their general properties of assembly. In this regard, the bacterial dynamin like protein (BDLP) from Nostoc punctiforme represents a starting point of choice. BDLP was not only used as the template to generate our previous Fzo1 structural model [70], but this bacterial DRP also displays well established cis-oligomerization properties [152, 153]. Upon binding of non-hydrolysable analogues of GTP, BDLP operates an extensive conformational change from a compact structure (that was used as the template to generate the closed Fzo1 model, figure 3.1a) to an extended conformation. This extended conformation allows insertion of BDLP into lipid bilayers through a membrane paddle and favours its cis-oligomerization through GTPase domain interactions (PDB-id 2W6D, [153]). The Fzo1 open-conformation model was obtained by superimposing our previously published closed conformation model piecewise onto the open structure of BDLP (figure 3.1b; see Methods for details). The final model of the Fzo1 dimer in open conformation was then superposed with respect to the template structure 2W6D using the Cα atoms (figure 3.1c). The resulting structural drift, measured as RMSD between the target and the template, is rather low at 0.63 Å, which underlines the similarity of both molecular systems.

The BDLP-like cis-oligomerization model (figure 3.2a) imposes an extensive conformational switch of Fzo1 with an opening angle of about 180° between the four-helix bundle and the trunk of the mitofusin. While a possible conformational reorganization has been experimentally documented in this hinge region of Fzo1 [57], a significantly lower angle of opening may not be excluded. GTPase domain contacts observed for human mitofusin [46] provide a template to form dimers with such a low opening angle (figure 3.2b). This assembly leads to a distinct head-to-head orientation with a notably narrow Fzo1 opening that differs from the significantly wide Fzo1 opening seen in the BDLP-like cis-oligomerization model.

In either the wide or narrow head-to-head models, the GTPase domain interface would take place after GTP binding [46, 153] to promote cis-dimerization of Fzo1. This contrasts with the current view that oligomerization of mitofusins through their GTPase domain promotes their association in trans [46, 271]. In this latter scenario of trans-association, cis-oligomerization should thus employ Fzo1 interaction regions distinct from the GTPase domain. Interestingly, in our recent Fzo1 model, the portion of HR2 belonging to the trunk

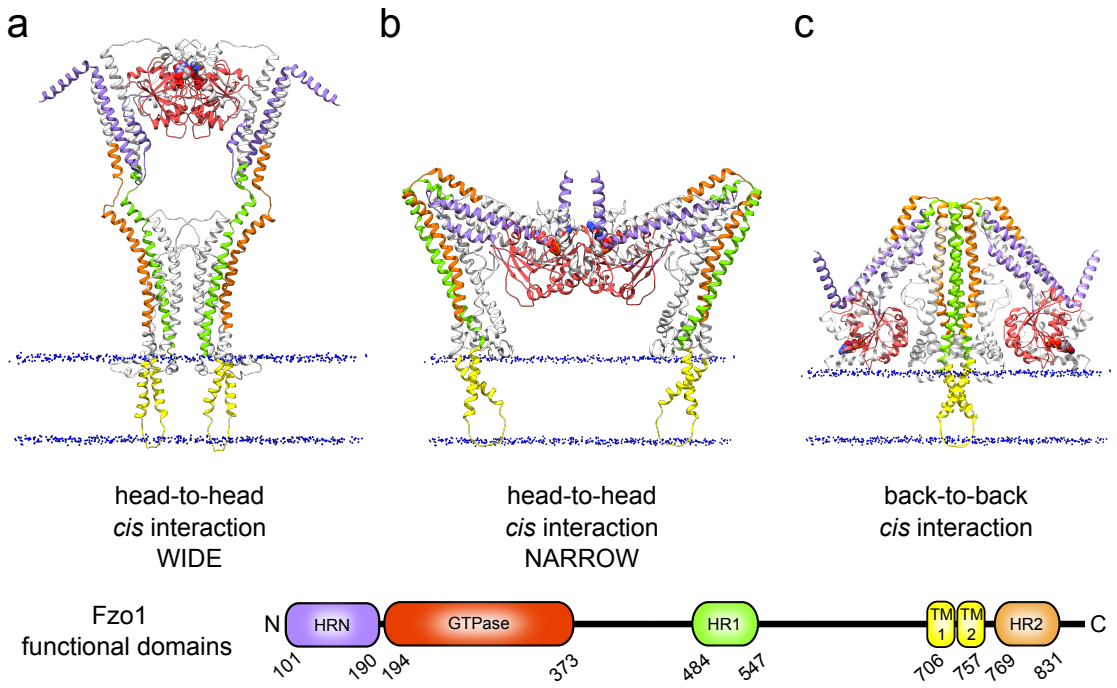

Fig. 3.2 Putative Fzo1 models for interaction in cis configuration. a. and b. Wide and narrow head-to-head complexes, respectively, in which two Fzo1 molecules interact via their GTPase domains, consistent with interactions observed for the bacterial BDLP [152, 153] and human mitofusin [46]. c. Back-to-back complex for the closed model where two Fzo1 molecules interact through their HR domains. The domains highlighted in colour are: violet, HRN (N-terminal heptad-repeat); green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (blue) from lipid bilayer headgroups and GDP nucleotides are depicted in the space-filling representation. A cartoon scheme showing the colouring and sequence positioning of the functional domains is shown at the bottom.

is exposed to the solvent, suggesting that it could be available for putative hydrophobic interactions with neighbouring Fzo1 molecules [70]. The coiled-coil structures would be positioned back-to-back to yield a cis-dimer with GTPase domains available for interactions in trans (figure 3.2c). This would notably be consistent with the crystal structure of the MFN1 HR2 domain [135], which supports the possibility that mitofusins could also associate through the hydrophobic side of their HR2 domain. However, in this crystal (PDB-id 1T3J), the interacting HR2 domains adopt an antiparallel orientation, suggesting that the transmembrane segments of two mitofusin molecules interacting through their HR2 domains would be located on opposite membranes.

We have tested the robustness of the three cis-dimer models through coarse-grained molecular dynamics (MD) simulations (see Methods). The final conformations obtained for all three MD simulations (for whole proteins see Supplementary Fig. 3.9;for protein

interface details see Supplementary Fig. 3.12 and Supplementary Figs. 3.15-3.17) stay close to their respective starting models (Supplementary Fig. 3.9). This observation suggests that the models are relatively stable, and therefore can all be considered as possible molecular assemblies (see Supplementary Fig. 3.14). Notably, the dynamics of the closed back-to-back model shows a decrease in the membrane-protein interaction surface as well as an increase in the protein-protein interface (figure 3.3, cyan curve).

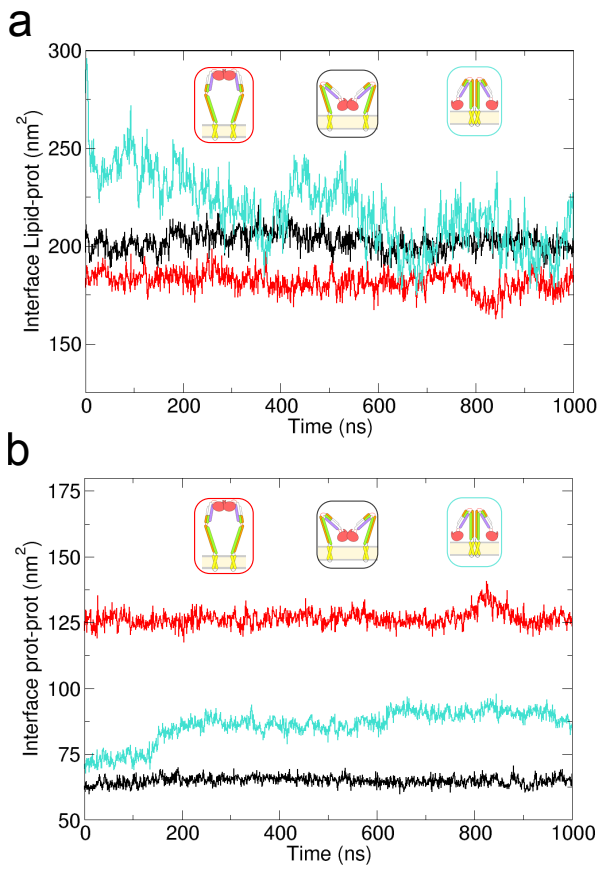

Fig. 3.3 Lipid-protein and protein-protein interface evolution during CG MD simulations for the dimeric complexes. a. Lipid-protein interfaces for the three different studied cis-dimers. b. Protein-protein interfaces belonging to the three different cis-dimers. The colours in the curves match the outline of the corresponding image in the inset (black: head-to-head narrow, red: head-to-head wide, cyan: back-to-back). The domains in the inset schemes are coloured in the same way as Fig. 2.

The latter is a consequence of the augmented interaction of their transmembrane regions with each other that consequently decreases the lipid-protein interface. At this point of our study, MD simulations alone do not allow privileging the validity of any model over the others. For this reason, we analysed the propensity of each cis-dimer to possibly associate with the mitochondrial outer membrane carrier protein Ugo1, as cis-dimerization of Fzo1

has been suggested to involve the participation of Ugo1 [8]. This 3-membrane spanning factor, which is essential for outer membrane fusion [60, 114, 220, 221, 269], interacts with Fzo1 through the region spanning residues 630-703 and 756-843 (Sesaki and Jensen, 2004) and may thus contribute to deciphering the differential likelihood of the three cis-models. Interestingly, it is obvious that in the wide and narrow head-to-head dimers, the region required for potential interactions with Ugo1 is exposed and fully accessible (figure 3.4a; 75-77 nm2 and 4b; 62-63 nm2). In contrast, this region becomes partially masked (44-47 nm2) and likely less accessible to potential Ugo1 molecules in the back-to-back model (figure 3.4c). If Ugo1 indeed participates in Fzo1 cis-dimers formation, these observations would tend to favour head-to-head models as opposed to the back-to-back configuration.

3.3.2 Proposed models for Fzo1 trans-tethered oligomers

Based on the three previously obtained cis-dimers of Fzo1, we aimed at modelling potential mitofusin trans-oligomers. We took here in consideration the possibility that cis-dimers from one membrane could engage in interactions with cis-dimers from an opposing membrane through the availability of either their GTPase or trunk domains. Importantly, the average 8 nm distance between outer membranes observed during Fzo1-mediated mitochondrial tethering [37] provides an experimental estimation of the membrane separation that Fzo1 trans-oligomers should favour.

Following GTP binding, wide opening of Fzo1 would take place and head-to-head cis-dimerization through GTPase domains would generate competency for trans-oligomerization. In this configuration, trans interactions could take place through GTPase domains using an interface of association distinct from that used during cis-dimerization (figure 3.5a). Such a surface of interaction remains to be established experimentally. Moreover, the resulting Fzo1 trans-oligomer would impose a tethering distance of 32 nm between the two juxtaposed outer membranes, which is four-fold higher than the average 8 nm distance experimentally observed [37]. In the alternative configuration in which GTP binding would induce narrow opening of Fzo1, the most exposed surface of trans-interaction between head-to-head cis-dimers would lie in the trunk region. Anti-parallel associations reminiscent of those observed in the crystal structure of the Mfn1 HR2 domain [135] would take place between HR2 portions located in the trunk of dimers from opposing membranes (figure 3.5c). This would impose a tethering distance of about 10 nm, which is compatible with the cryo-ET experimental measures [37]. In the context of the back-to-back HR-parallel cis-dimers (figure 3.2c), GTP binding would induce the conformational switch of Fzo1, which would allow the GTPase domain to engage in trans association with the GTPase domain of a back-to-back cis-dimer from an opposing membrane (figure 3.5b). Although this Fzo1 trans-oligomer imposes a tethering distance

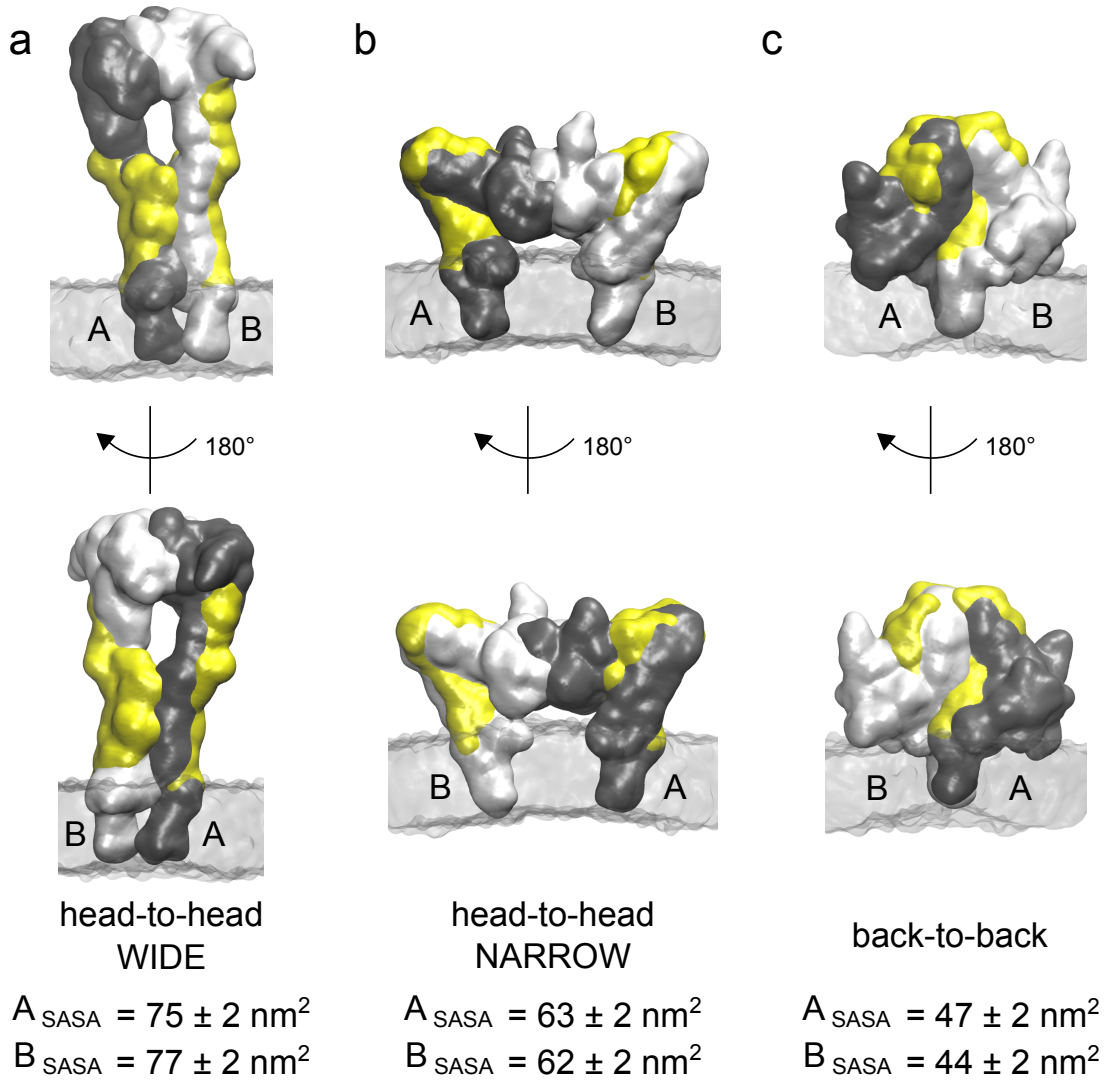

Fig. 3.4 Exposed Ugo1-interacting regions in Fzo1 putative cis-dimers. a., b. and c. Final snapshots from the simulations showing the Ugo1-interacting surfaces (yellow) for each of the cis-dimers. Each Fzo1 monomer is represented with a different colour (white or grey) and by a letter (A or B); the membrane bilayer is depicted as a silver transparent surface. The solvent accessible surface areas for the Ugo1-interacting regions corresponding to each monomer are reported at the bottom.

of 9 nm, which is compatible with experimental observations, an extensive manipulation of the initial closed conformation involving hinges 1a and 1b was required to generate the GTPase domain interface (Supplementary Fig. 3.10). Without this modification, the starting orientation of the GTPase domain would prevent the formation of the canonical G-interface observed in the dynamin superfamily [66] and Mfn1 dimers [46, 271].

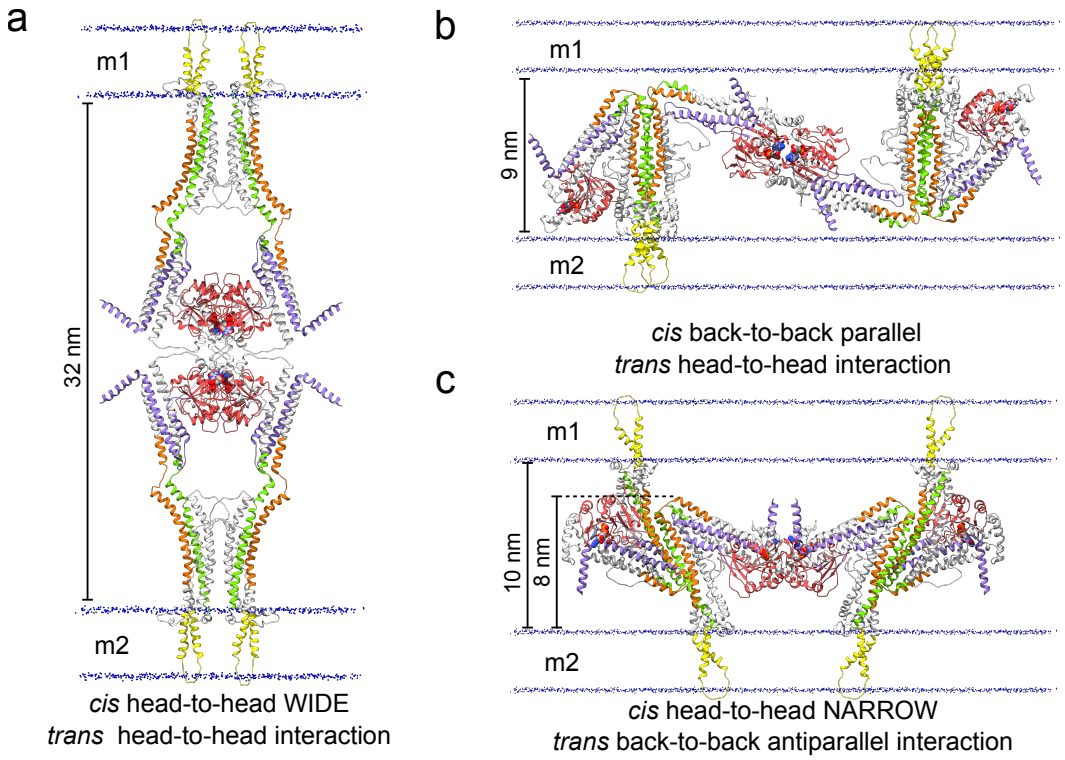

Fig. 3.5 Putative model complexes for Fzo1 trans-tethering interactions. a. The Fzo1 models in open conformations, as suggested from analogy to the BDLP system, interacting through their GTPase domains. b. The Fzo1 trans-tetramer in which the HRs are oriented in a parallel fashion and the trans interaction occurs towards the GTPase domain. c. The Fzo1 trans-tetramer in closed conformation. Here the interaction occurs towards the GTPase domain as well as through the respective HR domains oriented in an antiparallel fashion. Outer membranes corresponding to two different mitochondria are labelled m1 and m2. The domains highlighted in colour are: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (blue) from lipid bilayer headgroups and GDP nucleotide are depicted in the space-filling representation.

3.3.3 Trans-dimer dynamics

The robustness of the tetramers was tested through coarse-grained molecular dynamics simulations. The final conformations obtained for all three MD simulations show a structural drift with respect to their respective starting models that, although significant (Supplementary Fig. 3.14), does not point to major issues in the structure (figures 3.6 and 7.1). The models are thus stable in a membrane environment and can consequently be considered as plausible (see detailed analysis in supplementary material).

Analysing the interface area between protein and lipids revealed that the trans head-to-head model with parallel HR cis-interactions has a higher interaction surface with the

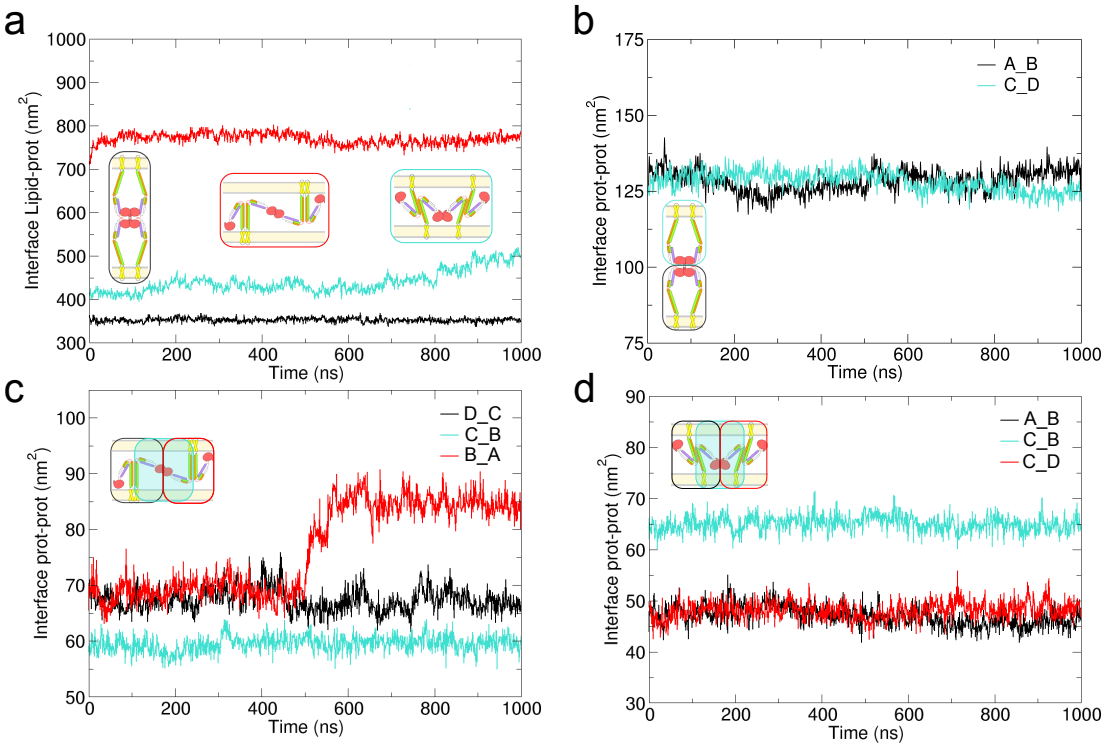

Fig. 3.6 Lipid-protein and protein-protein interface evolution during CG MD simulations for the tetrameric complexes. a. Lipid-protein interfaces for the three different studied tetrameric complexes (black: head-to-head trans open, red: head-to-head trans parallel, cyan: back-to-back trans antiparallel). b., c. and d. depict the evolution of the characteristic monomer-monomer interfaces belonging to the trans-tetramer in open conformation, the trans-tetramer in which the HRs are oriented in parallel and the trans-tetramer with HRs oriented in an antiparallel fashion. Legends A, B, C and D represent each of the monomers in the complex, following the colouring of the image inset.

membrane as compared to the other two tetramers (figure 3.6a, red line). In particular, the region including residues 380-386 in the GTPase domain of the back-to-back cis-dimer appears to interact strongly with membranes. For this model, we also observed a jump in the protein-protein interface in one of the two back-to-back cis-dimers that resulted from an increased interaction between the transmembrane regions of each monomer (figure 3.6c, red curve). The cis head-to-head trans antiparallel model showed an increase in membrane curvature during the simulation (figure 7.1c). This is explained by the strong interactions between the protein N-terminal regions with lipids, even though the first 100 residues of Fzo1 are missing from the models (figure 7.1c).

In addition, the region including residues 380-386 in the GTPase domain also contributed to the interaction with membranes. An analysis of the distance evolution between opposing

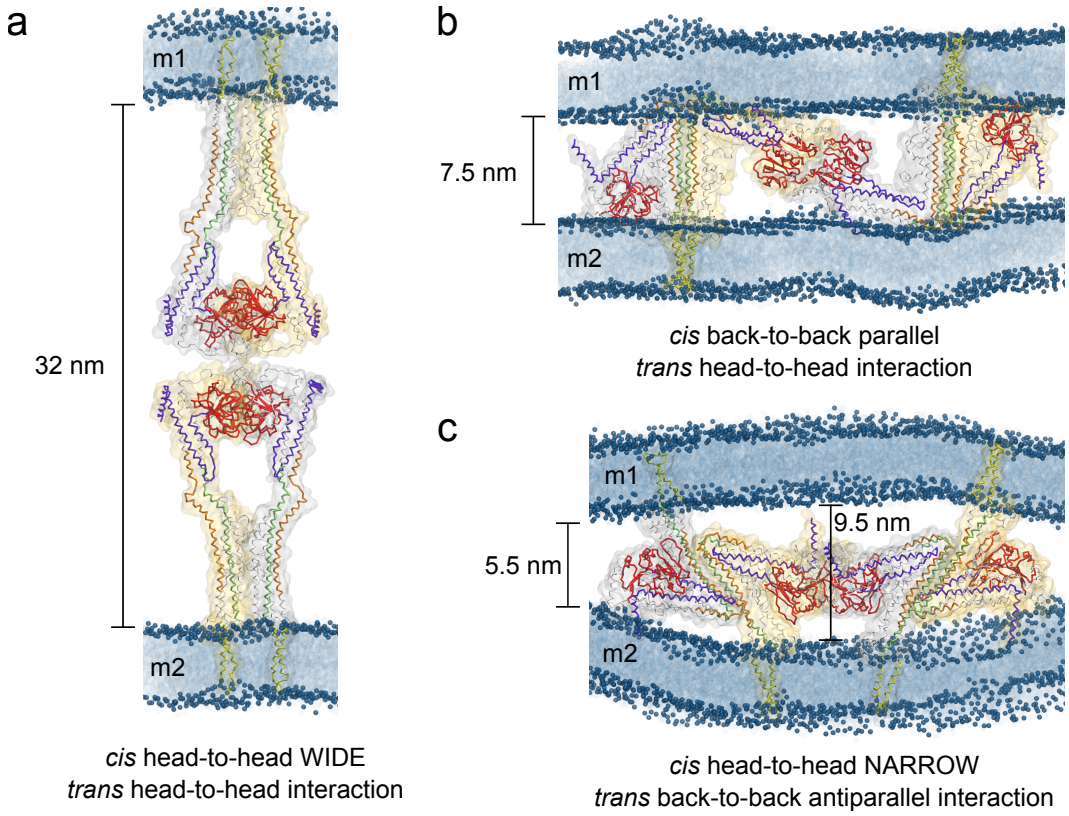

Fig. 3.7 Final configurations of the putative model complexes for Fzo1 trans-tethering interactions after relaxation through molecular dynamics simulations. a., the Fzo1 models in open conformation, as suggested from the BDLP system analogy, interacting through their GTPase domains. b., the Fzo1 trans-tetramer in which the HRs are oriented in a parallel fashion and the trans interaction occurs towards the GTPase domain. c., the Fzo1 trans-tetramer in closed conformation. Here, the interaction occurs towards the GTPase domain as well as through the respective HR domains oriented in an antiparallel fashion. In all figures the domains are coloured: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (dark blue) from the lipid headgroups of two mitochondria outer membranes (m1, m2) are depicted as spheres, whereas the tails are highlighted as a blue shadow. Intermembrane distances between the outer phosphate layer of the membranes are indicated.

membranes for all the tetrameric complexes showed slight changes during the MD run. Most notably, the initial distance between the membranes was 10 nm for the cis head-to-head trans anti-parallel interaction. After 1 μs, this value remained close only in the middle dimer (9.5 nm). At the extremes of the tetramer, however, the intermembrane distance stabilised at approximately 5.5 nm. This narrower value appeared as a consequence of the increased curvature in the membrane caused by the strong protein interactions with the lipid (figure

3.6c, Supplementary Fig. 3.11). We also investigated the evolution of the accessibility to the Ugo1 binding area upon formation of the tetramers (figure 3.8).

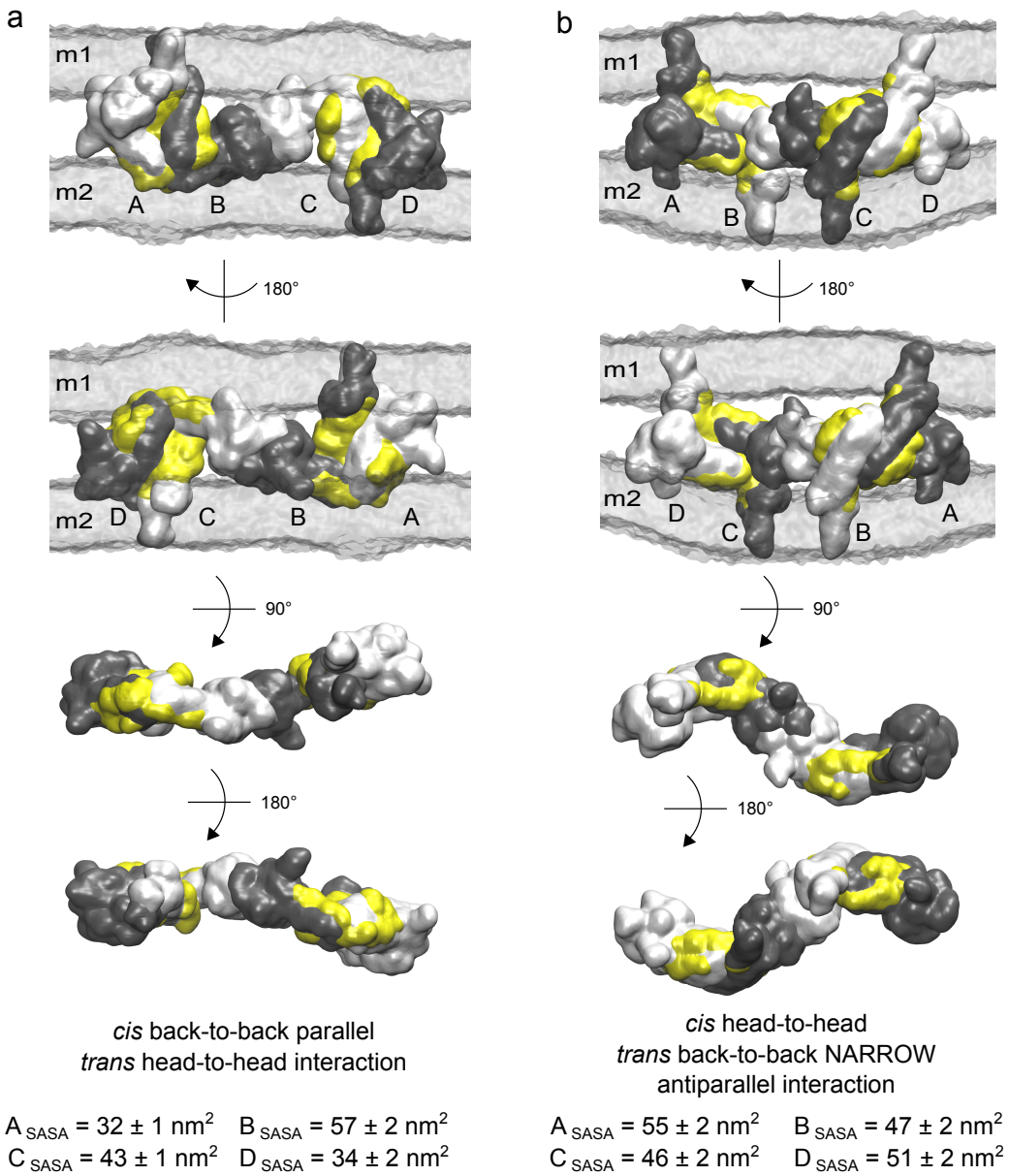

Fig. 3.8 Exposed Ugo1-interacting regions for Fzo1 trans-tethering complexes. a., b. and c. Final snapshots from the simulations showing the Ugo1-interacting surfaces (yellow) for each of the trans-tethering complexes. Each Fzo1 monomer is represented with alternate colours (white or grey) and by a letter (A, B, C or D); the membrane bilayer is depicted as a silver transparent surface. The solvent accessible surface area for the Ugo1-interacting regions corresponding to each monomer are reported at the bottom.

We did not include the BDLP-like model in this analysis as Ugo1 accessibility would obviously be unchanged upon trans-association of the dimers. In contrast, the Ugo1 accessibility after trans-association decreased in the cis head-to-head dimer (before 62-63 nm2; after 46-47 nm2) whereas it either increased or remained stable in the cis back-to-back dimer (before 44-47 nm2; after 43-57 nm2). Provided that Ugo1 interacts exclusively in cis with the Fzo1 trunk but not in trans with the four-helix bundle, these observations suggest two possible modes of Ugo1 recruitment. Formation of the back-to-back cis-dimer would not depend on Ugo1, which would only be recruited following head-to-head trans-interactions. Alternatively, Ugo1 may favour formation of the head-to-head cis dimer, before being released to allow back-to-back trans-interactions. Taking into consideration the involvement of Ugo1 in the coordination of outer and inner membrane fusion, its release during Fzo1 trans-oligomerization may facilitate its local interaction with Mgm1 to prepare inner membranes for subsequent homotypic docking. Further experimental data are clearly required to discriminate these distinct possibilities.

3.4 Conclusion

In analogy to the macromolecular spirals formed during DRP-mediated membrane fission (Jimah and Hinshaw, 2019), DRP-mediated membrane fusion may require the formation of macro-oligomers as observed with the Fzo1-dependent docking ring structure [37]. Such assemblies require building blocks as previously shown for fission DRPs that associate with their cofactors in the cytosol before recruitment to outer membranes and formation of the spirals [133, 142, 165]. In the context of mitofusins, these building blocks would form on the same membrane imposing the assembly of cis-oligomers. This rationale led us to build three independent cis-assembly Fzo1 models that gave rise to three independent trans-oligomeric models of mitofusin, which resulted in three distinct scenarios of Fzo1 oligomerization that are not connected to each other. Each scenario involves two main components of mitofusin auto-oligomerization: the GTPase domain interface [46, 271] and the trunk domain interface [135]. A plethora of experimental data demonstrates that the integrity of both regions is indeed essential for mitofusin-mediated mitochondrial fusion [57, 77, 101, 110, 113, 135, 207, 215]. If considering BDLP, the GTPase domain interface should trigger oligomerization in cis of Fzo1 [153]. Nonetheless, this interface is currently thought to trigger trans-oligomerization of mitofusins [46, 271]. Only one of our three final trans models employs this strategy. In the back-to-back cis-dimers (with parallel-oriented HRs), two dimers from opposing membranes would engage in trans interactions through their GTPase domains (figure 3.2c, and 3.5b). Yet, the generation of this trans-tetramer required

an extensive rotation of the GTPase domains around the axis of their four helix bundles to reconstitute the canonical G-interface where GTPase domains mirror each other as for other members of the dynamin superfamily [46, 66]. Moreover, a large area of the Ugo1 binding domain in each Fzo1 monomer was hindered by the HR interacting interface required to form the back-to-back cis-dimers (figure 3.4c). This model thus presents significant limitations upon confrontation to the experimental literature. Importantly, the G-interface of MFN1 was shown to promote membrane tethering, suggesting a trans-oligomerization function [46, 199]. However, this was observed with constructs of MFN1 artificially bound to membranes that lack the whole predicted trunk and transmembrane regions of mitofusins. The possibility that full-length mitofusins employ the G-interface for cis-oligomerization purposes must therefore not be excluded. The two remaining models employ such head-to-head cis-oligomerization properties. In this configuration, GTP-bound Fzo1 would undergo opening that could be narrow to wide. As seen for the stabilization of GTP-bound fission DRPs with their cofactors [142], this conformational switch around the hinges 1a and 1b would be stabilized by the concomitant binding of Ugo1 through the available trunk and a neighbouring mitofusin molecule through the activated GTPase domains. In both wide and narrow models, the canonical G-interface is maintained. From then, wide or narrow cis-oligomers would engage in trans-association with Fzo1 cis-complexes from the opposing membrane. The BDLP-like wide oligomers would associate in trans through a G-domain interface that is distinct from that employed during cis-dimerization. To our knowledge, such an interface has never been observed for other members of the dynamin superfamily, which casts strong doubts on its validity. Moreover, these Fzo1 trans-oligomers would impose a tethering distance between outer membranes of 32 nm, which is significantly larger than the distance observed experimentally between docked mitochondria (8 nm, [37]). Therefore, similar to the back-to-back system, the wide opening head-to-head model is difficult to reconcile with the literature. In contrast, the narrow opening head-to-head mechanism employs a trans-oligomerization strategy that turns out to fit particularly well with the past and current literature. After a possible release of Ugo1 (as suggested by figure 3.8), the trunk region in the cis-oligomers would expose a hydrophobic spine from its HR2 portion that would be available for putative interactions [70]. These interactions would take place in an anti-parallel fashion with an equivalent HR2 spine belonging to Fzo1 cis-oligomers from an opposing membrane. This would be consistent with anti-parallel HR2 interactions observed in 2004 for Mfn1 [135]. Moreover, the tethering distance of 5.5 to 9.5 nm imposed by this Fzo1 oligomerization model would agree with the 8 nm separation between outer membranes observed experimentally twelve years later [37]. The narrow opening head-to-head cis-oligomerization followed by the antiparallel back-to-back trans-associations system thus allows speculation on the

possible outcome of the fusion complex. Following GTP hydrolysis, hinges 1a and 1b might restore the displacement of the GTPase domain back to the trunk (i.e., a closed state). Similar to the dynamin BSE domain that may transfer energy from the GTPase domain to the trunk during DRP-mediated fission [49, 66], this movement could induce a sliding of interacting HRs with respect to each other. This would contribute to further reducing the distance between outer membranes to 2-3 nm as previously observed [37]. This cycle would reiterate around this initial region of minimal contact to reach the docking stage characterized by an extended area of membrane apposition delimited by the ring-shaped mitochondrial docking complex (MDC). It is tempting to propose that this MDC, which stands in regions where the distance between outer membranes reaches 6 to 8 nm [37], corresponds to a macromolecular assembly of Fzo1 oligomers reminiscent of those described in the present study. In addition, our models provide sufficient detail to derive experimentally testable hypotheses for future work, for instance in terms of possible crosslinks and surface-exposed residues. Similarly, our models could be further improved based on new experimental observations. Regardless of these possibilities, we hope that our hypotheses will provide food for thoughts before elucidating the structure of full length mitofusins and their known or yet to be described cofactors, as well as their precise oligomerization properties.

3.5 Supplementary Material associated to this publication

The following figures and tables are the supplementary data related to this article.

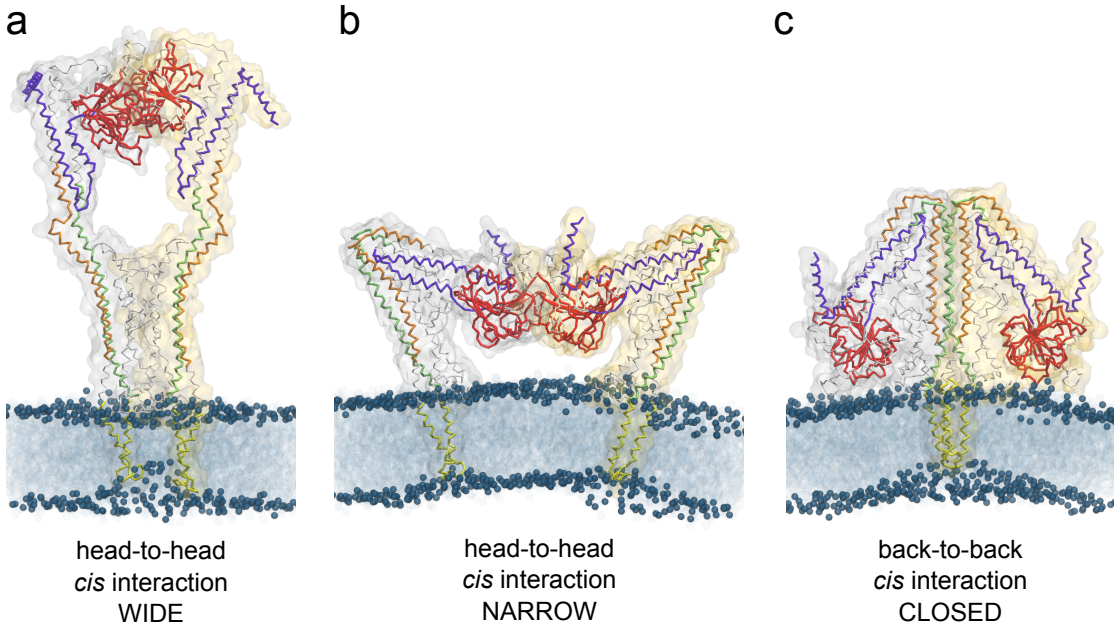

Fig. 3.9 Putative membrane-inserted Fzo1 models interacting in cis configuration after molecular dynamics relaxation. a and b. Wide and narrow head-to-head complexes, respectively. c. Back-to-back complex of the closed model. The domains are coloured: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (dark blue) from the lipid bilayer headgroups are depicted as spheres, whereas the bilayer is highlighted as a blue shadow.

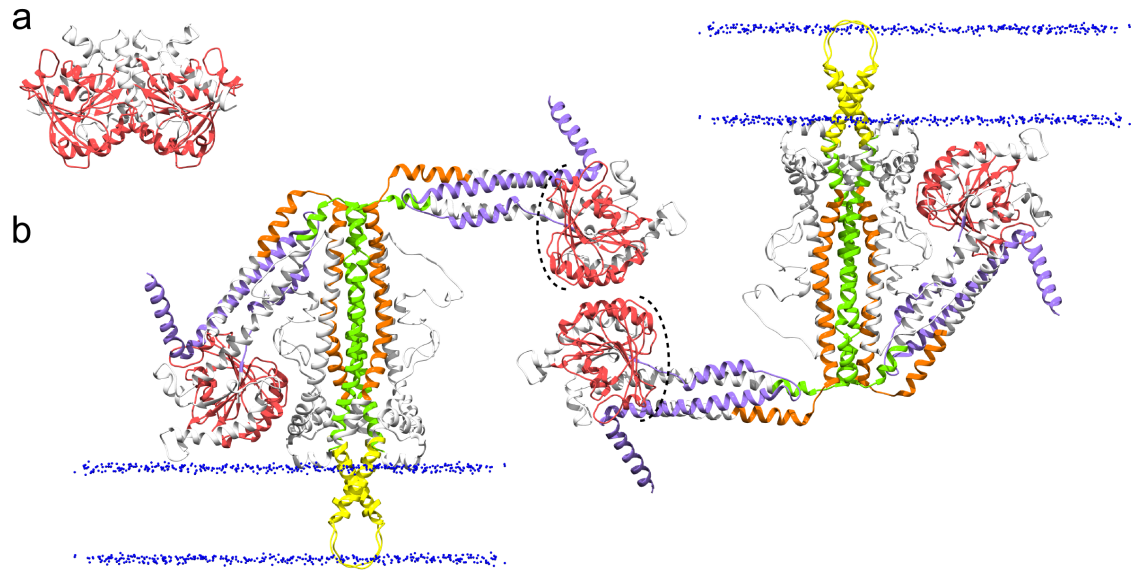

Fig. 3.10 a. The Fzo1 GTPase dimer construct in which the canonical G-interface commonly observed in the dynamin superfamily [66] is reconstituted. b. The back-to-back HR-parallel trans-tetramer during the modeling procedure (see Method). The Figure shows how the two GTPase domains that interact in trans share an apparent incorrect orientation (compared with a.). Without an extensive conformational rearrangement that would involve both hinges 1a and 1b, the formation of the canonical G-interface observed in the dynamin superfamily [66] would be impeded. In particular, the dotted line should be located on the same side. The domains are: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Phosphorus atoms (blue) from lipid bilayer headgroups and GDP nucleotide are depicted in the space-filled representation.

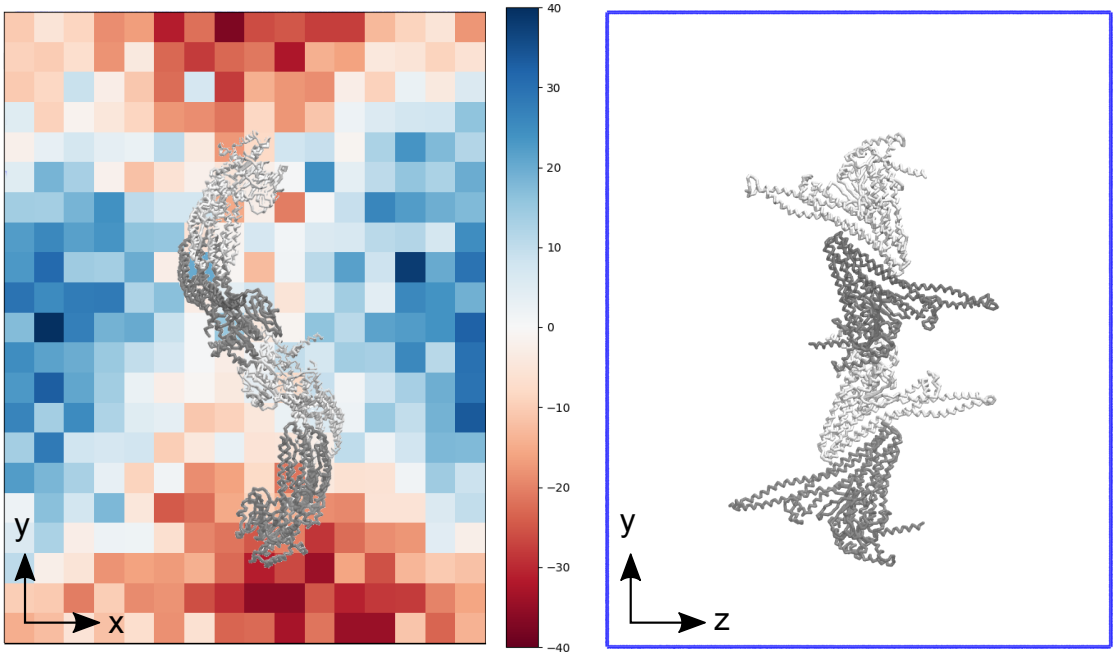

Fig. 3.11 Inter-bilayer difference distance matrix (Å). Left panel shows the final difference in distances compared to the initial structure, red regions correspond to regions that got closer in space compared to the initial structure of the production run whereas blue regions represent areas that moved further apart. The right panel shows a rotation of 90 degrees around the y-axis to visualize the position of the complex.

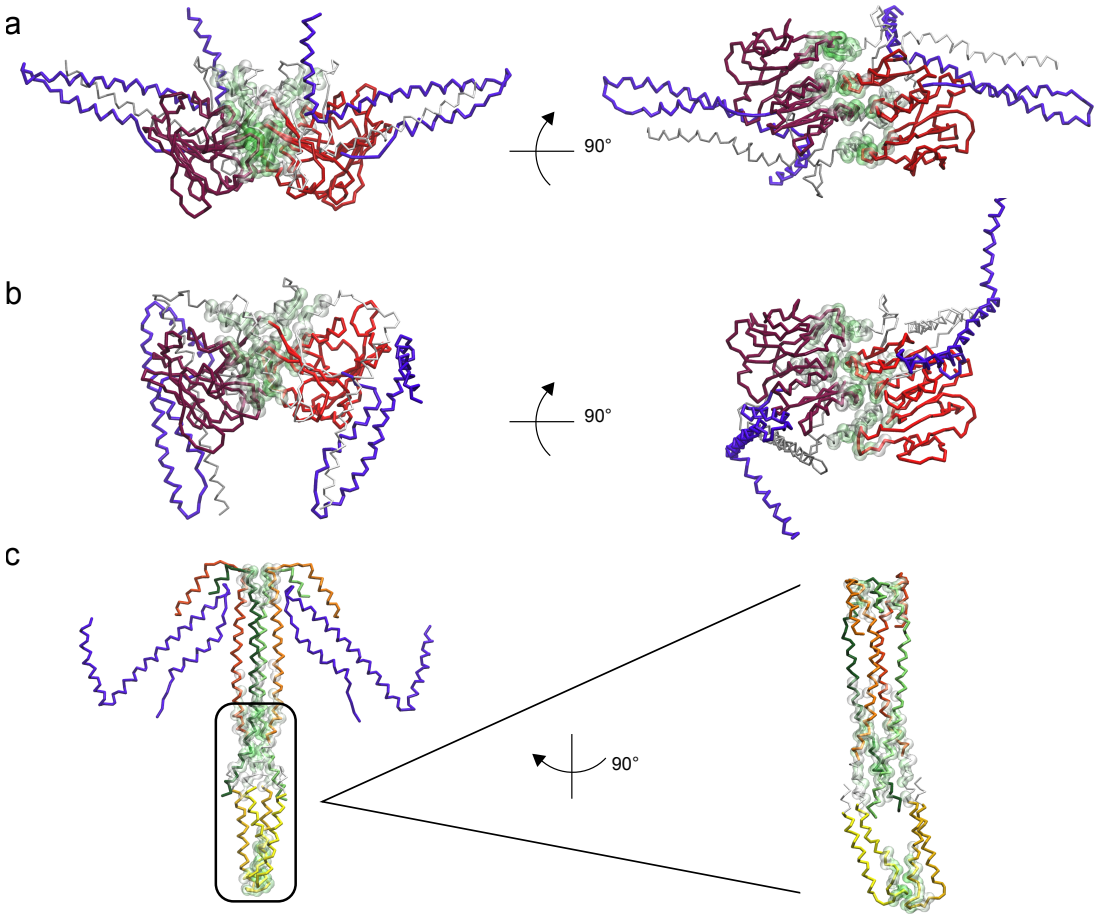

Fig. 3.12 Interface details from the cis-dimers. a. Narrow head-to-head complex, b. Wide head-to-head complex. c. Back-to-back complex. The domains are coloured in: violet, HRN; green, HR1; orange, HR2; red, GTPase and yellow, transmembrane. Darker shade of the same colours are used to differentiate between monomers when needed. The residues in contact between both monomers are highlighted with a transparent surface coloured by occupancy with a green gradient (grey: minimum value, green: maximum value). For clarity, only the regions corresponding to the interface are shown.

	Box size (x,y,z [nm])	Number of particles	
		All	Protein + lipid
Dimers			
Head-to-head wide	25.1474 x 25.1474 x 28.9065	157017	27418
Head-to-head narrow	24.6357 x 24.6357 x 29.8118	156752	27058
Back-to-back	25.2349 x 25.2349 x 28.7671	157232	27298
Tetramers			
Head-to-head *trans*	29.4832 x 29.4832 x 67.4888	523921	78836
Head-to-head *trans* parallel	49.4896 x 39.5917 x 28.3762	497044	166292
Head-to-head *trans* antiparallel	30.4625 x 40.6166 x 32.2698	347636	101324

Fig. 3.13 Molecular dynamics simulation details for each system.

		Figure 2a		Figure 2b		Figure 2c	
		RMSD (nm)	std (nm)	RMSD (nm)	std (nm)	RMSD (nm)	std (nm)
Full protein complex BB RMSD		0.860	0.068	0.401	0.039	0.469	0.013
chain A	all	**0.567**	**0.075**	**0.332**	**0.029**	**0.330**	**0.016**
	frag1	0.261	0.099	0.145	0.036	0.190	0.060
	frag2	0.179	0.013	0.129	0.007	0.233	0.023
	frag3	0.168	0.022	0.137	0.014	0.101	0.012
	frag4	0.431	0.034	0.350	0.021	0.386	0.013
	frag5	0.177	0.040	0.150	0.024	0.132	0.026
chain B	all	**0.688**	**0.062**	**0.284**	**0.025**	**0.463**	**0.025**
	frag1	0.388	0.052	0.290	0.069	0.363	0.094
	frag2	0.169	0.010	0.155	0.011	0.204	0.023
	frag3	0.147	0.023	0.101	0.010	0.109	0.011
	frag4	0.397	0.030	0.266	0.026	0.590	0.030
	frag5	0.243	0.038	0.164	0.024	0.101	0.012
		Figure 3a		Figure 3b		Figure 3c	
		RMSD (nm)	std (nm)	RMSD (nm)	std (nm)	RMSD (nm)	std (nm)
Full protein complex BB RMSD		0.637	0.065	0.303	0.006	0.905	0.068
chain A	all	**0.408**	**0.037**	**0.220**	**0.013**	**0.397**	**0.035**
	frag1	0.236	0.069	0.123	0.010	0.256	0.067
	frag2	0.161	0.011	0.194	0.009	0.185	0.012
	frag3	0.171	0.023	0.095	0.011	0.164	0.017
	frag4	0.335	0.022	0.231	0.028	0.420	0.043
	frag5	0.188	0.033	0.089	0.010	0.153	0.022
chain B	all	**0.364**	**0.030**	**0.353**	**0.014**	**0.283**	**0.018**
	frag1	0.500	0.041	0.351	0.019	0.143	0.037
	frag2	0.197	0.012	0.136	0.006	0.135	0.012
	frag3	0.141	0.023	0.131	0.012	0.097	0.009
	frag4	0.262	0.020	0.390	0.022	0.333	0.021
	frag5	0.266	0.042	0.094	0.013	0.119	0.015
chain C	all	**0.513**	**0.055**	**0.243**	**0.011**	**0.528**	**0.036**
	frag1	0.442	0.112	0.144	0.011	0.249	0.035
	frag2	0.179	0.011	0.155	0.005	0.174	0.015
	frag3	0.180	0.030	0.129	0.014	0.201	0.016
	frag4	0.338	0.021	0.272	0.013	0.421	0.031
	frag5	0.179	0.048	0.162	0.019	0.210	0.021
chain D	all	**0.414**	**0.037**	**0.245**	**0.010**	**0.310**	**0.029**
	frag1	0.361	0.051	0.111	0.009	0.190	0.050
	frag2	0.192	0.010	0.185	0.011	0.193	0.016
	frag3	0.177	0.029	0.104	0.009	0.112	0.011
	frag4	0.362	0.037	0.289	0.014	0.320	0.031
	frag5	0.147	0.027	0.092	0.012	0.131	0.016

Fig. 3.14 RMSD values averaged over the last 100 ns of simulation for each dimer (Fig. 2a, b, c) and tetramers (Fig. 3a, b, c). Reported values include a full RMSD calculated over all the backbone particles (BB) as well as by every chain and structural fragment separately.

ch	resid		occup
A	ASP	195	3.66
A	VAL	196	4.86
A	ASN	197	4.05
A	THR	198	2.15
A	GLY	199	0.77
A	ALA	202	0.69
A	GLU	215	2.43
A	ASP	216	4.35
A	GLN	217	5.33
A	LEU	218	2.34
A	PRO	219	0.74
A	ILE	267	1.39
A	HIS	268	4.50
A	THR	269	6.88
A	LEU	270	9.00
A	LYS	271	7.34
A	GLU	272	5.65
A	LEU	273	5.61
A	PRO	274	3.20
A	ASP	275	0.39
A	GLY	316	0.41
A	LEU	317	0.92
A	VAL	341	0.86
A	ASN	342	2.00
A	ALA	343	1.05
A	GLU	344	6.31
A	ASN	345	7.34
A	GLN	346	3.82
A	LEU	347	1.07
A	THR	348	4.77
A	LEU	349	3.36
A	SER	350	3.03
A	ALA	351	0.35
A	ASP	373	0.50
A	ASP	377	2.09
A	LYS	378	3.61
A	GLN	379	1.82
A	ARG	380	1.02
A	LYS	387	0.38
A	ARG	390	2.45
A	ASP	391	3.00
A	LEU	392	0.93
A	SER	393	1.85
A	PRO	394	4.40
A	GLU	395	0.97
A	THR	396	0.52
A	TYR	397	3.79
A	LYS	398	3.03
A	ARG	399	2.38

ch	resid		occup
B	ASP	195	2.40
B	VAL	196	4.95
B	ASN	197	4.05
B	THR	198	2.35
B	GLY	199	0.98
B	GLU	215	1.83
B	ASP	216	4.09
B	GLN	217	4.73
B	LEU	218	1.07
B	HIS	268	1.19
B	THR	269	5.54
B	LEU	270	6.10
B	LYS	271	6.00
B	GLU	272	1.18
B	LEU	273	0.85
B	PRO	274	1.15
B	LEU	317	0.31
B	ASN	342	1.98
B	ALA	343	1.98
B	GLU	344	7.11
B	ASN	345	8.59
B	GLN	346	3.20
B	LEU	347	2.60
B	THR	348	3.24
B	LEU	349	3.58
B	SER	350	2.99
B	ASP	373	0.23
B	ASP	377	2.71
B	LYS	378	3.22
B	GLN	379	2.05
B	ARG	380	3.58
B	LYS	387	0.69
B	ARG	390	3.06
B	ASP	391	3.09
B	LEU	392	2.15
B	SER	393	5.99
B	PRO	394	6.41
B	GLU	395	3.12
B	THR	396	3.18
B	TYR	397	7.16
B	LYS	398	6.90
B	ARG	399	3.94
B	ALA	400	4.12
B	ALA	401	5.39
B	ASP	402	3.33
B	ASP	428	2.74
B	ARG	429	1.81
B	PRO	431	0.71
B	ASP	432	2.76

ch	resid		occup
A	ALA	400	2.75
A	ALA	401	4.83
A	ASP	402	2.15
A	HIS	426	0.89
A	GLY	427	1.43
A	ASP	428	4.00
A	ARG	429	1.07
A	PRO	431	1.17
A	ASP	432	1.37

Fig. 3.15 Interface residues for the narrow head-to-head cis-dimer by chain (ch) residue number (resid) according to UNIPROT as shown in previous figures. Occupancy (occup) values are calculated on the last 100 ns of the simulations.

ch	resid		occup
A	ASP	195	3.08
A	VAL	196	4.63
A	ASN	197	3.95
A	THR	198	2.38
A	GLY	199	0.89
A	GLU	215	1.92
A	ASP	216	3.55
A	GLN	217	3.77
A	LEU	218	0.87
A	HIS	268	0.63
A	THR	269	3.61
A	LEU	270	2.35
A	LYS	271	2.55
A	GLU	272	3.39
A	LEU	273	1.85
A	PRO	274	0.40
A	LEU	317	0.32
A	ASN	342	0.20
A	ALA	343	0.36
A	GLU	344	4.21
A	ASN	345	5.47
A	GLN	346	2.29
A	LEU	347	1.27
A	THR	348	4.67
A	LEU	349	2.72
A	SER	350	2.73
A	ALA	351	0.41
A	LYS	371	0.94
A	ASP	373	1.16
A	LYS	374	0.86
A	ASP	377	1.11
A	LYS	378	1.87
A	GLN	379	2.95
A	ARG	380	2.62
A	ARG	390	1.64
A	ASP	391	0.85
A	SER	393	1.99
A	PRO	394	4.02
A	GLU	395	0.96
A	TYR	397	2.52
A	LYS	398	1.15
A	ALA	400	1.88
A	ALA	401	5.01
A	ASP	402	3.27
A	VAL	404	1.30
A	HIS	405	1.00
A	HIS	426	0.65
A	GLY	427	0.63

ch	resid		occup
B	ASP	195	2.38
B	VAL	196	5.02
B	ASN	197	3.43
B	THR	198	2.02
B	GLY	199	0.73
B	GLU	215	1.84
B	ASP	216	3.40
B	GLN	217	4.06
B	LEU	218	1.51
B	THR	269	3.61
B	LEU	270	1.60
B	LYS	271	4.07
B	GLU	272	2.95
B	LEU	273	2.36
B	PRO	274	0.50
B	LEU	317	0.27
B	ASN	342	0.95
B	ALA	343	1.31
B	GLU	344	4.29
B	ASN	345	4.97
B	GLN	346	2.89
B	LEU	347	1.42
B	THR	348	3.23
B	LEU	349	3.54
B	SER	350	2.95
B	ASP	373	1.84
B	LYS	374	0.58
B	ASP	377	3.44
B	LYS	378	1.63
B	GLN	379	1.59
B	ARG	380	2.02
B	CYS	381	0.30
B	ARG	390	1.53
B	ASP	391	2.33
B	SER	393	1.56
B	PRO	394	3.41
B	GLU	395	0.81
B	TYR	397	1.27
B	LYS	398	1.21
B	ALA	400	1.13
B	ALA	401	5.07
B	ASP	402	4.02
B	PHE	403	0.90
B	VAL	404	1.31
B	ASP	428	2.82
B	ARG	429	2.45
B	PRO	431	0.45
B	ASP	432	1.87

ch	resid		occup
A	ASP	428	3.46
A	ARG	429	0.28
A	PRO	431	0.62
A	ASP	432	1.56
A	LEU	549	0.46
A	HIS	550	1.77
A	LYS	590	2.09
A	LEU	593	2.02
A	LEU	594	0.27
A	LYS	597	0.33
A	ILE	601	0.66
A	ASN	604	0.95
A	GLU	605	0.73
A	GLY	607	0.31
A	ASP	608	0.84
A	MET	611	1.21
A	CYS	612	3.30
A	GLU	613	2.74
A	ARG	614	1.33
A	VAL	615	3.72
A	PHE	616	3.93
A	ARG	617	0.91
A	PRO	645	0.60
A	THR	646	3.18
A	TRP	647	0.94
A	GLY	649	0.53
A	PHE	650	1.74
A	LEU	678	0.74
A	MET	679	0.89
A	LYS	680	2.22
A	TYR	681	2.60

ch	resid		occup
B	SER	510	0.48
B	VAL	547	0.87
B	HIS	550	1.12
B	LYS	590	0.92
B	LEU	593	2.17
B	LEU	594	0.49
B	LYS	597	1.06
B	GLU	600	0.36
B	ASN	604	1.25
B	GLU	605	3.35
B	LEU	606	0.34
B	ASP	608	2.02
B	ASP	609	1.19
B	MET	611	1.37
B	CYS	612	2.79
B	GLU	613	0.92
B	VAL	615	1.49
B	PHE	616	1.68
B	THR	646	0.26
B	TRP	647	3.23
B	LYS	648	5.43
B	GLY	649	4.16
B	PHE	650	2.11
B	LEU	651	0.77
B	SER	652	0.22
B	TYR	681	0.68

Fig. 3.16 Interface residues for the wide head-to-head cis-dimer by chain (ch) residue number according to UNIPROT (resid) as shown in previous figures. Occupancy (occup) values are calculated on the last 100 ns of the simulations.

ch	resid		occup
A	ARG	803	0.24
A	LEU	728	0.24
A	SER	804	0.26
A	ALA	540	0.29
A	LYS	736	0.29
A	TYR	769	0.32
A	ILE	808	0.45
A	VAL	741	0.60
A	GLU	775	0.63
A	GLN	524	0.63
A	LEU	734	0.67
A	ASN	543	0.71
A	MET	809	0.74
A	ILE	742	0.77
A	ALA	526	0.78
A	LYS	496	0.82
A	GLU	806	0.98
A	LEU	776	1.00
A	PRO	758	1.00
A	LEU	781	1.02
A	LEU	765	1.02
A	LYS	538	1.02
A	GLU	499	1.08
A	TYR	778	1.10
A	LYS	811	1.64
A	MET	763	1.64
A	LEU	737	1.67
A	GLU	494	1.73
A	SER	738	1.86
A	GLU	539	1.95
A	MET	532	1.97
A	GLN	774	2.00
A	ILE	530	2.01
A	VAL	743	2.06
A	ASN	536	2.72
A	ASP	777	2.81
A	ASP	523	2.88
A	LYS	770	2.92
A	TYR	534	2.97
A	PRO	762	3.04
A	ASN	498	3.04
A	ILE	495	3.13
A	SER	766	3.25
A	ILE	807	3.38
A	GLN	528	3.47
A	SER	733	3.48
A	SER	729	3.71
A	ASP	810	3.82

ch	resid		occup
B	GLU	494	1.48
B	ILE	495	3.49
B	LYS	496	0.85
B	ILE	497	0.50
B	ASN	498	3.54
B	GLU	499	1.08
B	GLU	502	0.60
B	ASP	523	1.22
B	ALA	526	0.96
B	GLU	527	5.61
B	GLN	528	3.69
B	ILE	530	2.32
B	THR	531	4.95
B	MET	532	2.25
B	TYR	534	3.30
B	ASN	535	6.66
B	ASN	536	2.99
B	LYS	538	1.03
B	GLU	539	1.66
B	ASN	543	0.46
B	SER	729	0.76
B	SER	730	0.93
B	TRP	731	1.25
B	SER	732	3.52
B	SER	733	3.32
B	LEU	734	4.20
B	LYS	735	3.82
B	LYS	736	2.89
B	LEU	737	2.37
B	SER	738	3.59
B	VAL	739	7.48
B	PRO	740	3.67
B	VAL	741	1.68
B	ILE	742	2.46
B	VAL	743	4.13
B	SER	746	0.65
B	PRO	758	1.56
B	PRO	762	3.03
B	MET	763	1.44
B	LEU	765	1.00
B	SER	766	3.94
B	TYR	769	1.42
B	LYS	770	2.81
B	ARG	771	0.34
B	LEU	773	3.08
B	GLN	774	2.64
B	GLU	775	0.27
B	LEU	776	0.93

ch	resid		occup
A	LEU	773	4.05
A	SER	730	4.12
A	PRO	740	4.73
A	VAL	739	5.08
A	GLU	527	5.70
A	THR	531	5.99
A	ASN	535	6.61
A	TRP	731	7.07
A	SER	732	10.66

ch	resid		occup
B	ASP	777	2.85
B	TYR	778	1.81
B	LEU	781	0.96
B	GLU	806	1.00
B	ILE	807	2.23
B	MET	809	1.67
B	ASP	810	3.96
B	LYS	811	1.04

Fig. 3.17 Interface residues for the back-to-back cis-dimer by chain (ch) residue number according to UNIPROT (resid) as shown in previous figures. Occupancy (occup) values are calculated on the last 100 ns of the simulations.

Chapter 4

Adding hydrodynamics to lipids

This chapter presents the work related to the coupling of hydrodynamic interactions to a lipid coarse-grained force field as currently submitted [36]. We present an overview of different applications where the section regarding the Nanoreactor was done in cooperation with Stepan Timr in the context of the Big Challenge grant. My contribution to that section was to prepare the initial equilibrated vesicle and give the coupling parameters between protein and lipids. Because of access to computational resources, the simulations run under shear conditions where done in cooperation with Fabio Sterpone (prepared, post-processed and analysed by me).

4.1 Introduction

The study of biomembranes based on experimental, theoretical and computational approaches occupies a central place in modern biophysics. These structures constituted by lipids of different types as well as other molecules (e.g., proteins, sterols and carbohydrates), allows the structural organisation of living matter by creating the outer envelope of cells and its intracellular compartments [258]. A biomembrane is a highly heterogeneous *milieu* where the different types of lipids and membrane proteins, which can account for nearly 50% of the total mass, are partitioned non-uniformly, giving rise to domains of varying structural order and internal dynamics [243]. Being more than mere protective walls, biomembranes form a highly dynamic environment hosting numerous essential metabolic and signalling processes. Furthermore, lipid bilayers that mimic cellular membranes are used by bioengineers as a basis for novel drug-delivery constructs and artificial organelles [41, 100]. Only recently have we started to grasp the full complexity of these lipid environments, giving rise to the field of lipidomics [226].

Nowadays, a great variety of experimental techniques allows characterising many aspects of the structure and dynamics of biomembranes, e.g. the extension of membrane domains [212], lipid diffusion [95], protein localisation [144], membrane deformation and fusion [227]. Moreover, membrane constructs such as vesicles are studied experimentally for their transport and content release capabilities [195, 248]. Their behaviour under flow has been studied to assess deformations [71] and mechanical-force-induced stress effects [132].

A strong support to understanding membrane systems is provided by computer modelling and simulation. During the last 30 years, the development of atomistic force fields combined with the availability of massive computer power have boosted the investigation of systems of growing size and complexity[155]. This evolution has made it possible to capture processes ranging from membrane permeation of small solutes [228], pore formation [130] up to the dynamics of a whole respiratory complex [222]. While the popularity of atomistic simulations of membrane systems is benefiting from improvements in algorithms and code efficiency, including the use of enhanced sampling techniques [118, 179], the length- and time scales associated with many membrane-related phenomena require the use of coarse-grained (CG) models [47, 106, 159, 160, 176]. A large body of biophysical processes has been investigated via CG models as reported in many in-depth review articles[19, 155]. Some examples include the study of mobility of lipids and proteins in membranes [50, 97, 262], membrane protein aggregation [9] and clustering [260], and vesicle fusion [10].

The success of CG models in exploring very complex processes relies on the substantial reduction in the number of degrees of freedom to handle computationally. Moreover the smoothness of the CG energy landscape enables the use of significantly increased integration timesteps. In this regard, implicit solvent models [10, 59, 63, 237] are particularly attractive, but, by default, they cannot describe the effect of dynamical correlations mediated by the solvent. However, these correlations not only play an important role in the aggregation process of lipids [7], and other biomolecules [54], but may critically influence biomolecule migration in the proximity of membranes [202] as much as vesicle mobility and encounters [209]. In the "Modelling Manifesto" put forward by Lyman, Hsieh and Eggeling[154] this aspect was carefully discussed.

Here we present a coupling of Dry Martini [10], an implicit-solvent coarse-grain lipid model, with the Lattice Boltzmann Molecular Dynamics (LBMD) technique [2, 3, 20, 239] to include hydrodynamic interactions (HI) in the implicit-solvent simulation of membranes. This work represents an alternative to a recently published coupling of Dry Martini to the multi-particle collision method [274] and extends our previous work where the OPEP implicit solvent model was coupled to HI [51, 54, 55, 143, 239, 240]. We show here that this coupling represents a general tool to treat the impact of HI in a broad range of processes involving

lipids, not only standard bilayer membranes, but also, for instance, to investigate the dynamics of a vesicle under the perturbative action of the external fluid flow. This technique will allow synergic interactions with experiments where lipidic systems are manipulated in fluid flows [11, 168]. In fact, the microscopic character of Dry Martini and the efficiency and the natural multi-scale character of LB to treat a fluid makes the LBMD an appealing method. LB can also handle fluids of different nature such as charged species in water in order to solve electrokinetic problems[28]. This feature, already tested for instance to characterise DNA translocation in pores[65], can be extended to investigate voltage-dependent processes at the proximity of neuronal membrane systems.

4.2 Methods

In this section we detail the implementation of the Dry Martini force field and its coupling with the Lattice Boltzmann Molecular Dynamics technique. An overview of the setup of the systems used in this work is presented in the last part of the section.

4.2.1 Dry Martini force field

Dry Martini is an implicit-solvent coarse-grained force field for lipids [10] derived from the original explicit solvent Martini force field [67, 160]. It is based on the same mapping with respect to an atomistic description of the lipids: approximately 4 atoms form a CG bead. The force field consists of bonding and non-bonding interaction terms where the former include only bond and angular potentials while the latter include van der Waals and electrostatic interactions. The van der Waals interactions are described by standard Lennard-Jones (LJ) potentials having the parameters re-calibrated with respect to the original Martini model in order to account for the lack of explicit solvation. Similarly some parameters of the bond and angular terms were re-parametrised to improve the properties of a simulated membrane bilayer. The electrostatic interactions are treated via a screened potential that is short-range in nature.

In order to properly describe the implementation of the Dry Martini model in our simulation package, Muphy [26], two features of the force field must be highlighted. First, concerning the angular term of the Hamiltonian, Dry Martini uses a very soft potential of the form $U_\theta = \frac{1}{2}K_\theta \cos^2(\theta - \theta_0)$, which is different from standard harmonic ones. The second aspect concerns electrostatic interactions: instead of considering a standard long-range $1/r$ potential, the interactions between charged groups are described by a screened short-range potential. The screening is encoded in a polynomial switch function acting on the potential

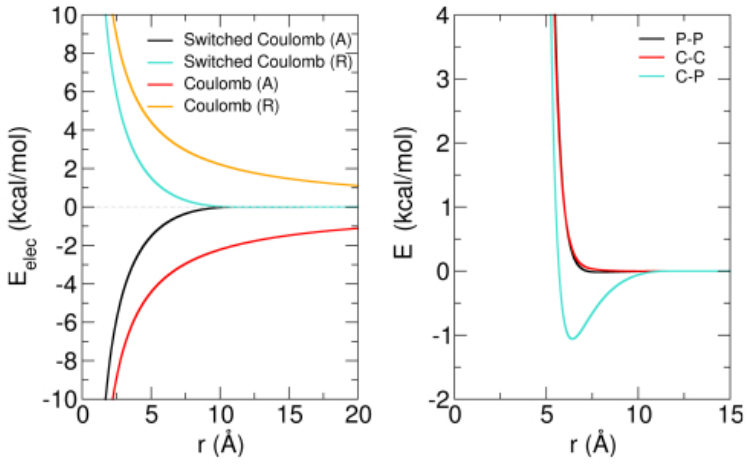

Fig. 4.1 Left panel. Attractive and repulsive Coulomb potential with relative dielectric constant $\varepsilon = 15$ compared to the switched electrostatic potentials used in the Dry Martini force field to describe the electrostatic interactions among charged groups. Right panel. Tabulated potentials that incorporate the electrostatic switched interactions and the short-range LJ interactions for the phosphate (P) and choline (C) charged groups of the lipids.

from $r = 0$ up to the cut-off r_c set equal to 12 Å. The comparison between standard and switched Coulombic potentials is provided in the left panel of Figure 4.1. The screening is intended to account for the lack of explicit solvation and ionic strength in the implicit solvent description [10]. In our implementation, we use tabulated potentials for the interactions between charged groups, adding together the screened electrostatic and the LJ terms (see right panel in Figure 4.1).

4.2.2 Lattice Boltzmann MD

In LBMD, the particle dynamics is coupled to the kinetic representation of the solvent, simulated via the Lattice Boltzmann (LB) technique [245]. This coupling allows to naturally include hydrodynamics in the simulation of biomolecules represented via implicit-solvent models. The coupling between particles and solvent arises from a Stokes-like drag force acting on each particle:

$$\vec{F}_i^D = -\gamma(\vec{v}_i - \vec{\tilde{u}}_i) \tag{4.1}$$

where $\vec{v}_i$ is the ith particle velocity, $\vec{\tilde{u}}_i$ is the fluid velocity $\vec{u}$ smeared over a finite extension of the i-th particle, and γ is the frictional coupling, an adjustable parameter in the methodology. The drag force adds up to the usual conservative forces derived from the Hamiltonian of the system, $\vec{F}_i^C = -\vec{\nabla}_i U(\{r\})$ and to a random white noise, $\vec{F}_i^R$, that represents thermal fluctuations. The reader can find more technical details in Ref. [54].

In our scheme, the LB implementation uses the BGK (Bhatnagar-Gross-Krook) collisional operator [20], with a lattice grid spacing of 3 Å, sufficient to resolve local hydrodynamic interactions for macromolecular systems. In our previous works [54, 239] we have provided a detailed discussion about the optimal size of the grid spacing to achieve an efficient coupling of LB with CG bio-molecule representations. The LB kinematic viscosity ν_0 was set to reproduce bulk water behaviour at ambient conditions. In these simulations the molecular and fluid dynamics were evolved synchronously using a physical time-step of $10\,fs$, a value dictated by the stiffness of the molecular forces. In some simulations, we explored the effect of different values of the friction coefficient γ in order for instance to tune protein diffusion in crowded conditions. Moreover when exploring the size effect on lipid diffusion in membranes we performed the simulations using a multiple-timestep scheme. In the present implementation all the simulations were performed using a constant volume/constant temperature scheme (NVT).

4.2.3 Systems and simulation setups

All simulations were performed using the code Muphy [26]. For verification purposes, comparisons with the original Dry Martini implementation in Gromacs [1, 256] were performed. The set-up of the studied systems is provided below.

Bilayers. The POPC lipid bilayers were generated using the tool Insane provided with Martini [266]. A total of 10^5 steps of steepest descent minimisation followed by NVT equilibration lasting $50\,ns$ were performed with the Gromacs software (v 4.5.5). The equilibrated systems were used to start the LBMD simulations. In order to study system size dependence of the diffusivity of the lipid molecules in the membrane we have assembled several bilayers with lateral sizes $L_x = L_y$ ranging from 100 Åup to 400 Å. All assembled systems have a surface area per lipid of 69.4 $Å^2$.

Nanotube. To simulate the adsorption of a carbon nanotube into a membrane we followed the approach designed by Bhaskara *et al.*[29] based on the Martini force field. The tube had a diameter of 15 Åand was assembled by stacking 11 "carbon" rings each constituted by 11 beads. The distal rings were formed by hydrophilic particles while the inner rings were formed by hydrophobic beads. The total length of the nanotube was $L = 43$ Å.

Vesicles and Nanoreactors. The vesicles simulated in this work were built with the help of the Martini vesicle maker utility in the CHARMM-GUI webserver [198] using the Dry Martini force field and equilibrated in Gromacs (v 5.1.4) following the CHARMM-GUI protocol. We built three systems of increasing size with initial vesicle radius equal to $r = 50$ Å, 60 Åand 160 Å, respectively (the radius is defined from the vesicle centre up to the centre of the lipid bilayer).

The smallest system was used to investigate the dynamics of a vesicle under shear flow as a function of the shear rate. The intermediate-sized vesicle was filled by 15 small tri-peptide molecules (alanyl-phenylalanyl-alanine) in order to reproduce substrate evacuation under the action of shear flow. The substrate molecules were added inside the vesicle using the Packmol software [161]. The largest vesicle was filled with 5 α-chymotrypsin enzymes, 6 crowding bovine serum albumin (BSA) proteins, and 100 substrate molecules (alanyl-phenylalanyl-alanine) in order to generate a model of a nano-reactor and to investigate protein-substrate encounters under the action of confinement, crowding and HI. The vesicle was filled using Packmol [161]. The structures of the enzyme (α-chymotrypsin, PDB code 1YPH) and of the crowder proteins (BSA, PDB code 4F5S[40]) were obtained from the PDB database[24]. In the intermediate and largest systems, proteins and peptides were modelled using an elastic network (EN) based on C_α and side-chain (SC) beads with a distance cutoff $r_c = 8$ Åand a force constant $k = 5 kcal/(mol$ Å$^2)$. The inter-molecular interactions were based on the OPEP v.4 force field [242]. In the EN the size of C_α beads was increased by 50% compared to the standard OPEP model. This choice was motivated by the fact that in the EN model the backbone of a residue is reduced to the C_α bead only. In the case of small peptides like the substrate because of their small size, the C_α beads can partially penetrate inside the external fold of the crowders altering the diffusivity. Our rescaling prevents this incorrect behaviour. The OPEP SC-SC non-bonded interaction energies were rescaled by $f = 0.545$, a factor which we found to describe more accurately the structure and dynamics of crowded protein solutions. The non-bonded interactions of the lipids with the content of the vesicle were parameterised by utilising existing OPEP potentials rescaled by the same factor f. More specifically, the positively charged choline group of POPC was treated as an OPEP side-chain bead of lysine for the interactions with amino acid beads, the negatively charged phosphate group as that of an OPEP glutamate bead, the two polar glycerol beads as the OPEP asparagine side-chain bead, and finally, the apolar beads of the POPC tails each as the OPEP leucine side-chain bead. In the simulation the friction coefficient was set to $\gamma = 0.0018\, fs^{-1}$ for α-chymotrypsin, $\gamma = 0.001\, fs^{-1}$ for BSA, and $\gamma = 0.05\, fs^{-1}$ for the substrate in order to reproduce the respective HYDROPRO [253] diffusion coefficients in dilute conditions.

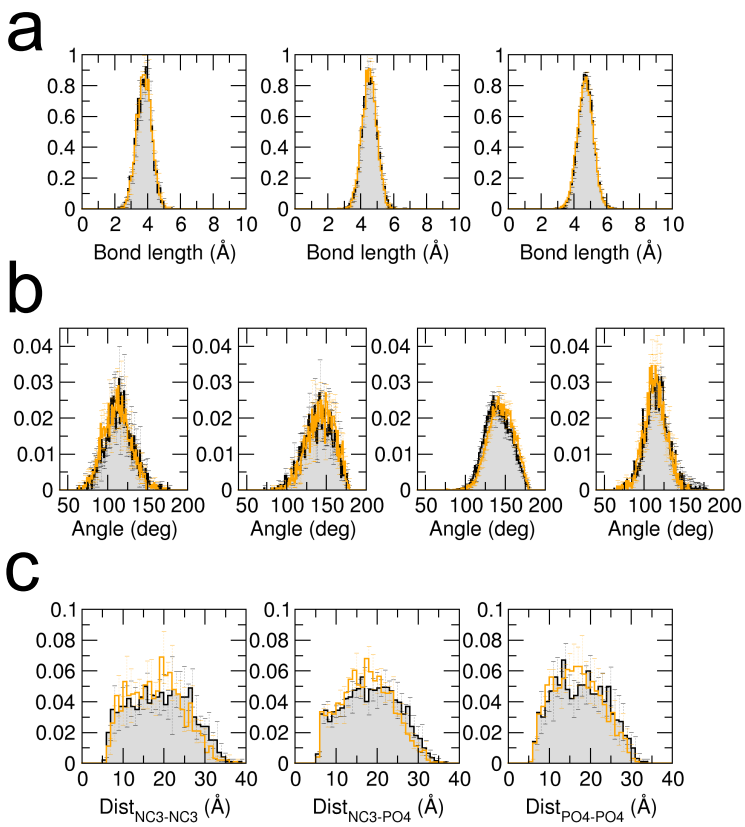

Fig. 4.2 Probability distributions of bonds, angles and charged beads distances. Black lines and gray surfaces represent results obtained from a simulation carried out using Gromacs, and the orange lines represent the results from our implementation in Muphy. Panels A and B present the probability distributions of each of the unique bond or angle types present in Dry Martini respectively. The panel C depicts the probability distributions of the distances between charged beads (NC3: choline bead, PO4: phosphate bead).

4.3 Results

4.3.1 The implementation of Dry Martini

In this section we present a test of the numerical implementation of the force field in the software Muphy [26]. We verify that the structural properties governed by the intra- and intermolecular interactions reproduce those obtained from the original implementation in the Gromacs software (version 4.5.5). A dimer of POPC lipids was simulated by standard MD simulation for a few nanoseconds with temperature controlled by the Langevin thermostat and set to $310\,K$. The probability distributions of the values taken by representative degrees of freedom in the dimer are plotted in Figure 4.2 and compared to those extracted from a reference simulation carried out using Gromacs. The results show excellent agreement.

4.3.2 Coupling with HI

The second type of tests was designed to inquire the effect of the hydrodynamic coupling on the global properties of a simple lipidic system. We have first verified that the coupling with LBMD does not alter the structural properties of a formed bilayer. We performed simulations with and without HI of a membrane bilayer oriented perpendicular to the Z-axis with dimension in the X-Y plane ranging from $10 \times 10\, nm^2$, up to $40 \times 40\, nm^2$. For all systems, the measured average thickness, $\simeq 38$ Å , is unaffected by the coupling with the external fluid. It is also worth noting that the fast fluctuations of the membrane thickness occurring on the picosecond time-scale are not affected by HI, see SI Fig S1. Finally, a small system composed by 128 POPC lipids randomly placed in a cubic box of size $L = 75$ Å was simulated with and without HI to explore the role of HI on the self-assembling of the lipids. In both cases a bilayer was formed, see Fig. 4.3 for a pictorial representation.

The timescale of the assembly process clearly depends on the choice of the friction parameter γ. For $\gamma = 2.5 \times 10^{-4}\, fs^{-1}$ a complete bilayer structure was obtained in about $10\, ns$, while for higher values of γ the kinetics was slower. When formed, the final thickness of the bilayer was between $38 - 39$ Å which is in agreement with experimental data [137]. Interestingly, for this highly concentrated small system HI have negligible impact on the aggregation process, but is rather controlled by the mechanical forces. To be noted, however, that in the investigation of other molecular systems (amyloid peptides, or simple lipids) at lower concentration it was demonstrated that HI accelerate the aggregation process [7, 54].

4.3.3 Lipid lateral diffusion

Here we focus on lipid dynamics inside a membrane. Recently, the effect of periodic boundary conditions (PBC) and of simulation box size on the calculation of the lipid lateral diffusion in simulations has been thoroughly investigated, and rationalised [45, 261, 262]. Because of reentrant hydrodynamic interactions under PBC (stemming from both the solvent and lipids), it was shown that the calculated diffusion constant for lipids (or an embedded rigid body like a protein) in a planar membrane simulated with periodic boundary conditions (D_{PBC}) diverges logarithmically as a function of the lateral box size $L_x = L_y = L$[45, 261, 262] for fixed transversal separation (L_z). It was however shown that the value of the diffusion coefficient for an infinite system (D_0) to be compared with experiments can in principle be obtained using adequate numerical procedures [45, 177, 262]. In previous studies the size dependence on the calculated lipid diffusion was estimated by considering either explicit solvent atomistic models or the explicit solvent CG model Martini [45, 261, 262]. Recently, a work by Lyman *et al.* [274] reported the coupling of the implicit CG Dry Martini model

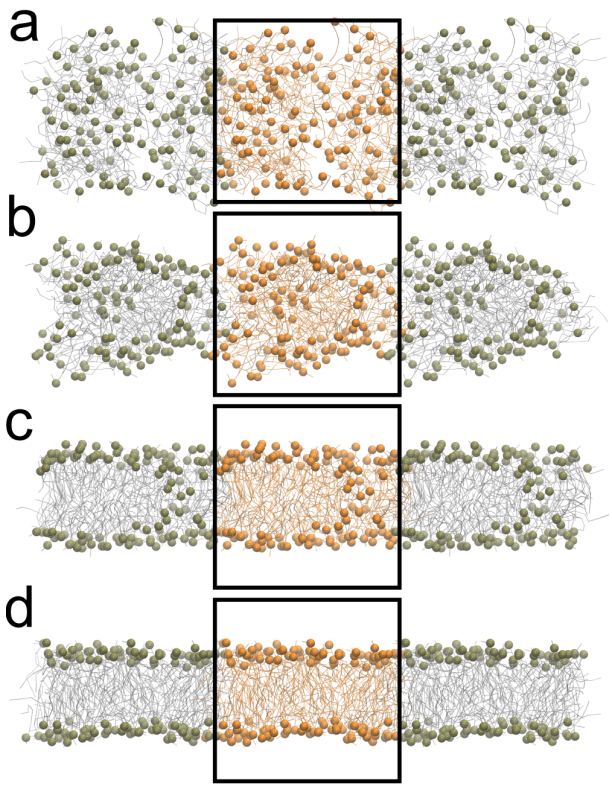

Fig. 4.3 Spontaneous bilayer assembly of 128 POPC lipids. The panels A-D show the time evolution of the assembly process. For the sake of clarity, the phosphate beads are represented as van der Waals spheres, whereas the rest of the lipid is shown as lines. The system is represented with periodic images, and the orange-coloured particles belong to the simulation box.

with the multiple-collision algorithm, also referred to as Stochastic Thermostatted Rotation Dynamics (STRD)[157], in order to integrate solvent mediated hydrodynamic interactions in implicit solvent particle simulations. With this coupling in hands the authors studied the size dependent lipid lateral diffusion in POPC membranes. The lateral dimensions of the membrane system were increased from $L_x = L_y = L = 100$ Å up to 400 Å while the separation between two periodic membranes in the transversal direction was 200 Å. In the same spirit we performed LBMD simulations of POPC membranes of different sizes (100 Å$< L <$ 400 Å) as in Ref. [274]. We set $L_z = 240$ Å to ensure a separation of ~ 200 Å between two periodic membranes along the transversal direction as in Ref. [274]. For each system the simulation was $1.2\gamma mus$ long, the last 800 ns of the trajectories were used for calculating the 2D mean square displacement (MSD). The diffusion coefficient was derived via linear fit of the MSD in the time window 40-70 ns (see SI Fig. S2). In Figure 4.4 we report the calculated average diffusion coefficients from the LBMD simulations and compare to the results from STRD

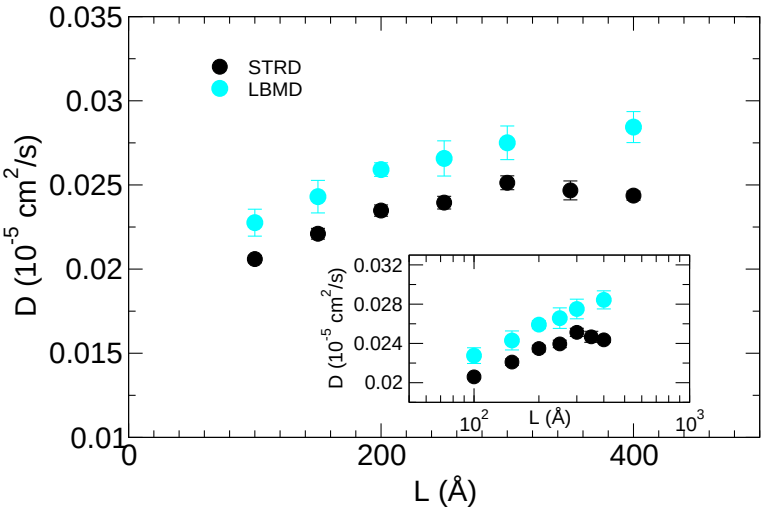

Fig. 4.4 **Box size effect on lipid diffusion.** Diffusion coefficient for POPC lipid extracted from 800 ns simulation of a planar membrane in different boxes of size $L_x = L_y = L$ and $L_z = 240$ Å(cyan circles). The diffusion coefficient D was obtained from a linear fit of the MSD computed during the last 800 ns of the simulation and using a time window of 40-70 ns for the fit. In the figure (black circles) we also report the data from Ref. [274] based on STRD simulations of the POPC lipid planar membrane with inter-membrane separation of 200 Å. Inset: Log-linear plot of the diffusion coefficients.

[274]. Our data reproduce the size dependency already observed for Dry Martini and using STRD [274], with the diffusion constant growing unbounded with the lateral box size, and start showing the logarithmic divergent behaviour (see inset graph of the Figure 4.4 where we show the log-linear plot of the diffusion coefficient D). Our finding shows that LBMD is a computationally appealing alternative strategy for using implicit solvent CG lipid models coupled to the hydrodynamic interactions from an embedding fluid. The approach allows to simulate large sized membrane systems at a cheaper computational cost with respect to explicit solvent CG models. A complete investigation of the effect of sizes (membrane dimension as well as the transversal separation) and fluid viscosity, as for instance presented in Refs. [45, 261, 262], is beyond the scope of the present modelling and it is reserved for a future study.

Nanotube adsorption

The contribution of HI stemming from the surrounding fluid can be appreciated by considering the adsorption dynamics of a biomolecule on the membrane surface. Two aspects should be considered, the first concerns the general effect of HI on biomolecule diffusivity [7, 54, 85, 239]. For a single protein, and even for aggregated objects, it was previously shown that the inclusion of HI speeds-up molecular transport. This effect can be rationalised [7] for instance by referring to the Zimm description of polymer dynamics [277] where with the inclusion of HI, the diffusivity scales more favourably with the polymer size than the correspondent case without HI (Rouse description [210]). The second aspect, which we do not explicitly explore here, relates to the bending oscillations of the lipid membrane that couple with the solvent dynamics and cause a repulsive effective force on particles moving along the membrane surface [202].

Here, as a simple test, we considered the adsorption process of a nanotube of length $L = 43.1$ Å and radius $r = 7.5$ Å on the surface of a flat membrane in a simulation box of lateral sizes $L_x = L_y = 150$ Å and transversal dimension $L_z = 300$ Å. Ten independent orientations of the nanotube were generated by rotating its axis and maintaining the centre of mass at a distance of 113 Å from the membrane surface. In our simulations the dynamics of adsorption is generally accelerated by the presence of HI interactions, with the average adsorption time being 10% shorter than without hydrodynamics, see SI Table S1. Since at the considered length scales the membrane does not exhibit visible transversal oscillations, we do not observe hydrodynamic repulsive effects stemming from the membrane motion [202]. An example of the adsorption process is represented in Figure 4.5 where the time series of the minimal distance between the centre of mass of the nanotube and the membrane surface is plotted for a simulation including or neglecting hydrodynamics. A pictorial representation of the nanotube dynamics including the surrounding fluid represented by its isokinetic streamlines is presented in the top part of the figure.

4.3.4 Vesicles in fluid flows

The capability to represent effectively a lipid membrane together with the surrounding solvent fluid opens up the possibility to investigate systems and processes of biological and technological relevance. One example is the behaviour of vesicles and liposomes under shear flow [25, 56, 178, 225, 264, 265]. Vesicles are simple models of cellular compartments, and their response to a shear flow has been studied to understand both the response of large biological structures, such as red blood cells, [264] or the design of drug delivery vectors [25, 56, 225, 265]

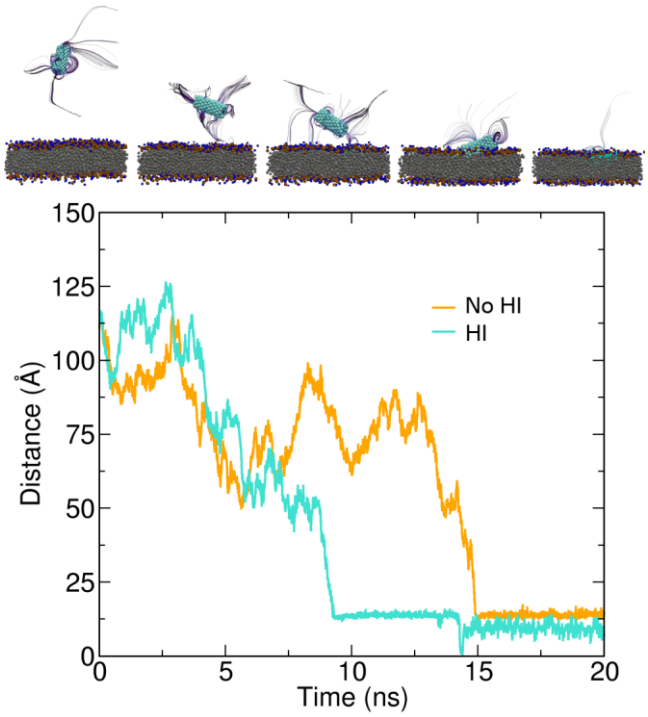

Fig. 4.5 Nanotube adsorption. Time evolution of the minimal distance of the nanotube centre of mass with respect to the external surface of the lipid bilayer for one of the simulated systems. Starting from the same initial condition two simulations were performed with (LBMD, cyan) and without (Langevin, orange) HI. On the top part of the figure we represent the streamline of the fluid surrounding the nanotube during its adsorption process.

Recently, it was shown experimentally that a vesicle, depending on external fluid viscosity and shear rate, can experience three different dynamical regimes, from tank treading to tumbling and trembling [71, 273]. The first two motions imply a preserved shape while the third one is associated to shape deformations. To show that the LBMD method can reproduce the behaviour of vesicles under a shear flow we have assembled a small POPC spherical vesicle formed by 394 lipids ($R \sim 34$ Å) immersed in a fluid in a cubic simulation box of size $L = 240$ Å. We observed how an external shear flow affects its dynamics and shape fluctuations. We scanned four different values of the shear rates, $\dot{\gamma} = 3, 7, 20, 30 \times 10^9 s^{-1}$ which allowed us to monitor the response of the systems on the nanosecond time scale. In Refs. [240, 241] the interested reader can find the technical details about how a shear flow is generated. To be noted that shear rates used here compare to the values applied in recent DPD simulations to inquire the effect of shear on protein migration in the membrane of vesicles [178].

In the simulations, the increase of shear rate causes a progressive deformation of the spherical vesicle towards an prolate shape. This transformation is quantified by the eigen-

values λ_i of the vesicle gyration tensor ($\lambda_1 \geq \lambda_2 \geq \lambda_3$), see Figure 4.6, and their relative ratio (see SI Figure S3). The shear rate not only causes a deformation but also controls the oscillation of the gyration eigenvalues λ_i. This frequency of the oscillation of λ_i increases due to the rotational component of the shear field, with characteristic time inversely proportional to the shear rate, $\tau \propto 1/\dot{\gamma}$. When the shear field becomes too high the vesicle breaks apart, and the lipids start forming dynamical entities which encounter, partially fuse and dissociate as they follow the fluid flow. A pictorial representation is reported in SI Figure S5.

In the Figure 4.7.A we have represented the typical evolution of the vesicle, and its membrane, during a period of the sliding dynamics ($\sim 1\,ns$). In order to characterise the dynamics of the lipids and the vesicle, we have projected the eigenvector $\Lambda^{(3)}$ of the gyration tensor along the Z axis along which the velocity gradient is aligned in the Couette flow. The eigenvector $\Lambda^{(3)}$ is associated to the main axis of the prolate. For this projection we considered the absolute value. Similarly we have considered the vector $\mathbf{v}$ connecting one polar bead of a lipid in the membrane to the centre of mass of the vesicle. For this vector too, we have considered the projection along the Z axis. The normalised projections are reported in Figure 4.7.B and 4.7.C and show two characteristic motions. The first motion concerns the oscillation of the main axis of the vesicle. After a transient initial phase, when the vesicle is deformed by the flow, the main eigenvector $\Lambda^{(3)}$ of the gyration tensor never aligns with the Z axis, in fact the value of the projection stays always < 0.5. The second motion relates to the sliding of the membrane which has a slower frequency with respect to the oscillation dynamics of the vesicle principal axis.

In order to model a more a realistic system and match the characteristic size of POPC small unilamellar vesicles [39] we have performed extra simulations of a larger vesicle of radius $\simeq 100$ Å, and formed by 3451 lipids. The vesicle was placed in a cubic simulation box of size $L = 450$ Å under the action of a shear flow of rate $\dot{\gamma} = 1.9$ and $3.7 \times 10^9 s^{-1}$ and $\dot{\gamma} = 3.7 \times 10^7 s^{-1}$. Again, as seen in the previous example, in Figure 4.8 we observe the shape deformation toward a prolate structure for high shear values (Panel B and C), while for this vesicle a shear rate $\dot{\gamma} = 3.7 \times 10^7 s^{-1}$ does not affect the spherical shape. The dynamics of the vesicle's characteristic vectors is reported in SI, Figure S6.

The response of a vesicle to the external shear flow [25, 265] can be exploited to design drug delivery systems or biological constructs that control chemical reactivity. We therefore considered an extra system -a small toy model- to test how the release of substrates can be simulated using our mixed particle/fluid approach. In this model a spherical POPC vesicle formed by $n = 923$ lipids with an initial radius $R = 60$ Å was filled with a small number of substrates, 15 small alanyl-phenylalanyl-alanine tripeptides. The substrate was modelled as an elastic network, see the Methods section, and the interaction parameters with the

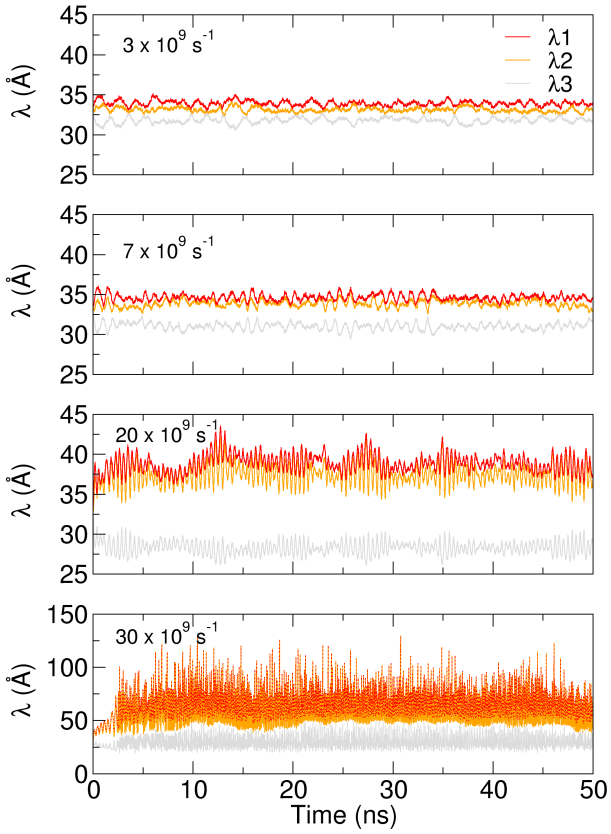

Fig. 4.6 Vesicle in shear flow. Time evolution of the three eigenvalues of the vesicle gyration tensor for shear rates $\dot{\gamma} = 3, 7, 20, 30 \times 10^9 s^{-1}$.

membrane were inspired by characteristic energy terms of the OPEP force field. The vesicle was placed in a cubic box of size $L = 350$ Å under a shear flow with $\dot{\gamma} = 6 \times 10^9\, s^{-1}$.

As discussed above, a spherical vesicle undergoes deformation by the shearing of the surrounding fluid, and acquires prolate conformations. As a consequence, the internal pool is deformed as well and the substrates under the action of the external perturbation accommodate in the flattened space, see Figure 4.9. After 40 ns, some of the substrates penetrate the membrane and are released in the solution. At the very beginning the evacuation involves just a few molecules but under the constant stress of the external fluid, a collective release eventually occurs. It is interesting to note that the substrate release does not require a disruption of the vesicle but it is the consequence of a forced permeability of the membrane. It seems that the drag caused by the external shear flow results in an effective force that induces the substrate to cross the membrane. In a separate work, we will inquire the molecular basis of the evacuation process by considering larger and realistic systems, and by investigating the different aspects controlling the process, e.g. how the vesicle shape deformation alters the

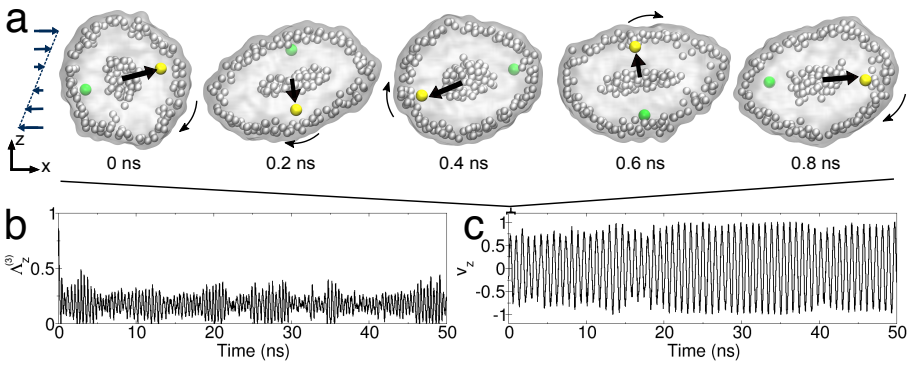

Fig. 4.7 Vesicle in shear flow. In the top layer of the figure (Panel A) we represent the short time rotational motion of the vesicle under shear flow. Two lipid beads corresponding to the outer layer are highlighted (yellow and green spheres) to help visualising the membrane rotation. The yellow and green beads remain aligned at a short time, but at a longer time, as we show in SI Figure S4, the different lipid mobility in the bilayer causes their mis-alignment. In panel B and C we represent the time evolution of the normalised projection on the Z axis of the system of two characteristic vectors of the vesicle: the eigenvector $\Lambda^{(3)}$ of the gyration tensor and associated to the longest axis of the prolate (Panel B), and the vector $\mathbf{v}$ connecting one polar bead of the lipids in the membrane (yellow sphere) to the centre of mass of the vesicle (Panel C). For the projection of the eigenvector we considered the absolute value.

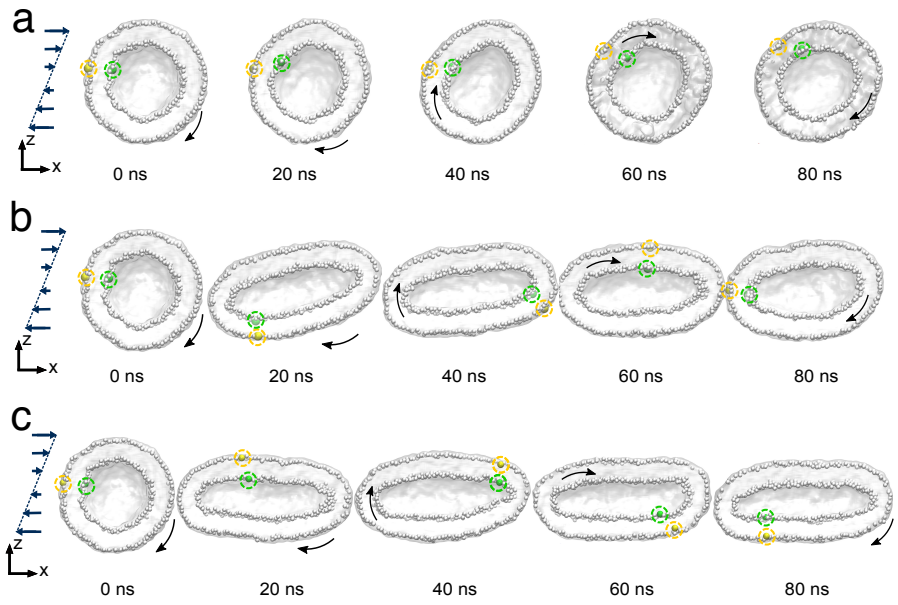

Fig. 4.8 Vesicle in shear flow. Time evolution of the shape of a larger vesicle ($R \sim 100$ Å) under the action of the shear flow. Panel A, B and C refer to simulations at different shear rate, $\dot{\gamma} = 3.7 \times 10^7 s^{-1}$, 1.9 and $3.7 \times 10^9 s^{-1}$, respectively. Two lipid beads corresponding to the outer and inner layer are highlighted (yellow and green spheres, respectively) to help visualising the membrane rotation and the lipid motion inside the membrane.

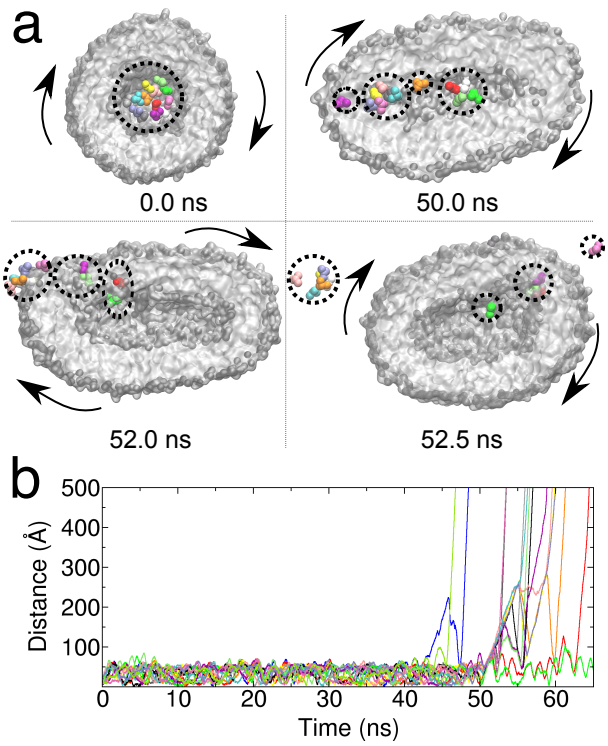

Fig. 4.9 Substrate release in shear flow. Time evolution of the single substrate distance with respect to the vesicle centre of mass. In the top part of the figure we represent the molecular detail of the substrate release. Clusters of substrate molecules are highlighted by dashed lines. Arrows indicate the tumbling motion of the vesicle.

local packing of the lipids in the membrane and changes the associated free-energy barrier for membrane penetration.

4.3.5 Nanoreactor

In order to demonstrate that the LBMD technique and the presented coupling with the Dry Martini lipid force field can be used to study very large systems, we considered a more complex construct aimed at modelling a vesicular nano-reactor. In fact, vesicles filled with enzymes and substrates are widely used to investigate biochemical reactions under confinement [138]. Here we employ the LBMD approach to simulate the first detailed model of a vesicular nano-reactor, inspired by constructs investigated in previous experimental studies [17, 53]. We selected α-chymotrypsin, one of the enzymes reported to exhibit increased activity under confinement, as the enzyme to be placed inside the nano-reactor. The nano-reactor was formed by a POPC vesicle with an inner radius of 120 Å, placed in a cubic simulation box of size $L = 410$ Å. The vesicle is filled with 5 α-chymotrypsin molecules

(238 residues) in the presence of 6 protein crowders (BSA, 583 residues), corresponding to a crowder concentration of 92 g/L, accompanied by 100 oligopeptide molecules (alanyl-phenylalanyl-alanine), forming the substrate for the enzymatic reaction.

Preliminary simulation (300 ns) allowed us to characterise the spatial distributions of the various molecules inside the nanoreactor as well as their diffusivities (see Figure 4.10). These factors affect the rate of the chemical reaction by governing the probability that an enzyme will meet a substrate. We observed that both the protein crowders as well as the substrate molecules were concentrated near the lipid membrane (Figure 4.10.B), with the large protein crowders diffusing tangentially to the membrane surface. The enzymes also interact with the lipid bilayer. However, rather than being in direct contact with the membrane surface, some of the enzymes preferred to bind to the crowder, which brought them closer to the centre of the vesicle. For all the three molecular species, we observed a slowdown in diffusion compared to dilute conditions in the absence of confinement. On the short timescale ($1-5\,ns$), the slowdown is $35\%-50\%$ depending on the molecule type. On the longer timescale ($10-50\,ns$), the slowdown of the diffusion became even more pronounced ($\sim 60\%$), reflecting the finite size of the vesicle interior and the interactions with the lipid membrane. For the BSA proteins, the slowdown inside the vesicle was larger by 30% than that measured via neutron scattering for a BSA solution at comparable crowding condition [208]. Interestingly, at the long timescale, the diffusion coefficient of the enzymes practically matched that of the significantly larger crowders (see Figure 4.10.C), as a consequence of their mutual interactions in the vicinity of the membrane. Finally, we did not observe any major changes in the shape of the vesicle in the course of the simulation, with the maximum deviations not exceeding $4-6\%$ of the vesicle size in each dimension. The results of our simulation confirm the capacity of the LBMD technique to provide a detailed molecular view of protein motion in large and realistic systems consisting of many biomolecules surrounded by a lipid membrane. Thus, the results pave the way for simulations of entire organelles or therapeutic constructs to capture the complex biological processes occurring in their interior.

4.4 Conclusions

In this work we coupled the coarse-grained Dry Martini model for lipids [10] to the Lattice Boltzmann Molecular Dynamics technique [2, 3, 239]. This is a precious combination since it naturally allows including hydrodynamic interactions while still preserving the computational gain related to the implicit solvent nature of the Dry Martini model. It was already demonstrated that the presence of solvent mediated correlations is a necessary ingredient to study biological systems at the mesoscale[6, 7, 54, 167], including important

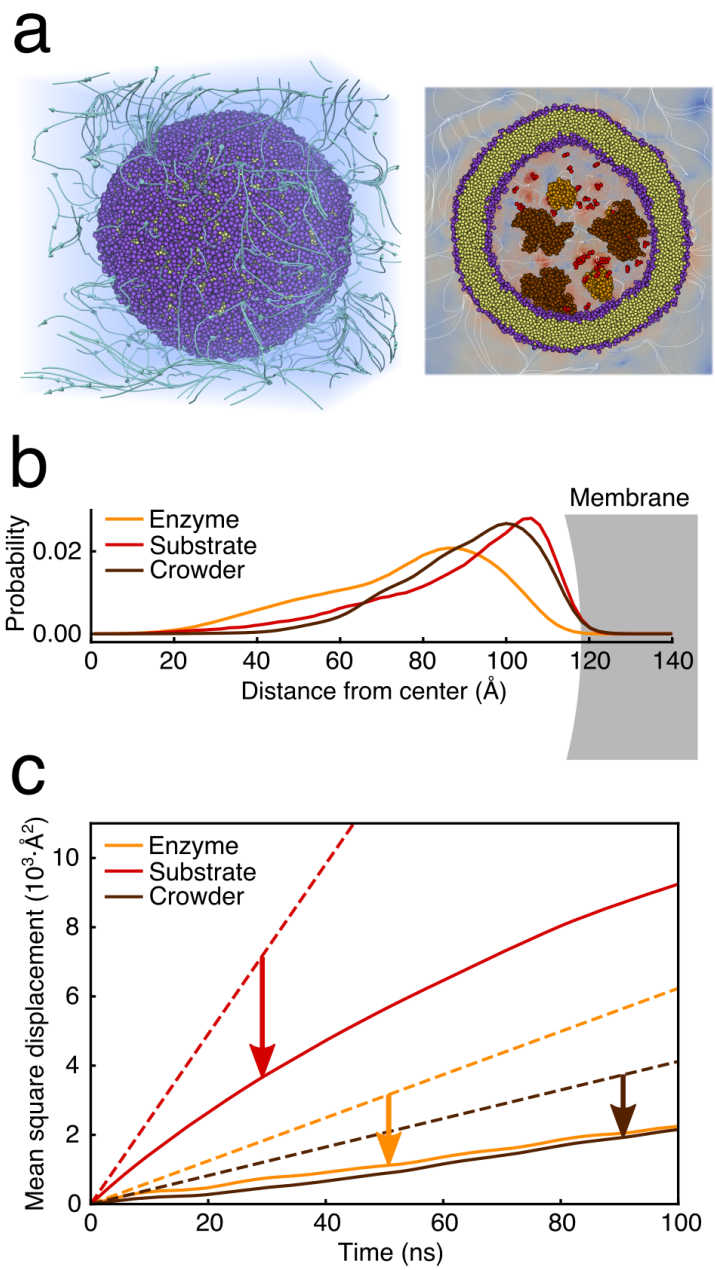

Fig. 4.10 Nanoreactor. Panel A. Pictorial representation of the nanoreactor and the internal and external fluid. Panel B. Probability distribution of proteins and substrate localisation inside the vesicle. Panel C. Proteins and substrate mean square displacement in dilute condition (dashed lines) and inside the vesicle (solid lines).

processes involving membranes, such as lipid aggregation [7], protein migration in proximity of membranes [202], vesicle encounter and fusion [209].

We first validated the implementation of the force field in the multi-scale simulation software Muphy [26]. Then, we explored different situations where the coupling with the

surrounding solvent is key for the behaviour of the lipid system. For instance, we examined the size effect in the estimate of lipid diffusion. In these regards, the LBMD technique based on Dry Martini could be computationally appealing with respect to explicit solvent models in order to approach the infinite-size limit necessary to calculate the transport properties of lipids or proteins in two-dimensional membrane systems [45, 261, 262]. Since the technique allows simulating a great variety of fluid flow conditions, we used the methodology to explore the behaviour of a vesicle under the external perturbation of shear flow mimicking experimental assays[25, 71, 265]. Finally, inspired by recent experimental work [138] we explored the possibility to simulate a nano-reactor so as to investigate the combined effect of confinement, crowding and hydrodynamics on substrate/enzyme encounters.

The presented coupling could be further extended. All our simulations were performed in the canonical ensemble at fixed volume and temeprature. However, for a membrane system it is well known that it is important to control the lateral pressure to ensure the accurate structural and dynamical behaviour specifically for laterally periodic systems [192, 250]. In our scheme it is possible to implement an efficient algorithm for the particle dynamics to sample the isobaric ensemble. The fluctuating simulation cell would be mapped on a fixed number of grid points used for the resolution of the fluid dynamics. According to our tests, at equilibrium even for a small membrane system ($L_x = L_y = 50$ Å), where the fluctuations are expected to be most pronounced, they remain below 2% of the mean lateral size. This implies that cell fluctuations modify the grid spacing by about 2% as well. When preparing a system in isobaric condition the transient relaxation leads to a shrinking of the cell that is on the order of 4% of the initial lateral size. Therefore even in this case one can utilise a constant resolution for the LB component.

A second aspect to mention concerns the application of multi-resolution meshes to increase efficiency. Our scheme is suitable for the study of transport properties in two-dimensional membrane systems in periodic boundary conditions where in order to approach the infinite-size limit the transversal dimension should be increased enormously. The computational cost increases linearly with the size of the transversal dimension with a single grid spacing, but a tremendous gain can be achieved by using a multi-resolution approach, where only a portion of the space surrounding the membrane is resolved at the fine scale while the remainder of the volume is mapped on a coarser grid. Also, the computation of the particle dynamics can be boosted by implementing new algorithms to exploit the vectorial capability of modern CPUs [1] and approaches for parallelism to handle more efficiently the sparse nature of some lipidic systems like large vesicles.

In conclusion, in this work we extended the operational capability of our multi-scale Lattice Boltzmann Molecular Dynamics technique [54, 55, 143, 239, 240], an alternative

to the scheme proposed by Lyman where Dry Martini was coupled to the STRD method [274]. With our approach we are now able to explore a great variety of biophysical processes involving lipid systems at the mesoscale where solvent mediated interactions are important. Namely, our approach can support the bio-engineering and biotechnology experimental investigation of lipidic systems like vesicles in fluid flow conditions [11, 168]. In a forthcoming work we will concentrate on the effective coupling between the Dry Martini force field and the coarse-grained model for protein OPEP in order to model complex protein/lipid systems with a flexible model for proteins.

4.5 Supplementary Material associated to this publication

The following figures and tables are the supplementary data related to this article.

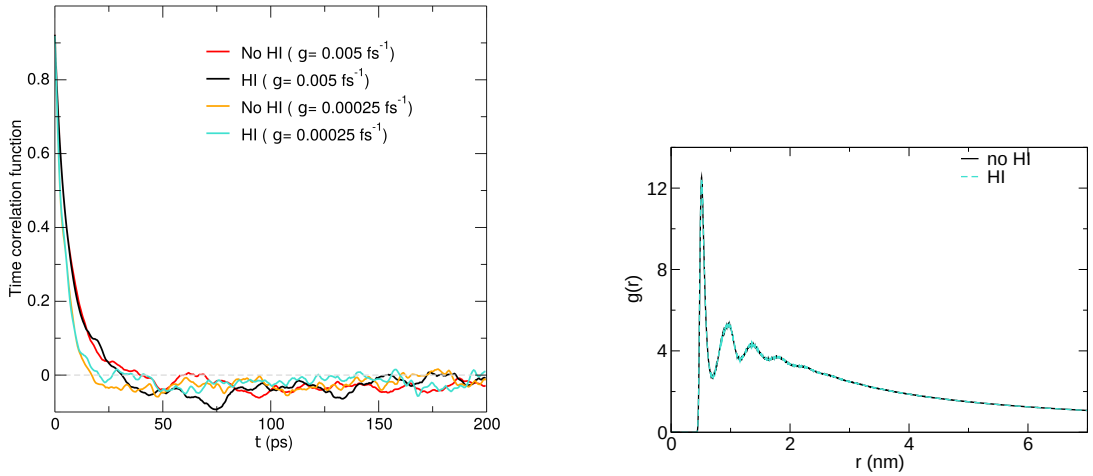

Fig. 4.11 Left. Time correlation function of the membrane thickness fluctuation for a membrane system simulated with and without hydrodynamics, and using two different value of the friction coefficient γ. Right. Radial distribution function of hydrophobic beads for a simulation of a POPC bilayer with (HI, dotted cyan line) and without (no HI, black line) hydrodynamic interactions.

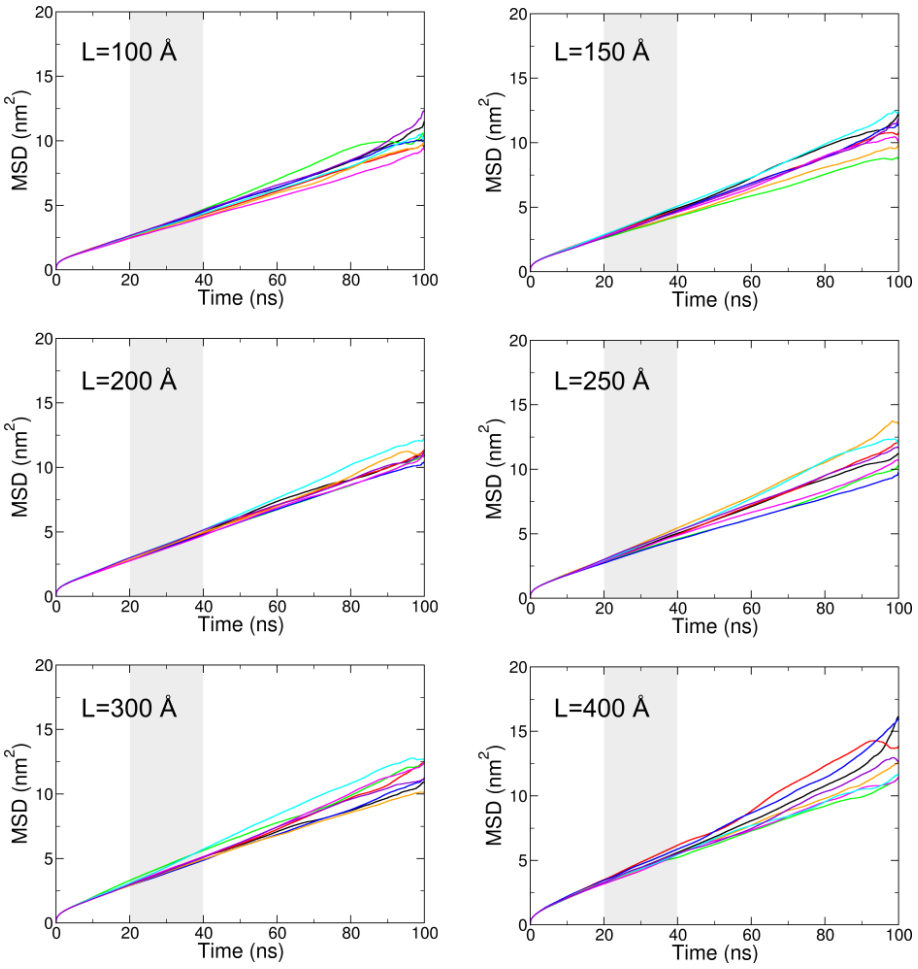

Fig. 4.12 Lateral mean square displacement calculated (MSD) for the lipids in a planar membrane of size $L_x = L_y = L$ and fixed $L_z = 243$ Å. For each system the MSD is calculated using independent blocks of the trajectory, each of 100 ns. The diffusion coefficient was derived by a linear fit in the time window 20-40 ns (gray zone).

System	τ [ns] (no HI)	τ [ns] (HI)
S1	8.20	6.37
S2	14.89	10.91
S3	7.09	6.41
S4	6.17	5.02
S5	12.34	12.36
S6	9.90	6.71
S7	10.62	10.95
S8	15.09	11.92
S9	7.14	9.87
S10	9.15	14.21
S11	37.81	37.97
S12	30.23	29.5
S13	21.33	30.77
S14	37.36	29.22
S15	34.95	20.9
S16	22.18	31.23
S17	28.79	10.52

Table 4.1 Nanotube adsorption times obtained in 17 different simulations with and without hydrodynamics.

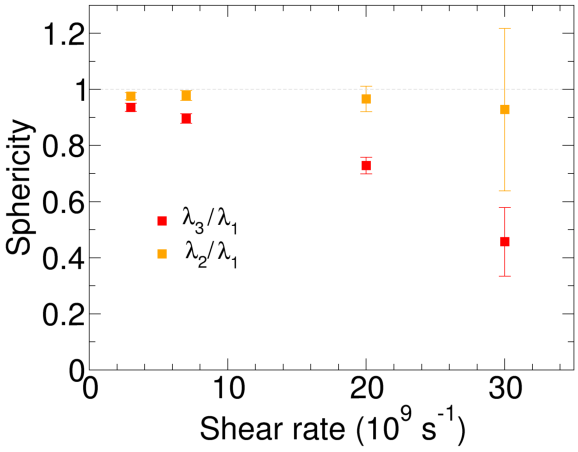

Fig. 4.13 Vesicle sphericity computed as the ratio of the eigenvalue of the vesicle gyration tensor as a function of the shear rate.

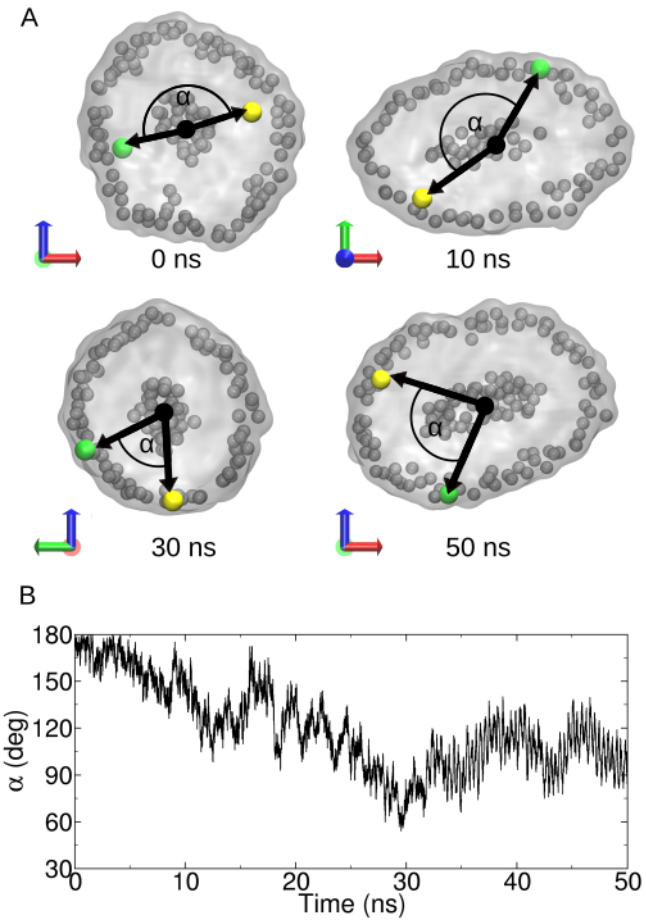

Fig. 4.14 Time evolution of the angle formed between the polar heads of two lipids in the vesicle under the action of shear flow with $\dot{\gamma} = 20 \times 10^9 s^{-1}$.

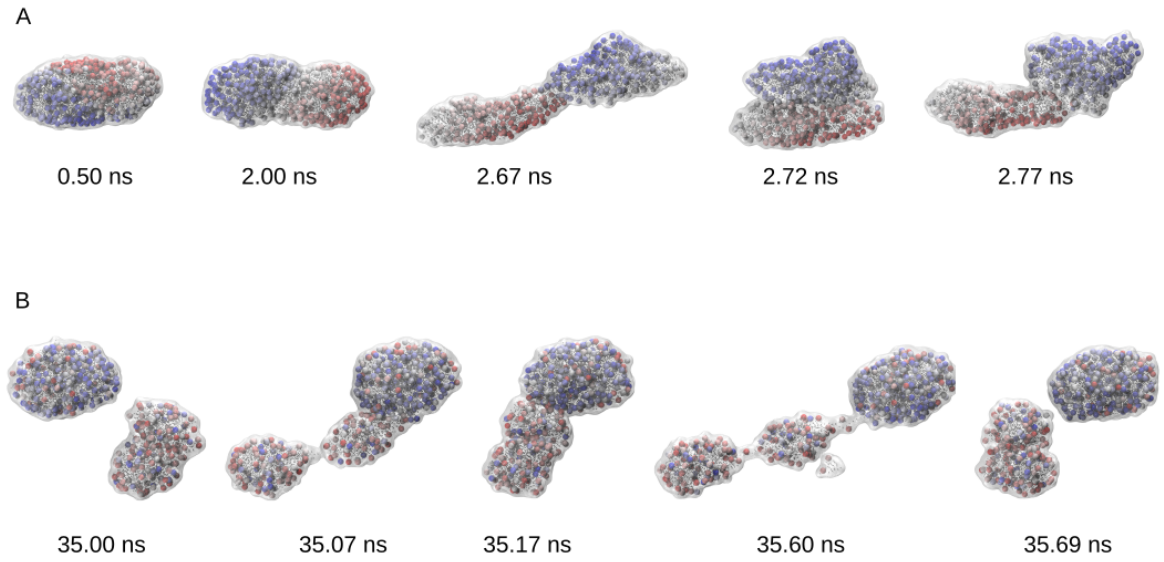

Fig. 4.15 Vesicle under shear flow. At very high shear rate, $\dot{\gamma} = 30 \times 10^9 s^{-1}$, the vesicle breaks in smaller entities that encounter, fuse and disconnect as effect of the solvent shearing. The polar heads of two set of lipids, initially separated in the vesicle, were coloured in red and blue. The time evolution of the system shows their mixing.

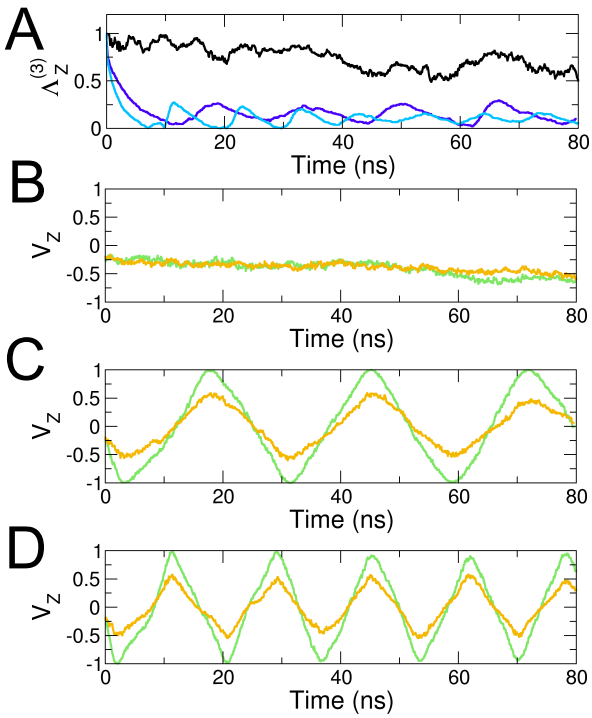

Fig. 4.16 Vesicle under shear flow. In panel A we represent the time evolution of the normalised projection on the Z axis of the system of the eigenvector $\Lambda^{(3)}$ of the gyration tensor associated to the longest axis of the vesicle. The absolute value of the projection is reported in the figure for the three shear rates $\dot{\gamma} = 1.9$ (violet) and 3.7 (cyan) $\times 10^9 s^{-1}$ and $\dot{\gamma} = 3.7 \times 10^7 s^{-1}$ (black). Panel B-D: projection of the the vector $\mathbf{v}$ connecting one polar bead of the lipids in the membrane to the centre of mass of the vesicle. In each panel the green and yellow curves refer to two different beads in the membrane. Panel B refers to the lowest shear rate while Panel D to the highest.

Chapter 5

Expanding OPEP force field to study protein-lipid systems

5.1 Generalities and Implementation

In this chapter we discuss a first systematic attempt to couple a coarse-grained force field for lipids (Dry Martini) with the OPEP force field for proteins. The OPEP coarse-grained force field for proteins was designed to study protein folding and has a very detailed backbone description, while keeping only one bead for the side-chain of every amino acid (except glycine). Currently it is limited to the study of proteins in solution, so the main goal of this chapter is to expand the universe of systems that can be studied with OPEP to include membrane proteins and peptides. For instance, it would be of interest to study conformational changes and folding of peptides or proteins in a membrane environment, since the OPEP force field is able to account for such changes in contrast to the common Martini force field where the secondary structure has to be fixed. From a more technical point of view, the idea of coupling the Dry Martini force field for lipids with the OPEP force field for protein was based on two main characteristics: (1) both are implicit solvent models and (2) the Dry Martini lipid force field has a mapping resolution that is comparable to the one used in OPEP to model amino acids side-chains. At the moment of the writing of this thesis, the Dry Martini force field has a set of parameters to model proteins in interactions with the lipids that have been published but are still under development. This protein model is based on the topology of the explicit solvent Martini coarse-grained model for protein but has a slightly different energy to account for the lack of explicit solvation.

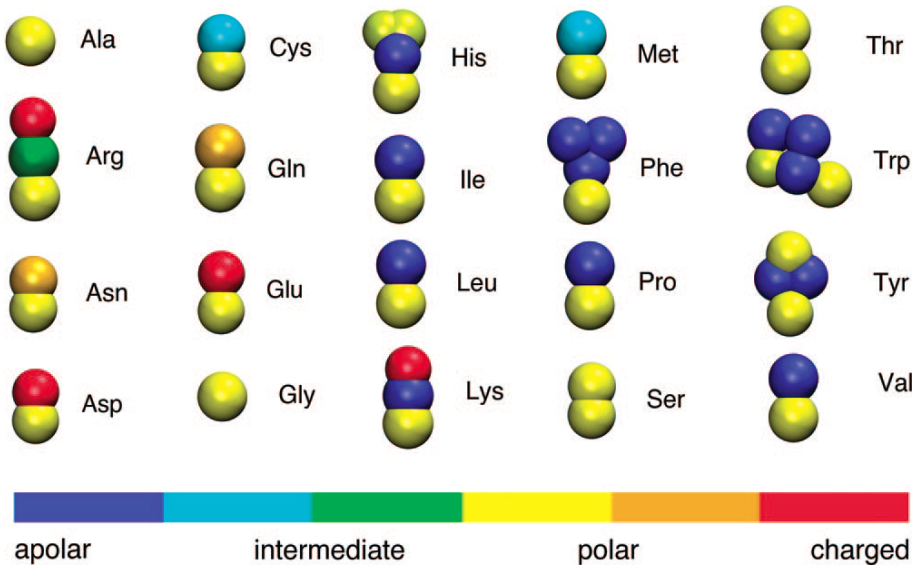

Fig. 5.1 Martini topology for amino acids colored by polarity (including backbone bead). *Taken from [171]*

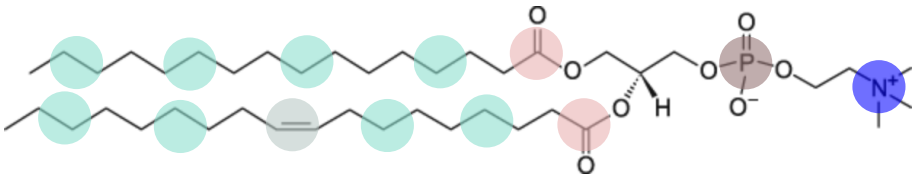

Fig. 5.2 Dry Martini v2.1 POPC mapping highlighting the different beadtypes by colours: C1 (cyan, representing the hydrophobic beads from the acyl chains), C3 (grey, representing the unsaturation in the acyl chain), Na (pink, representing the glycerol moiety), Qa (brown, representing the phosphate) and Q0 (blue, representing the choline moiety)

5.1.1 Initial energy comparisons between the different force fields

Because Dry Martini and OPEP force fields were parametrised with different strategies, we firstly decided to compare the energy scale between these models. To have an idea of the characteristic energies involved, we have plotted the Lennard-Jones interactions present in a typical lipid (POPC) and compared them to their analogous beads in the OPEP force field. A plot of this comparison is shown in figure 5.3.

The analogous protein side-chain beads were chosen based on the chemical moiety they represent and are named by the residue's side-chain they belong to. In particular, the non-bonded energy potential describing the interaction between two lipid hydrophobic beads (C1-C1) has a much higher energy minimum than its protein analogue (LEU-LEU), see the panel with the asterisk in figure 5.3. For this reason, and to potentially avoid the disruption of the membrane by the presence of the protein, we decided to use the energy scale of Dry

Martini as a reference for the cross interaction between OPEP protein and Dry Martini lipids [10]. It is important to note that, as shown in figure 5.1, some side-chains of amino acids in Dry Martini are represented by more than one bead. For these amino acids we have derived a new single bead effective potential describing the interactions between the amino acid side-chain and the lipid beads. The amino acids described in Dry Martini by multiple beads

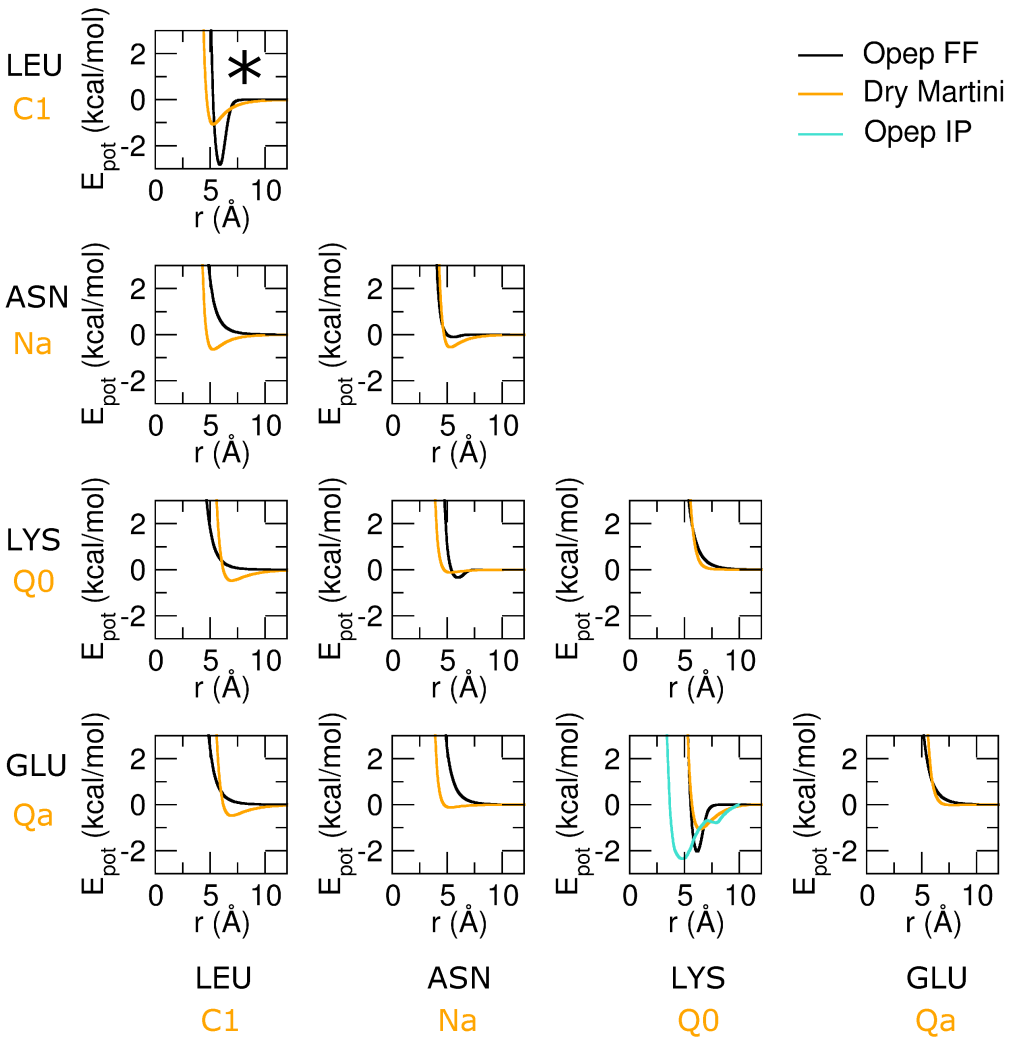

Fig. 5.3 Non-bonded interaction potentials between the different POPC lipid beads in Dry Martini (v2.1) (orange).The non-bonded interaction potential fo their analogous beads in OPEP protein force field (v4) are also plotted to be compared (black). The cyan line shows the attractive ion-pair potential for charged-charged interactions as implemented in the OPEP force field (v5, OPEP IP). The asterisk highlights the case with a marked energetical difference: the interaction between the hydrophobic bead for the lipid (C1) and its counterpart in the protein model (LEU). *For an extended comparison of small polar and apolar interactions see figure 5.15 in the Appendix*

are: the tryptophan (TRP), the phenylalanine (PHE), the tyrosine (TYR), the histidine (HIS), the arginine (ARG) and the lysine (LYS). For the amino acids that in Dry Martini protein force field are described by a single bead, we directly used the already existing parameters tuned in the Dry Martini context. On the other hand, to optimise the Dry Martini force field for the case of amino acids with more than one side-chain bead we decided to have a consistent thermodynamic approach as used originally in the development of the Martini force field for protein. First, for each of the amino acid side-chains with multiple beads, we computed the potential of mean force to pull out an amino-acid side-chain from the center of a POPC membrane along the direction orthogonal to the bilayer surface and using the Dry Martini force field for both lipid and protein. Umbrella sampling simulations were run in Gromacs to obtain the potential of mean force for each of these six side-chains as a function of the distance from center of mass of the POPC bilayer. The POPC bilayer was formed by 64 lipids per leaflet and the amino acid was placed initially at the center of the membrane. A harmonic potential was used to restraint the coordinate orthogonal to the membrane surface (z) of the distance between the centre of mass of the amino acid side-chain and the one of the bilayer. The spacing between the centres of the biasing potential was 0.05 nm and the number of windows was large enough to account for the amino acid being fully outside the bilayer. The simulations were unbiased using the g_wham tool in Gromacs and the error bars assigned via the bootstrap methodology [119] using a number of 200 bootstraps and considering each complete histogram from the biased simulation as independent data points. The obtained potential of mean force, that we indicate as PMF^{ref}, was used as a target reference to optimise the cross-interaction terms designed to couple the side-chain of amino acids based on the OPEP resolution (1 bead) and the lipid modelled by Dry Martini. In all cases treated in this section the parameters correspond to a protonation at pH 7.

In order to generate the potential of mean force for the OPEP single bead side-chains, a first guess of the cross interactions with the lipids was necessary. For this purpose we designed two strategies. In the first strategy (which we will refer to as *strategy 1*) we individuated the leading interaction that controls the shape of the PMF. In fact, when considering the multi-beaded side-chained amino acids the energetics of the side-chain interacting with the lipid depends on the sum of the pairwise interactions between each side-chain bead and the lipid particles. The leading contribution was individuated by constructing several PMFs, each one generated by setting all but one side-chain bead interactions to zero. In practice, this has the effect of only studying one interaction at a time between side-chain beads and lipid beads by switching off all the others.

The second strategy (which we will refer to as *strategy 2*) was intended to generate an effective potential for the side-chain single bead based on energy evaluation on a grid.

Namely, the multi-beaded amino acid side-chain was placed in the middle of a box, and concentric spheres with increasing radii r were created around it (figure 5.4). For each value of r several grid points were evenly distributed on the surface of the sphere with radius r. For each beadtype that is present in the lipid model we have evaluated the energy interactions with the side-chain by placing the lipid bead sequentially on the grid points. For each value of r an angular average was done, and the resulting distance-dependent potential was then fitted according to a LJ functional form to obtain effective average parameters. As we will discuss in detail later on, we anticipate that the first strategy was more successful, and the second one was successfully applied only in few cases.

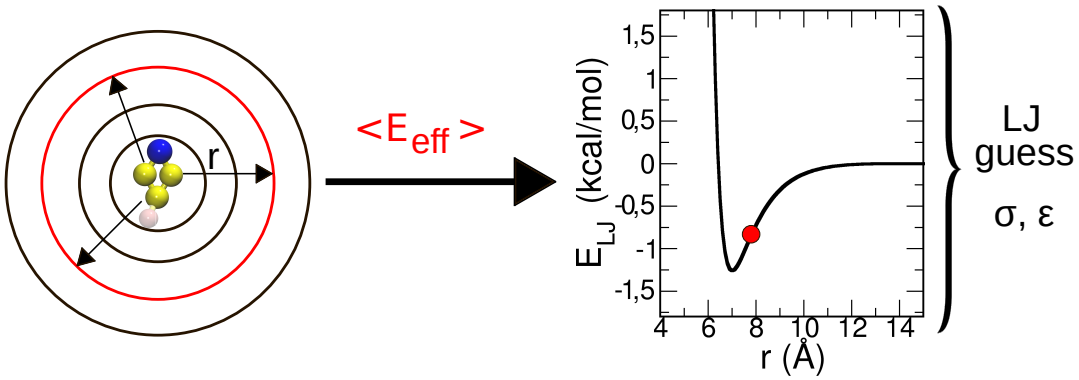

Fig. 5.4 Second strategy to obtain an effective potential between a side-chain and each of the lipid beads. A set of concentric spheres is located around the side-chain of the studied amino acid of interest. Lipid beads (in this example C1) are located along each sphere surface of radius r. For each sphere, an average effective potential is calculated in order to plot $E_{LJ}(r)$ (exemplified in red for a value of r). Finally, the potential is fitted using a LJ functional form to obtain a first set of parameters, σ and ε.

5.2 Parametrization of lipid-protein interactions

Here we present the results of the re-parametrisation for the lipid-protein cross interactions of the amino acids that have more than one side-chain bead in Dry Martini's model, and were re-mapped to have only one side-chain bead. The protocol followed for each amino acid that was parametrised consisted of the following steps:

1. Obtention of initial LJ parameters for the protein-lipid coupling, either strategy 1 or 2.
2. Calculation of PMF of amino acid extraction from the bilayer with the initial LJ parameters and comparison to the reference PMF^{ref} obtained using the Dry Martini force field.

3. Re-calculation of PMF with different epsilon values for the LJ (initial epsilon value multiplied by a factor). This step was repeated until the PMF matches the reference profile, particularly at the highest z values.

4. Check of radial distribution function of the side-chain bead and the beads from the hydrophobic core of the membrane to make sure the local packing in the membrane is preserved. Fix a new value of sigma if needed.

5. Re-calculation of the PMF with the new sigma value. If needed, this step was repeated with the fixed sigma and changing epsilon until the PMF matches to the reference profile.

See figure 7.3 as an example for the full process in the case of tyrosine.

To help follow the logic behind the first strategy to obtain a first set of Lennard-Jones interactions we have plotted the PMFs obtained when switching off each different beadtype one at a time for the different multi-beaded side-chain amino acids (Figure 5.6) at pH 7. In the case of HIS we have also plotted the resulting PMF obtained after using the Lennard-Jones parameters obtained from the effective potential (strategy 2 mentioned in previous paragraph) since it was used as a base to obtain the final coupling, although for the other amino acids the average potential was not giving good results. In all the cases, regardless from the parameter's origin, the initial epsilon values were systematically multiplied by a factor until fitting the reference value of the free energy of insertion of the amino acid side-chain (Figure 5.7). If needed, sigma was also altered afterwards. Secondly, the radial distribution function was checked to ensure that the interaction with the lipid was also treated correctly from a topological point of view and not just only energetic (Figure 5.8). Finally, a brief description for each amino acid follows.

5.2.1 Tryptophan (TRP)

In the case of tryptophan, there are only two different beadtypes that form the side-chain. Firstly, we followed the strategy of switching off all the interactions but one between one of these beadtypes and compared it to the reference PMF. As shown in Figure 5.6, the main contribution to the free energy of insertion of the amino acid in the membrane was the SC4 beadtype, which is an apolar bead. As a consequence, we used the already existing interactions between this beadtype and the lipid beads as a first guess for the new protein-lipid coupling. Several iterations were run until the best fit with the reference bead was found (see Figure 5.7)

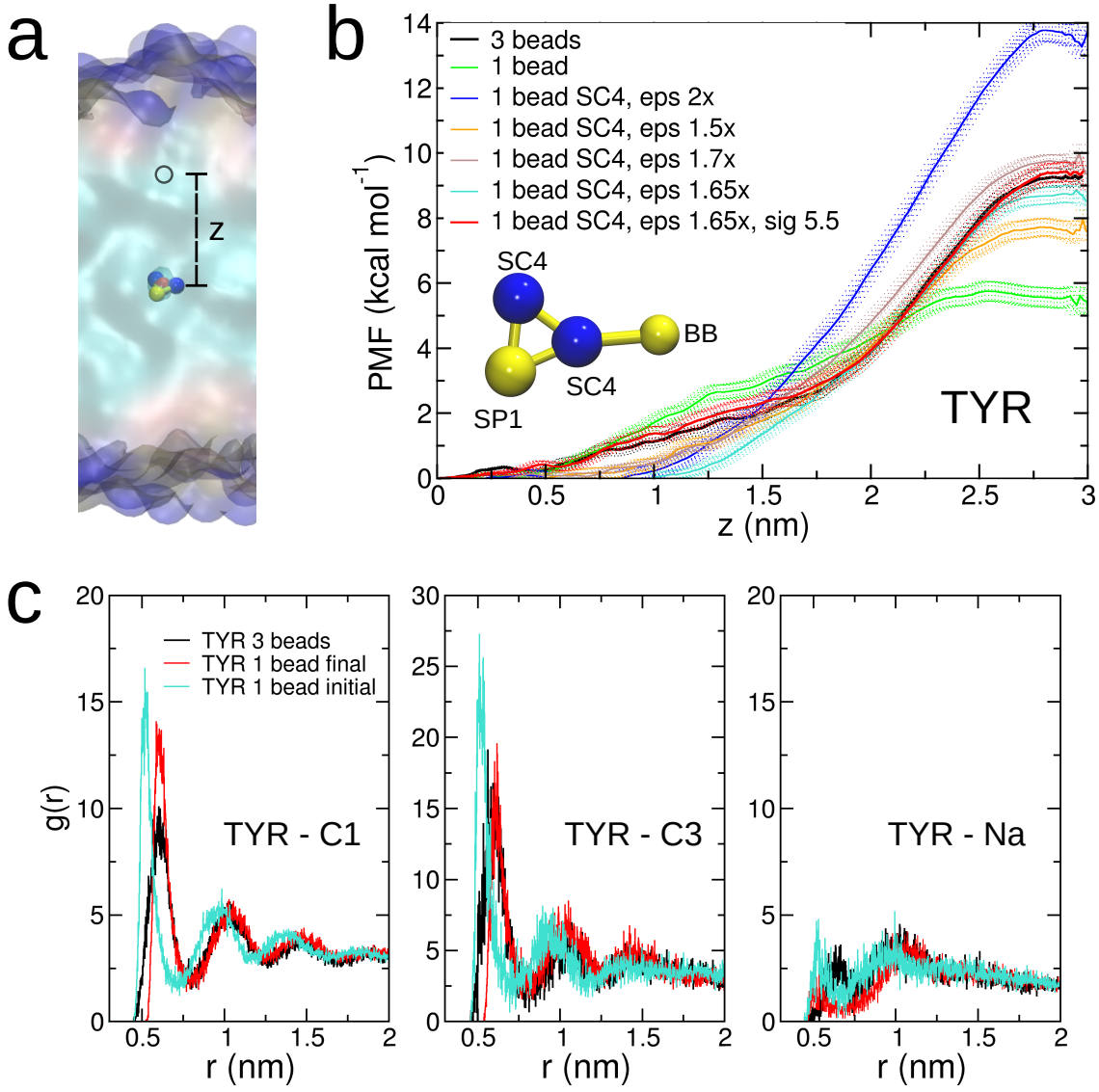

Fig. 5.5 Example of the whole parametrisation process in the case of tyrosine (TYR) as a representative amino acid. Panel a: Setup for PMF calculations showing the initial coordinates of the window corresponding to z = 0 nm (zoom on the region surrounding the amino acid side-chain). Panel b: Example of different PMFs obtained from the reference simulation and different LJ guesses either with strategy 1 (1 bead SC4) or strategy 2 (1 bead guess). For the first strategy several tests are shown with different ε (eps) and σ (sig) values. Panel c: Radial distribution functions between the side-chain and hydrophobic core beads present in the lipid, for the reference system (black), best PMF result after fitting when modifying ε values only (turquoise), and final result (red)

Consequently, it is important to also consider the distance distributions between the different beadtypes, so to account for the geometry of the system and not only the energy. To

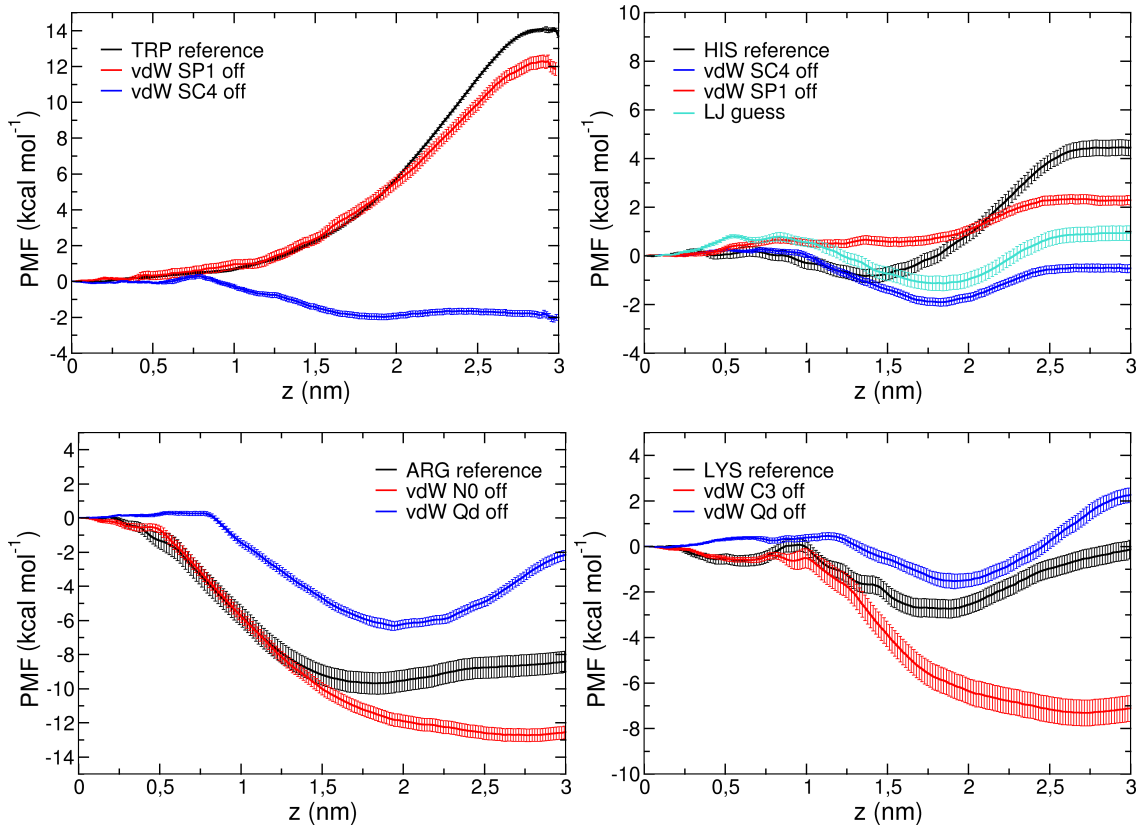

Fig. 5.6 PMFs obtained from switching off each of the beadtypes in TRP, HIS, ARG and LYS amino acids. The PMF obtained with the original parameters is plotted in black as a reference. TYR and PHE side-chain's PMF are omitted in these graphs as TYR is composed by the exact same bead types as TRP (and has an analogous curve) and PHE's side-chain beads belong to the same bead type.

do so, we have compared the radial distribution functions of the 1-beaded tryptophan versus the reference simulation. Changes in sigma were done accordingly to fit the reference radial distribution function correctly (Figure 5.8).

5.2.2 Other hydrophobic cyclic amino acids: tyrosine (TYR) and phenylalanine (PHE)

These amino acids were parametrised following the same strategy as tryptophan, since they have the exact same bead types. Thus, the first initial guess was based on the bead that contributed the most to the PMF curve (i.e SC4). A handful of simulations under different conditions for sigma and epsilon were done until the PMF was similar enough to PMF^{ref}.

(Figure 5.7) Radial distribution functions of the side-chain and different non-charged beads from the lipids were also checked to validate the implementation.

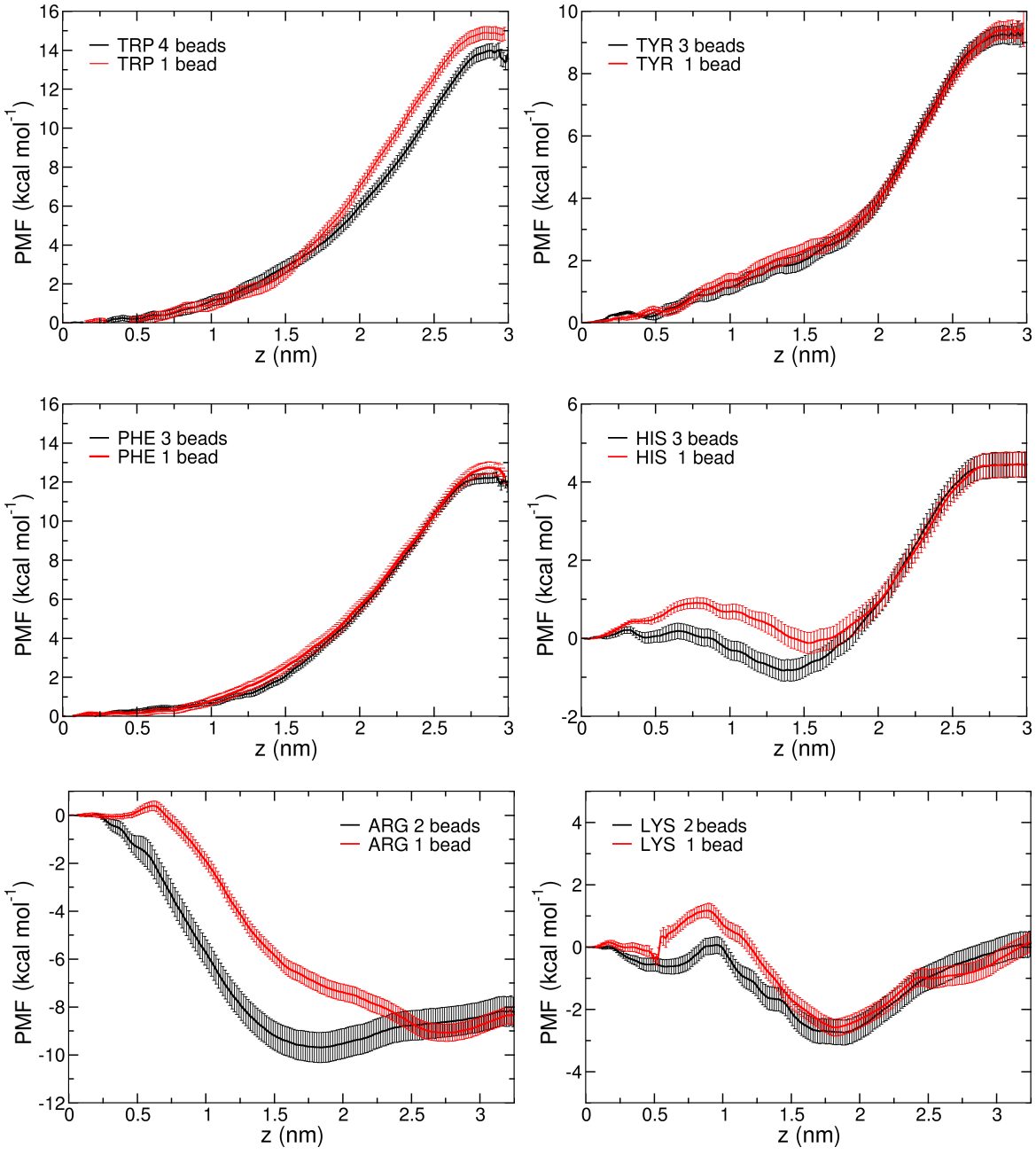

Fig. 5.7 Final PMF curves for the six treated amino acids. Black lines correspond to the reference simulation (PMF^{ref}) whereas red lines correspond to the simulation with the final parameters as shown in table 5.2

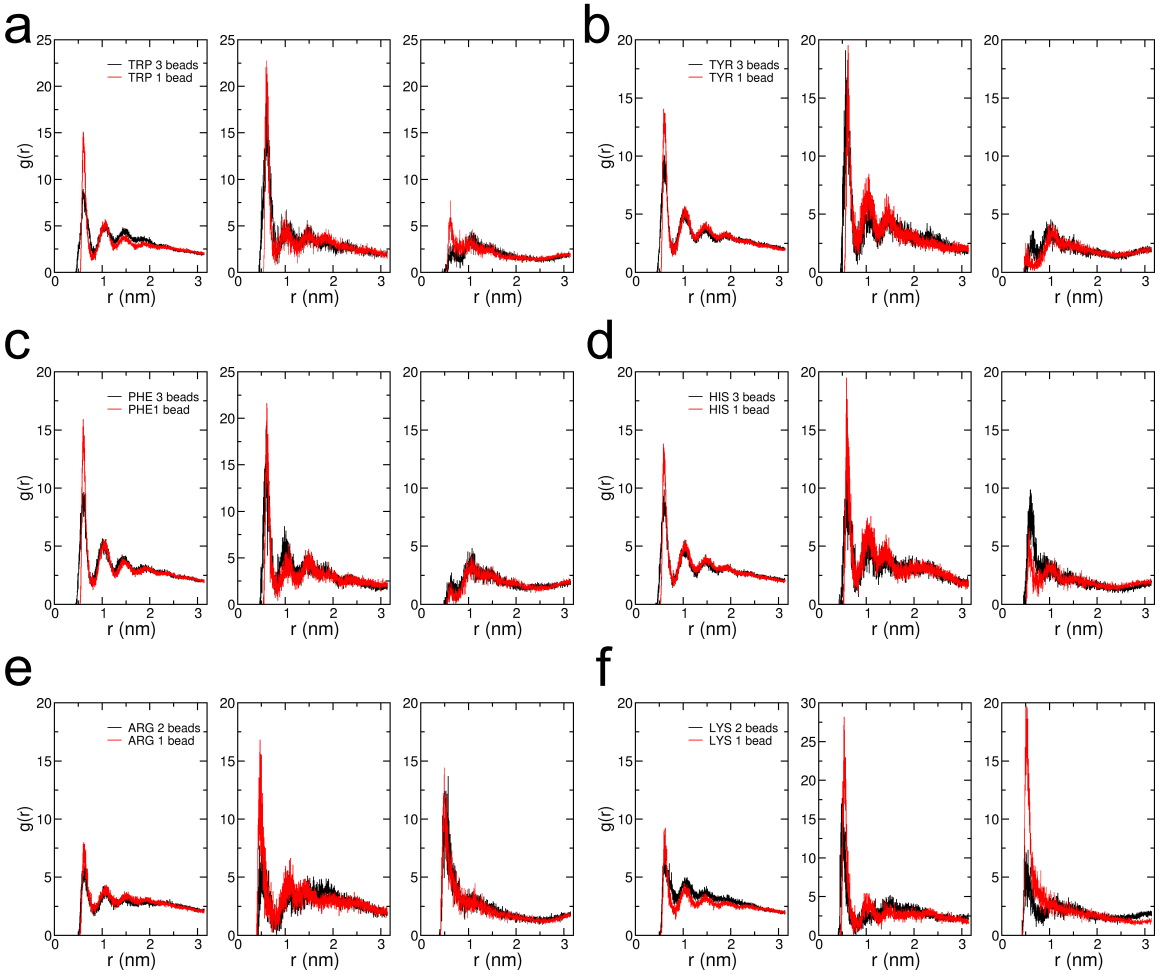

Fig. 5.8 Radial distribution functions between the side-chains of all the six multi-beaded side-chain amino acids and characteristic beads from the hydrophobic core of the membrane. From left to right the lipid beads that are being plotted are: C1, C3 and Na. Panels a-f correspond to the amino acids: tryptophan, tyrosine, phenylalanine, histidine, arginine, lysine

5.2.3 Histidine (HIS)

In the case of histidine, even if the polar bead (SC4) was still the main contributor to the free energy of insertion of the amino-acid in the bilayer, the values of sigma and epsilon obtained from the mapped effective potential was used as a first guess (LJ guess, turquoise line in Figure 5.6). This choice was made since the parameters obtained from the effective potential turned out to be a global better fit than iterating over the different bead type. As a consequence, not only the free energy of insertion was the same with respect to the reference run (Figure 5.7) but also the shape of the profile was in better agreement.

5.2.4 Charged amino acids: Lysine (LYS) and Arginine (ARG)

For these charged amino acids, the initial guess for the Lennard-Jones parameters came from the different bead types and subsequently refining the interactions between all the lipid beads and the side-chain bead of the amino acid. Both, residues kept the +1 charge for the coulombic interactions. Nonetheless, the interactions between the lipid tails and lysine were derived from the original non-charged bead that was present in the original 2-bead Dry Martini topology, however keeping the total positive charge. On the other hand, in the case of arginine, the PMF profile that fit best to the reference PMF was generated by the simulations done with cross-interactions based on the original charged bead type and lipids. It is relevant to emphasize however that even in the case where the parameters for the Lennard-Jones interactions between protein and lipid derived from the non-charged beads, the total charge of the residue at pH 7 was always kept. Neutral arginine and lysines were not considered at this point of the development.

5.2.5 Remarks and final considerations about re-parametrisation

The choice of parameters was made so that the resulting PMFs fit the the reference PMF generated with Dry Martini standard parameters (version 2.1) in Gromacs. Additionally, radial distribution functions between the amino acid side-chain beads and different lipid beads were checked to assure the correct geometrical distribution of the amino acid in the bilayer. Noticeably, radial distribution functions between side-chain and the beads that form the hydrophobic lipid tail (C1, C3) show an excellent agreement. Particularly for the position of the first peak. The full final set of Lennard-Jones parameters for the cross-interaction of these amino acids and POPC is presented in the following table 5.2

Lipid	TRP		TYR		PHE		HIS	
bead	σ	ε	σ	ε	σ	ε	σ	ε
type	(nm)	(kJ/mol)	(nm)	(kJ/mol)	(nm)	(kJ/mol)	(nm)	(kJ/mol)
C1	0.550	9.270	0.550	6.600	0.550	7.600	0.547	5.442
C3	0.550	9.270	0.550	7.425	0.550	8.550	0.548	4.172
Na	0.550	5.562	0.470	4.455	0.550	5.130	0.520	4.599
Q0	0.550	1.030	0.470	0.825	0.550	0.950	0.571	0.944
Qa	0.550	1.030	0.470	0.825	0.550	0.950	0.571	0.944

	ARG		LYS	
	σ	ε	σ	ε
	(nm)	(kJ/mol)	(nm)	(kJ/mol)
C1	0.580	2.400	0.570	3.600
C3	0.470	0.600	0.490	4.500
Na	0.470	0.850	0.470	4.050
Q0	0.600	3.400	0.600	3.600
Qa	0.700	4.590	0.700	4.860

Table 5.0 Table with final coupling parameters for all the six remapped amino acids with POPC beads

Finally, these parameters as well as the ones for the amino acids whose topologies have a single side-chain bead were used to run simulations with OPEP/Dry Martini in Muphy, the MD engine where OPEP is already available. Since at this moment of software development electrostatic interactions are not fully available in Muphy, tabulated potentials were used for the coupling between charged amino acids and the charged beads of the lipid. These tabulated potentials incorporate the electrostatic interactions and the Lennard-Jones interactions that were parametrised in this section.

5.3 Validation: Testing small transmembrane peptides

After obtaining a set of parameters to treat the non-bonded interactions between lipids and protein, we tested them in a more complex and realistic system such as transmembrane peptides. All the protein-protein interactions are treated internally via the protein force field (OPEP). Nonetheless, if the balance between protein and lipid interactions is incorrect it could affect the stability of the proteins. Moreover it could affect the orientation or motility inside the membrane. At this point it is worthy to mention some considerations that need to be taken when using the final lipid-protein parameters in the context of full proteins and not just side-chains. Interactions of the protein backbone with the lipids need to be defined. In OPEP protein force field the backbone is represented by 4 atoms: N, H, Cα, C and O are CG beads, however for the cross-interaction with Dry Martini lipid we opted to define just one interaction point at the position of Cα. The parameters for these interactions were taken from the ones that already exist for Dry Martini between lipid beads and backbone beads. Since Dry Martini backbone beads depend on the secondary structure assigned to the residue they belong, we initially opted to use the default values, i.e: the one corresponding to an isolated amino acid, which corresponds as well to the parameters associated to a coil structure. This decision was made to simplify the cross interaction parameters, taken into consideration that

the intra-protein interactions are already well taken care of by OPEP. Another special case to consider are the interactions between lipid and the protein caps (acetyl, amide, etc) that are sometimes used at the protein termini. In this version of the coupling they are neglected.

In the following sections we present three different systems that were put to a test.

5.3.1 System preparations and simulation setup and analysis

The peptides were modelled as capped canonical alpha helices with Swiss-PdbViewer [103] according to the following sequences: $acetyl-L_{10}-amide$ (for polyleucine), $acetyl-GWW(LA)_8LWWA-amide$ (for WALP23) or $acetyl-GKK(LA)_8LKKA-amide$ (for KALP23). Subsequently, they were manually embedded in a pre-equilibrated POPC bilayer that already had a hole in the middle. A short equilibration was run, keeping the protein fixed to allow the membrane to adapt to the newly inserted protein without distorting the protein. Consequently, production runs were run for 50 ns to analyse the system's dynamical behaviour. All of the systems simulated with standard Martini v2.1 with a timestep of 10 fs, or Dry Martini v2.1 with a timestep of 3.5 fs were run in Gromacs 4.5.5 [256] for compatibility purposes. All atom simulations with charmm36m force field [117] were run in Gromacs 5.1.4 [1] with a timestep of 2 fs using LINCS and the simulations with our new coupling (Dry Martini lipids coupled with OPEP protein) were run in Muphy [26] usign mass-rescaling for the H atoms and with a timestep of 3.5 fs. On top of the normal bonded/non-bonded terms of any force field, Martini and Dry Martini force fields use an elastic network and define different bead-types for the backbone beads according to the secondary structure of the protein to ensure the protein's stability. Because of this, in the version of this manuscript we have treated the cross-interaction of protein backbone and lipids as the default standard parameters for the backbone of the amino acid without a particular secondary structure (from now on referenced as *stdBB* in the plots). In the case of WALP, we have also run a reference simulation with Dry Martini where the backbone beadtype was assigned according to the one that comes from the secondary structure assignment in the "wet" Martini force field (referenced as *BBaswet* in the plots). For the systems run with OPEP/Dry Martini in Muphy, the OPEP parameters corresponded to the v 4.0 [201] with a scaling factor of 1.8. All simulations were run at 310 K in a 65x65x130 box with 3D periodic boundary conditions. Simulations were run in the NVT ensemble, except for the atomistic simulation that was run in NPT only. In the case of standard wet Martini, simulations were also run in the NPT ensemble, since it will be the ensemble of choice for membrane systems. For the systems that were simulated in the NPT ensemble, pressure was coupled semi-isotropically, with different coupling for the plane of the membrane and the plane normal to the bilayer. To characterise the simulations we followed mainly three different elements: the location of the helix inside the bilayer and

the orientation of it regarding the bilayer normal and the end-to-end distance of the peptide. Another descriptor was used just for the simulations run in all atom and with the OPEP/Dry Martini coupling where we also analysed the structural stability of the embedded peptides by computing the secondary structure with STRIDE [92] as a function of time. The positioning of the helix inside the bilayer was plotted as the center of geometry of all the alpha carbon (or backbone beads in the case of Martini or Dry Martini force fields) of the protein inside the bilayer along the whole simulation. Additionally, the center of mass of the upper and lower phosphate layer was also plotted to easily follow the relative location of the peptide inside the membrane. The end-to-end distance was simply calculated as the distance between the initial and final alpha carbon of the helix under study (or backbone bead for Martini and Dry Martini). Finally, the tilt angle was obtained as the dot product between the vector of the end-to-end distance and the normal to the bilayer.

5.3.2 The case of polyleucine: a stable polypeptide inside the bilayer

Several experimental studies have been carried out to characterise a minimal α-helix motif that can stably stay embedded in the bilayer [68]. Such is the case of polyleucine, a peptide that is known to precipitate in solution, but adopts an α-helix motif inside the membrane. Particularly here we focused on $acetyl - L_{10} - amide$ (L10), a 10-residue long polyleucine peptide that is acetylated in its N-terminal and amidated in its C-terminal.

Unbiased simulations run with either Martini or Dry Martini reported that the L10 peptide, even if it remains stably inside the membrane, has the tendency to stay in the edge of it (Figure 5.9,left, orange and turquoise lines). The same behaviour was observed for the OPEP/Dry Martini coupling (black line). On the other hand, the all atom simulation reported a very stable behaviour in comparison, where the polyleucine helix has barely moved from the initial position. Following the same trend, for the tilt angle the fully atomistic simulation reported a very stable behaviour (Figure 5.9, right, blue line). All the other coarse-grained simulations showed a much more fluctuating behaviour. Interestingly, the simulation run with OPEP/Dry Martini coupling showed the most similar behaviour with respect to the fully atomistic simulation, although the fluctuations are much higher (black line). Moreover, also its initial α-helical secondary structure was conserved during the simulation (as well as the all atom reference simulation).The results obtained for this minimalistic system show that in all the simulations the peptide stays inside the membrane, but in the coarse-grained simulations it shows more freedom to rearrange its orientation inside the bilayer. Since it is such a short peptide and composed of just leucines we have expanded the study to other more challenging peptides.

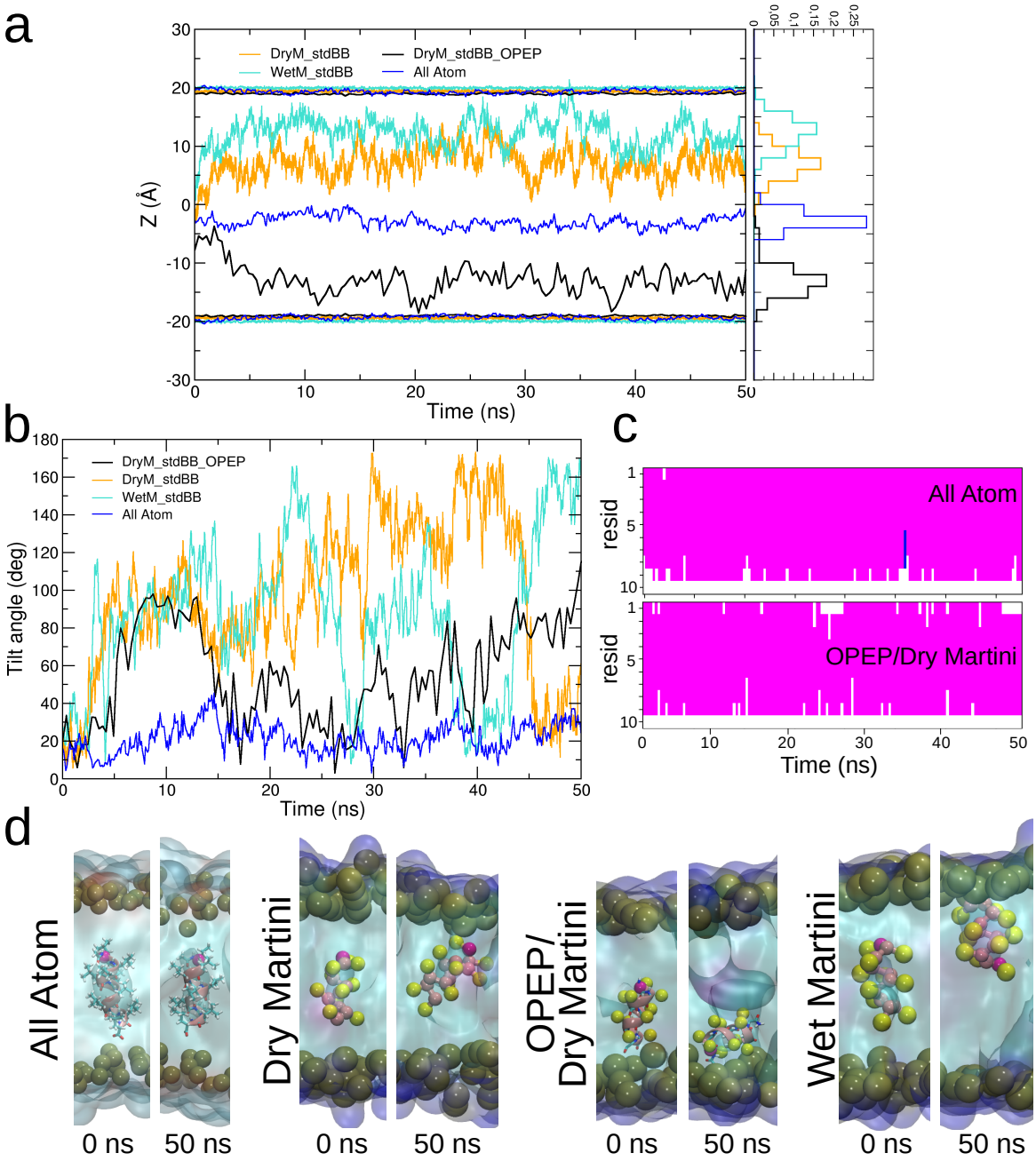

Fig. 5.9 a: Position in Z-axis of the center of mass of L10 during the simulation with its associated histogram. The average position of the center of mass of the phosphates in the upper and lower monolayer are also plotted to have as a reference. b: L10 tilt angle with respect to the normal of the bilayer. c: Secondary structure assignment for simulations in All Atom and with OPEP/Dry Martini coupling. d: Initial and final snapshots showing the peptide inside the bilayer for each of the force fields used. The protein backbone is shown as pink and side-chain beads in yellow for all the coarse-grained models. The first backbone atom for the peptide is coloured in dark pink to recognise the orientation of the helix. To help visualisation phosphate beads are shown as vdW spheres in tan and the membrane is depicted as a transparent surface.

5.3.3 Flanked peptides

Another set of engineered transmembrane peptides have been created to help the orientation inside the bilayer[68]. To do this, certain amino acids are put in the extremities of the peptide to anchor it in the right position to the bilayer, either by charged residues or hydrophobic ones.

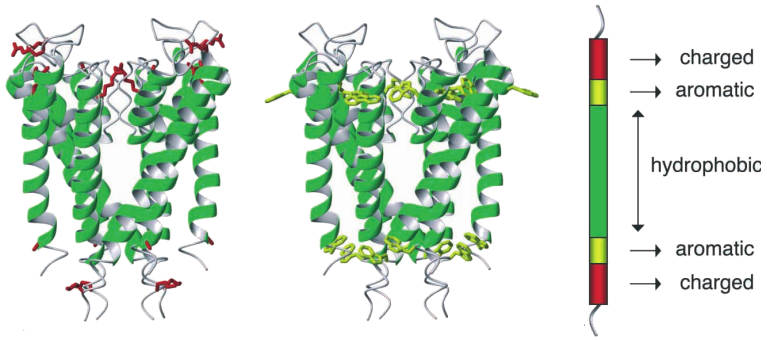

Fig. 5.10 Example of amino acid distributions in the transmembrane region of a potassium channel (*Streptomyces lividans KcsA potassium channel*). Charged amino acids are shown in red whereas hydrophobic amino acids are depicted in green and aromatic in yellow. *Taken from [68]*

WALP

The so called WALP peptides owe their name to their amino acidic composition. These peptides have an hydrophobic core of alanine-leucine repeats and are flanked by tryptophan in their extremities. In this work, we have particularly focused on WALP23, which is a 23-residue long peptide with the following sequence: $acetyl-GWW(LA)_8LWWA-amide$. This peptide had been chosen to test the behaviour of our first coupling parameters since it has experimental information available and has also been assessed before in different force fields.

WALP23 simulations were processed in the same way as polyleucine, but we have added the study of the end-to-end distance as well as the secondary structure in this case since this peptide is longer and could potentially suffer deformations due to a very strong coupling with the lipid. The end-to-end distance of all the systems remained stable during the simulation (Figure 5.11). In particular, for the simulation with OPEP/Dry Martini that means that the newly defined cross interactions between the peptide and the lipids did not alter the inner stability of the peptide, at least regarding its length. This stable behaviour was also supported by dynamics of the peptide's secondary structure (Figure 5.11, right), where the α-helix

was not deformed during the simulation with OPEP/Dry Martini and has virtually the same behaviour as the reference all atom simulation.

When studying the positioning of the peptide inside the membrane, the only simulation that showed more movement inside the bilayer was the one run with Dry Martini, where the vdW interactions of the backbone beads with the system were defined as if they were forming part of a coil (or an isolated amino acid, label:stdBB). In Figure 5.12, the orange line refers to that system where the protein tends to go to one extreme of the bilayer. However, if the backbone beads' interactions were defined by the secondary structure that they have assigned accordingly (usual Martini way of defining the backbone beads, labelled BBaswet), the peptide remains stable in the middle of the bilayer (green line), as it also happens with standard Martini simulations (WetM, cyan line) and all atom (blue line).

Finally, the plot depicting the tilt angle vs time (Figure 5.12, right) shows two main different behaviours. The fully atomistic simulation and the simulations run with Martini (WetM_stdBB, turquoise line) or Dry Martini with the backbone beads treated according to the assigned secondary structure (DryM_BBaswet, green line) follow the same trend. In these cases the tilt angle averaged between 30-50ns is $18° \pm 5°$ (fully atomistic), $13° \pm 6°$ (standard Martini) and $34° \pm 6°$. These values are mostly in agreement with experimental values obtained by fluorescence that lay around 20-25 degrees [112]. However, the tilt angle obtained in the simulations run with the OPEP/Dry Martini coupling as well as the simulations run in Dry Martini with the secondary structure-independent parameters show a clear difference. In these cases, the peptide rotates to a position where it is closer to be perpendicular to the normal axis of the bilayer. The average values in the region between

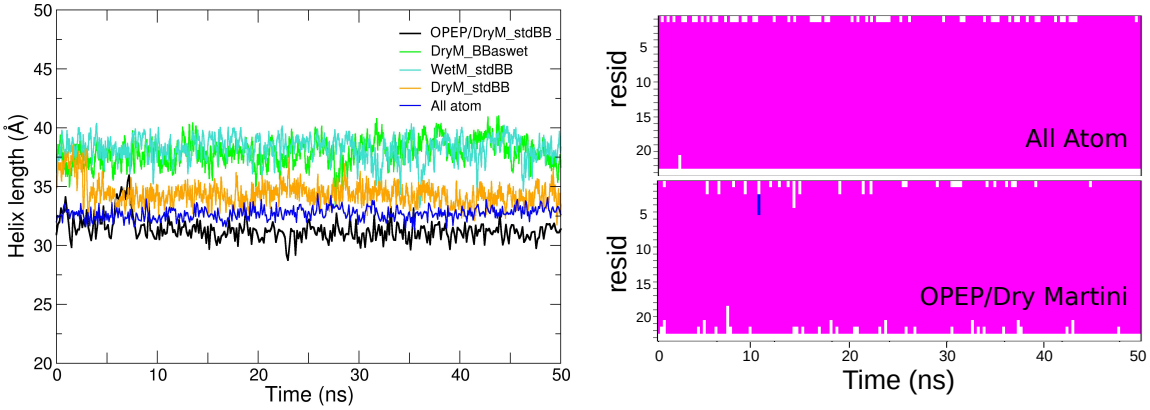

Fig. 5.11 Left:End-to-end distance of WALP helix: distance between first and last alpha carbon in the peptide (All atom, OPEP) or between the first and last backbone beads (Dry Martini, Martini). Right: Secondary structure assignment plot for the WALP23 run in All atom and with OPEP/Dry Martini coupling; magenta: α-helix, blue: 3-10 helix, white: coil

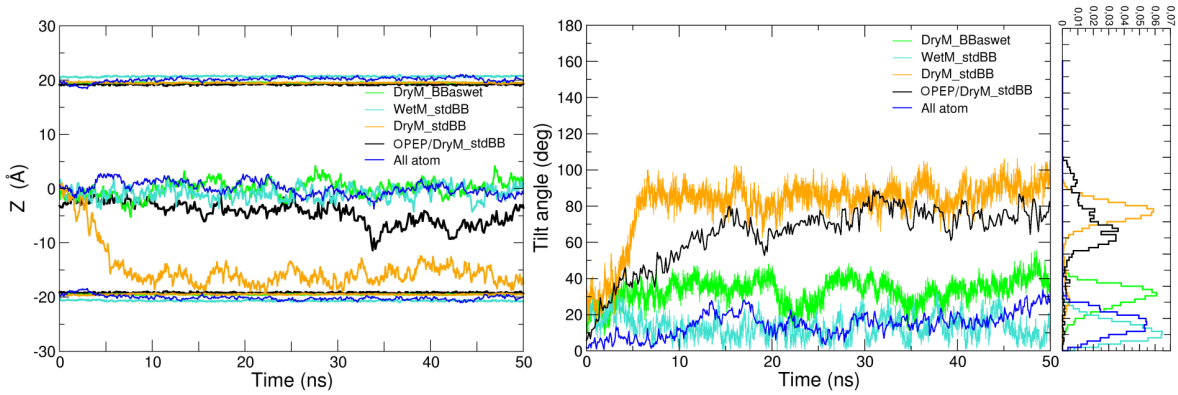

Fig. 5.12 Left:Position in Z-axis of the center of mass of WALP23 during the simulation. The average position of the center of mass of the phosphates in the upper and lower monolayer are also plotted to have as a reference. Right: WALP23 tilt angle with respect to the normal of the bilayer with the associated histogram from the angle population.

30-50 ns are $88^\circ \pm 7^\circ$ and $76^\circ \pm 6^\circ$ for Dry Martini and the simulation with the new coupling respectively. These values imply that the tilt angle is closer to 90 degrees with respect to the bilayer normal, which is in complete disagreement with either experimental results [112, 244] or other reported simulations [129, 170]. These results suggest that the current coupling is not good enough to study this system with default backbone parameters for the cross interaction, but that secondary-structure-dependent parameters are a better fit.

KALP

KALP23, is another type of transmembrane peptide that has been extensively assessed and frequently used as a model to check parametrisation of force fields [98, 129, 170, 244]. It is analogous to WALP in its sequence, but using lysines instead of tryptophans: $acetyl - GKK(LA)_8LKKA - amide$. By using, lysines it ensures positive charges that will interact strongly with the phosphate beads of the bilayer. Following the same strategy as for WALP23, the peptide was initially embedded perpendicular to the bilayer and its orientation and length followed along the simulation.

The first thing that we have evaluated here was the helix length to have an overview of the conformational stability of the helices during the simulation. As it can be seen in Figure 5.13, all the coarse-grained systems have a very stable length throughout the simulation, whereas the fully atomistic simulation has more fluctuations and an reports an increased value in the end-to-end distance but remaining stable nonetheless throughout the simulation(blue line). This increased in the end-to-end distance could be assigned to the local loss of α-helix in the C-terminal, that happened during the equilibration process. Furthermore, the α-helix

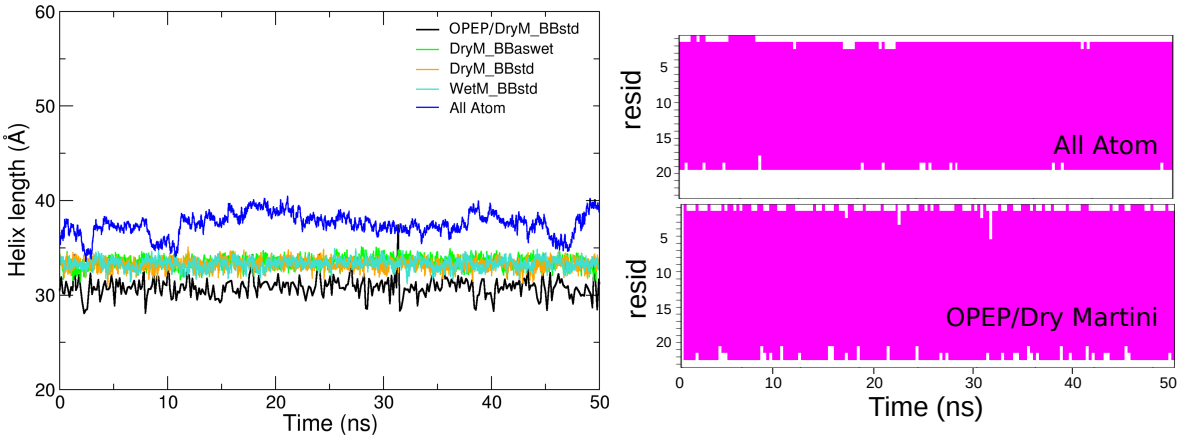

Fig. 5.13 End-to-end distance of KALP helix: distance between first and last alpha carbon in the peptide (All atom, OPEP) or between the first and last backbone beads (DryMartini, Martini). Right: Secondary structure assignment plot for the KALP23 run in All atom and with OPEP/Dry Martini coupling; magenta: α-helix, blue: 3-10 helix, white: coil.

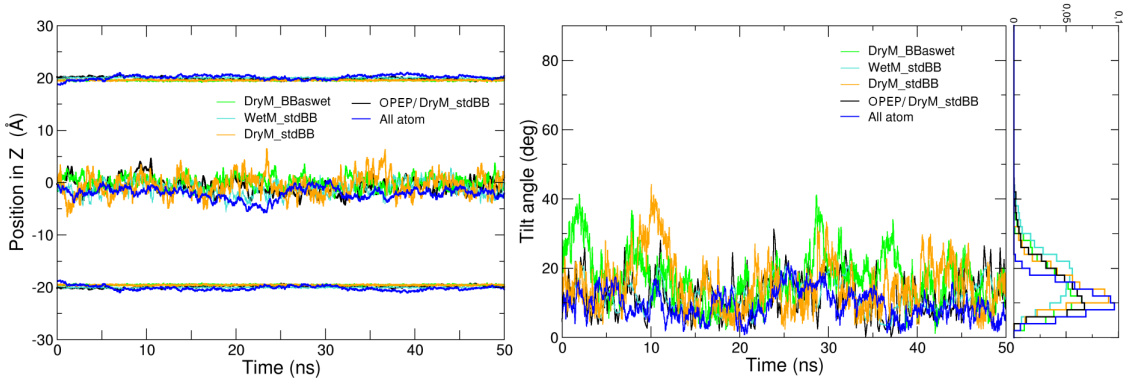

Fig. 5.14 Left: Position in Z-axis of the center of mass of KALP23 during the simulation. The average position of the center of mass of the phosphates in the upper and lower monolayer are also plotted as a reference. Right: KALP23 tilt angle with respect to the normal of the bilayer with its corresponding histogram.

was also conserved for the OPEP/Dry Martini coupling respect to its initial conformation, following the same trend as the all atom reference simulation. If we analyse the positioning inside the membrane, it can be seen that for all the systems the center of mass of the protein fluctuates minimally around the center of the bilayer (Figure 5.14, left). Moreover, the tilt angles observed for all the simulations have the same qualitative behaviour and even similar distributions (Figure 5.14, left). Particularly, for the case of the new parametrisation the peak of the distribution is around 10° with an average value of $12.6° \pm 6.6°$ comparable to the all atom simulation ($9.1° \pm 4.0°$). More importantly, these values are also in agreement

with previously reported angles in the range of $7^\circ - 17^\circ$ [98] that resulted from molecular dynamics simulations.

5.4 Conclusion

In this chapter we have described the strategy and guided the reader through the search and implementation of a set of parameters to couple Dry Martini lipid forcefield to protein OPEP force field. In this instance we have used the Dry Martini energy scale as a basis for the coupling. After obtaining a first set of cross-parameters we have put them to test for a series of known transmembrane peptides. At this point, it is important to consider some technicalities regarding the comparison of the simulations with different force fields. Although the systems that were run had the same initial configuration (same initial coordinates, or same equivalent coarse-grained coordinates), the time steps, initial velocities, solvation and softwares differ between the different force fields. However, at this stage of the development we aim to compare trends and qualitative behaviours of the peptides' properties when embedded in a membrane. Currently, we are simulating more replicas for each peptide/membrane system to be able to statistically analyse the results. Preliminary results for OPEP/Dry Martini show the same trends in orientation and secondary structure assignment. We have shown that given these first sets of parameters it is possible to run simulations of peptides embedded in the membrane without altering them (e.g, neither the peptide length nor the secondary structure were significantly altered). Unfortunately, in the case of the flanked peptide WALP23, the tilt angle shows an unrealistic value. This could be improved by using the Dry Martini backbone parameters derived from the secondary structure for the cross interaction and not the general ones, as simulations run with Dry Martini and these parameters show similar behaviour to atomistic simulations (figure 5.12, green line). However, we aim to have a set of structure-independent parameters for the cross interactions between lipid and proteins to have a generalisable force field. For that reason, a new set of better Lennard-Jones parameters should be derived for tryptophan, that seems to be the original cause of that behaviour. First optimising the backbone parameter only to have a better sequence-independent one, and then refine the side-chain parameters if it is still needed. On the other hand, the results derived from the simulation with KALP23 were very promising, and we are extending the case to study the dimerisation of that monomer inside the bilayer.

As a general final note, since the Dry Martini parameters for protein are still under development by their authors it could be wise to refine these set of parameters in a second round using atomistic PMFs as a more reliable reference to generate a universal set of parameters for the cross interaction. Furthermore, more advanced and efficient approaches

can be used to further refine this first set of coupling parameters presented here, such as neural network and artificial intelligence approaches like machine learning that have shown to be useful lately in similar parametrisation problems [21, 32, 147, 163].

5.5 Supplementary Material associated to this chapter

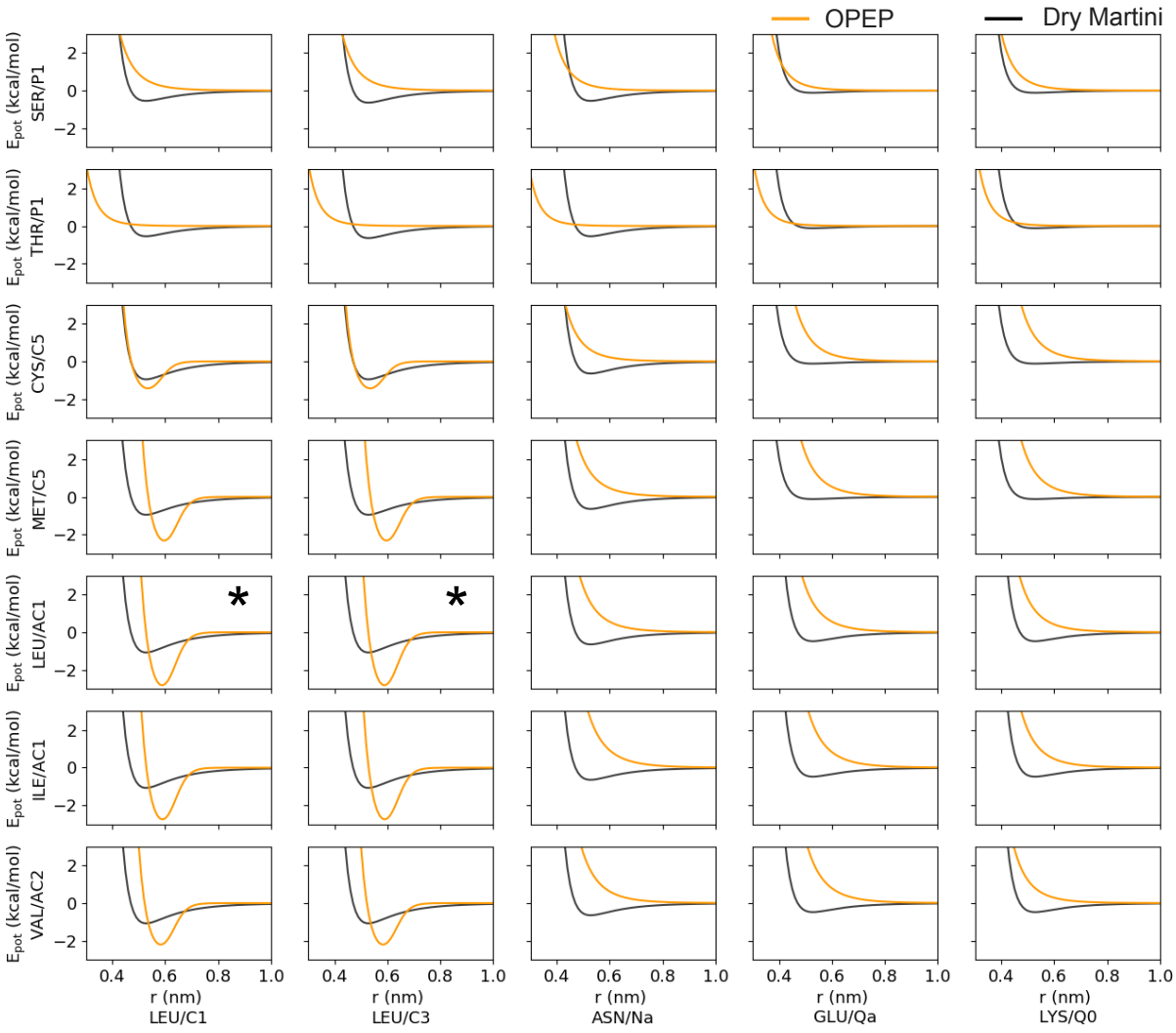

Fig. 5.15 Plot with all the small polar and apolar amino acids with POPC lipid beads.

Amino acid	Side-chain beadtype
ASP	Qa
GLU	P5
ASN	P5
GLN	P4
THR	P1
SER	P1
MET	C5
CYS	C5
PRO	AC2
VAL	AC2
LEU	AC1
ILE	AC1
ALA*	P4
GLY*	P5

Table 5.1 Table assigning the side-chain beadtype for each amino acid that has only one side-chain bead in Dry Martini. The asterisk marks the special case of amino acids with no side-chain bead that have a particular beadtype assigned to their backbone bead. In all the other cases the default backbone bead is P5.

	Qa		AC1		AC2		C5	
	σ (nm)	ε (kJ/mol)	σ (nm)	ε (kJ/mol)	σ (nm)	ε (kJ/mol)	σ (nm)	ε (kJ/mol)
C1	0.620	2.000	0.470	4.500	0.470	4.500	0.470	4.000
C3	0.470	0.500	0.470	4.500	0.470	4.500	0.470	4.000
Na	0.470	0.500	0.470	2.700	0.470	2.700	0.470	2.700
Q0	0.600	2.000	0.470	2.000	0.470	2.000	0.470	0.500
Qa	0.600	2.300	0.470	2.000	0.470	2.000	0.470	0.500

	P1		P4		P5	
	σ (nm)	ε (kJ/mol)	σ (nm)	ε (kJ/mol)	σ (nm)	ε (kJ/mol)
C1	0.470	2.300	0.470	0.500	0.470	0.500
C3	0.470	2.700	0.470	1.000	0.470	0.500
Na	0.470	2.300	0.470	1.000	0.470	0.500
Q0	0.470	0.500	0.470	0.500	0.470	0.500
Qa	0.470	0.500	0.470	0.500	0.470	0.500

Table 5.2 Table with all the already existing cross-interactions between POPC lipid beads (C1, C3, Na, Q0, Qa) and all the different beadtypes associated to the one-beaded amino acids.

Chapter 6

Conclusions and perspectives

In the work presented in this PhD thesis we have covered a variety of aspects that are of interest for the membrane protein field.

Firstly, we have explored the initial steps in membrane tethering in the context of yeast mitochondrial membrane fusion by the means of coarse-grained simulation approaches that are already available (Chapter 3). Using the previously published homology model for the monomer, we built three independent cis-assemblies of Fzo1 models that gave rise to three independent trans-oligomeric models of mitofusin, which resulted in three distinct scenarios of Fzo1 oligomerization that are not connected to each other. Each scenario involves two main components of mitofusin auto-oligomerization: the GTPase domain interface [46, 271] and the trunk domain interface[135]. We have shown that the narrow opening head-to-head mechanism could employ a trans-oligomerization strategy that turns out to fit particularly well with the past and current literature. The proposed HR2 antiparallel trans interaction in the Fzo1 complex is in agreement with a similar antiparallel heptad repeat interaction observed for Mfn1 [135]. Moreover, the tethering distances between outer membranes have been reported to reach 6 to 8 nm [37], corresponding to a macromolecular assembly of Fzo1 oligomers reminiscent of those described in the present study. They are comparable to the tethering distance of 5.5 to 9.5 nm imposed by this Fzo1 oligomerization model. In addition, our models provide sufficient detail to derive experimentally testable hypotheses for future work, for instance in terms of possible crosslinks and surface-exposed residues. Similarly, our models could be further improved based on new experimental observations. More generally, by making these models (dimers, and tetramers) available to the community we are expanding the knowledge and giving realistic structural hypothesis of full length mitofusins and their known or yet to be described cofactors, as well as their precise oligomerization properties.

Secondly, we have implemented a new coupling between an implicit solvent coarse-grained model for lipids simulated at the molecular dynamics level and a lattice Boltzmann

representation of the solvent hydrodynamics (Chapter 4). This implementation was presented with a showcase of applications, that can be of advantage to solve biological questions in the future.
In this work we extended the operational capability of the multi-scale Lattice Boltzmann molecular dynamics technique developed in the laboratory [54, 55, 143, 239, 240]. With our approach we are now able to explore a great variety of biophysical processes involving lipid systems at the mesoscale where solvent mediated interactions are important. More interestingly, our technique can complement the experimental investigation of lipidic systems for bio-engineering and biotechnology applications, such as vesicles in fluid flow conditions [11, 168]. In this context, we have presented a rudimental example of a Nanoreactor where we studied the effect of crowding on the diffusion of proteins. The idea to study lipid systems that contain proteins lead to the development of a more rigorous effective coupling between the Dry Martini force field for lipids and the coarse-grained protein model OPEP.

Finally, we have initiate a systematic coupling between the Dry Martini lipid model and the OPEP force field for proteins (Chapter 5). To create this coupling it was necessary to design an effective strategy to treat the cross-interactions of the lipids and proteins, and in particular a new resolution scheme had to be implemented. A first version of this coupling is proposed but must be refined using more statistics and test-cases.
The advantage of our coupling is that, since OPEP allows for protein conformational changes and do not require extra bias to stabilize protein structures, it will allow to inquire biophysical processes where these conformational fluctuation are important in membranes. This is at variance with other exiting coarse-grained force fields that allow to simulate protein and lipids together at quasi-atomistic resolution [15, 22, 170].

In summary, at the final stage of this project we produced a lipid model that can be used for simulations in Muphy and that could be coupled to hydrodynamic interactions. This model has been extensively tested and is up and running. On the other side, the parameter set obtained for the protein-lipid coupling are in its first version (v.0) so, further refinement will be needed to obtain better agreement with experimental results, in particular for the case of tryptophan, which we are still working to improve prior to publication. In contrast, the study of mitochondrial fusion with "standard" coarse-grained models gave a first glimpse into the possible organisation of the tethering complex under very simple conditions (just Fzo1 alone and membrane) however relevant enough to give a structural feedback to the community working on mitochondrial dynamics. The short-term plan is to continue with a more detailed representation of the system of choice (the trans-antiparallel, cis head-to-head parallel tetramer) as well as the modelling of the heptad-repeats alone and their

oligomerization as fragments in a membrane context (on-going work by our experimentalist collaborators).

In conclusion, we have characterised the first stages of mitochondrial fusion and debunk existing theoretical models that have been extensively shown in the literature without any rigorous structural consideration. From a methodological point of view, by considering the coupling with hydrodynamics via the Lattice Boltzmann technique, we have expanded the biological processes that can be investigated using the effective Dry Martini force field including situations where the external perturbative fluid is key. Finally we have started to couple the OPEP protein force field to Dry Martini lipid model in order to explore biological processes where membrane protein conformational changes are important.

Chapter 7

Résumé de la thèse

7.1 Introduction générale au résumé

Dans les pages suivantes, nous donnerons un aperçu des principaux sujets traités dans cette thèse et des résultats obtenus. Pour faciliter la compréhension, nous suivrons le même ordre que dans le texte principal. Pour cette raison, ce résumé est divisé en trois parties, liées aux trois principaux axes de contenu. Le point commun est dans tous les cas l'utilisation de modèles gros grain pour représenter les biomolécules, en particulier les membranes et les protéines membranaires. L'approche multi-échelle est également présentée en relation avec les simulations lipidiques.

7.2 Partie 1: Complexes de fusion de membranes mitochondriales externes

Les mitochondries sont des organites dynamiques organisées en réticulum cytoplasmique. Elles fusionnent leurs membranes externes et internes pour former des tubules. Ces tubules mitochondriaux peuvent s'interconnecter par fusion mais peuvent aussi se fragmenter par fission et ainsi générer un réseau extrêmement dynamique. Ensemble, la fusion et la fission régulent la morphologie et l'évolution du réseau mitochondrial, ce qui rend ces processus essentiels au maintien de l'intégrité et des fonctions mitochondriales [89, 230, 252, 267].

D'un point de vue mécanistique, la fission mitochondriale ainsi que la fusion des membranes externe et interne sont médiées par des membres de la famille des protéines liées à la dynamine (DLP ou DRP) [203]. Ces grandes GTPases agissent comme remodelatrices des bicouches lipidiques intracellulaires grâce à deux propriétés : leur capacité à se lier aux membranes biologiques et leur propension à l'oligomérisation en structures macromoléculaires

d'ordre supérieur. Conformément à ces caractéristiques, les DRP de fission sont recrutées par des protéines adaptatrices spécifiques au niveau des membranes externes mitochondriales, où elles s'auto-oligomérisent grâce à la liaison du GTP pour former des spirales macromoléculaires qui s'enroulent autour des tubules mitochondriaux [252]. L'hydrolyse subséquente du GTP induit des réarrangements conformationnels des dynamines, ce qui entraîne une réduction du diamètre des spirales et un étranglement des tubules mitochondriaux jusqu'à leur séparation. Les DRP de fusion sont des protéines transmembranaires qui favorisent la fusion homotypique des bicouches lipidiques dans lesquelles elles sont insérées [58]. Les mitofusines MFN1 et MFN2 (Fzo1 dans la levure) permettent la fusion des membranes externes mitochondriales, tandis que la fusion entre membranes internes est médiée par OPA1 (Mgm1 dans la levure). L'atlastine ATL-1 (Sey1 dans la levure), qui fusionne les membranes des tubules ER, est un autre membre clé des DRP de fusion. Les mitofusines et les atlastines ont en commun la capacité d'auto-oligomériser en trans (c'est-à-dire à partir de bicouches lipidiques opposées), de façon à lier et à hydrolyser les bicouches lipidiques opposées et à promouvoir leur fusion [58]. Les structures cristallines récentes de portions de protéines dépourvues des régions transmembranaires ATL-1, Sey1 et MFN1 permettent en outre d'interpréter l'assemblage des DRPs pendant la fusion membranaire comme des dimères protéiques interagissant en trans à travers leurs domaines GTPase respectifs [30, 44, 46, 172, 270, 271] Cependant, la mitofusine de levure Fzo1 forme également des complexes de poids moléculaire élevé [204]. En particulier, il a été démontré que Fzo1 déclenche la formation d'un complexe macromoléculaire en forme d'anneau pendant l'amarrage mitochondrial [37], ce qui suggère que l'assemblage des DRP pendant la fusion des membranes pourrait impliquer la formation de complexes de grande envergure plutôt que seulement des homodimères transmembranaires.

Conformément à cette possibilité, les mitofusines forment des oligomères à la fois en cis (sur la même bicouche lipidique) et en trans pour médier la fixation et la fusion des membranes [8, 46, 101, 120, 135, 214, 230]. En plus du domaine GTPase, les mitofusines comportent également deux domaines de répétitions d'heptades (HR1 et HR2), pouvant être impliqués dans des interactions homotypiques pendant la fixation mitochondriale [70, 101, 135]. Il reste à savoir si les interactions GTPase et HR se produisent en cis ou en trans, avant l'ancrage mitochondrial ou pendant l'ancrage mitochondrial [84, 135]. Dans cette thèse nous nous appuyons sur un modèle récent de Fzo1 [70] pour étudier des hypothèses distinctes concernant la formation des oligomères de mitofusines en cis et en trans. En plus d'être le seul modèle presque complet actuellement, ce modèle contient un domaine membranaire, ce qui est nécessaire pour étudier les orientations de la mitofusine au sein d'une bicouche. Nous avons utilisé une méthode de modélisation guidée par les données expérimentales disponibles

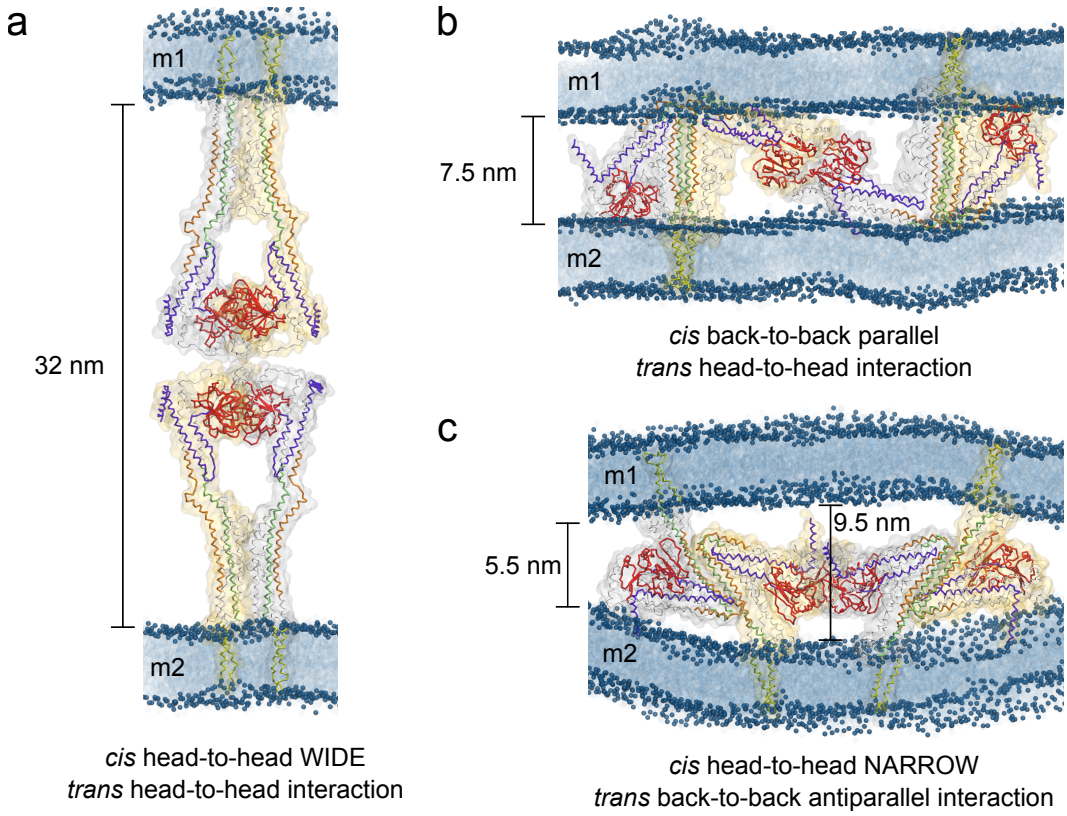

Fig. 7.1 Configurations finales des complexes de modèles putatifs pour les interactions trans théoriques Fzo1 après relaxation par simulations de dynamique moléculaire. a. Les modèles Fzo1 en conformation ouverte, comme suggéré par l'analogie du système BDLP, interagissent à travers leurs domaines GTPase. b. Le transtétramer Fzo1 dans lequel les HRs sont orientés en parallèle et l'interaction trans se produit vers le domaine GTPase. c. Le trans-tétramer Fzo1 en conformation fermée. Ici, l'interaction se produit vers le domaine GTPase ainsi qu'à travers les domaines HR respectifs orientés de manière antiparallèle. Dans toutes les figures, les domaines sont colorés : violet, HRN ; vert, HR1 ; orange, HR2 ; rouge, GTPase et jaune, transmembrane. Les atomes de phosphore (bleu foncé) des deux membranes mitochondriales externes (m1, m2) sont représentés sous forme de sphères, tandis que le restant des membranes est souligné par une ombre bleue. Les distances intramembranaires entre la couche externe de phosphate des membranes sont indiquées.

dans la littérature. Les modèles proposés dans cette étude ont fait l'objet d'une évaluation poussée de leur stabilité et de leur dynamique dans un environnement membranaire par des simulations de dynamique moléculaire (DM) gros grain. Cette étude ouvre de nouvelles voies pour l'exploration des conformations cis et trans pour Fzo1 et les mitofusines en général.

Par analogie avec les spirales macromoléculaires formées lors de la fission membranaire médiée par les DRP (Jimah et Hinshaw, 2019), la fusion membranaire médiée par les

DRP peut nécessiter la formation de macro-oligomères comme observé avec la structure en anneau Fzo1-dépendante [37]. Dans le contexte des mitofusines, ces éléments constitutifs se formeraient sur la même membrane, imposant l'assemblage de cis-oligomères.

Ce raisonnement nous a amenés à construire trois modèles d'assemblage de Fzo1, qui ont donné naissance à trois modèles de trans-oligomères de mitofusine et trois scénarios distincts d'oligomérisation de Fzo1. Chaque scénario implique deux composantes principales de l'auto-oligomérisation de la mitofusine : l'interface du domaine GTPase [46, 271] et l'interface du domaine tronc [135]. Une grande quantité de données expérimentales démontre que l'intégrité des deux régions est essentielle à la fusion mitochondriale par mitofusine [57, 77, 101, 110, 113, 135, 207, 215]. Si on considère la protéine BDLP (*bacterial dynamin-like protein*), l'interface du domaine GTPase devrait déclencher une oligomérisation de Fzo1 en cis [153]. Néanmoins, il est aussi suggéré que cette interface déclenche la trans-oligomérisation des mitofusines [46, 271].

Un seul de nos trois modèles de trans-tétramères correspond à cette stratégie. Dans les cis-dimères qui interagissent « dos à dos » (avec des répétitions heptades (HR) orientées en parallèle), deux dimères de membranes opposées s'engageraient dans des interactions trans via leurs domaines GTPase. Pourtant, la génération de ce trans-tétramère nécessite une rotation extensive des domaines GTPase autour de l'axe de leurs quatre faisceaux d'hélices afin de reconstituer l'interface G canonique où les domaines se font face en miroir, comme observé pour les autres membres de la superfamille des dynamines [46, 66]. De plus, une grande partie du domaine de liaison Ugo1 dans chaque monomère Fzo1 est entravée par l'interface d'interaction HR nécessaire pour former les cis-dimères dos à dos. Ce modèle présente donc d'importantes limites lorsqu'il est confronté à la littérature expérimentale. Il a été démontré que l'interface G du domaine MFN1 favorise l'ancrage de la membrane, ce qui suggère une fonction de trans-oligomérisation [46, 199]. Cependant, cela a été observé avec des constructions de MFN1 liées artificiellement à des membranes et ne possédant pas le tronc entier et les régions transmembranaires prédits pour les mitofusines. La possibilité que les mitofusines complètes utilisent l'interface G pour l'oligomérisation en cis n'est donc pas à exclure.

Les deux autres modèles utilisent ces propriétés de cis-oligomérisation en tête-à-tête. Dans cette configuration, le Fzo1 lié au GTP subirait une ouverture, étroite ou large. Comme on l'a vu pour la stabilisation des DRP de fission liées au GTP avec leurs cofacteurs [142], ce changement de conformation serait stabilisé par la liaison concomitante d'Ugo1 via le tronc disponible et d'une molécule de mitofusine adjacente via les domaines GTPase. Dans ces deux modèles, « large » et « étroit », l'interface G canonique est maintenue. Par la suite, les cis-oligomères larges ou étroits s'engageraient dans une trans-association avec

les cis-complexes Fzo1 de la membrane opposée. Les oligomères larges de type BDLP s'associeraient en trans par l'intermédiaire d'une interface dans le domaine G distincte de celle employée pendant la dimérisation cis. À notre connaissance, une telle interface n'a jamais été observée pour les autres membres de la superfamille des dynamines, ce qui met sérieusement en doute sa validité. De plus, ces trans-oligomères Fzo1 imposeraient une distance de fixation entre les membranes extérieures d'environ 32 nm, significativement plus grande que la distance observée expérimentalement entre mitochondries amarrées (8 nm, [37]). Par conséquent, le modèle tête-à-tête à large ouverture est difficile à concilier avec la littérature. A contrario, le mécanisme tête-à-tête à ouverture étroite utilise une stratégie de trans-oligomérisation qui s'avère en excellent accord avec la littérature passée et actuelle. Après une éventuelle libération d'Ugo1, la région du tronc dans les oligomères cis exposerait une colonne hydrophobe de sa portion HR2 qui serait disponible pour de possibles interactions [70]. Ces interactions seraient réalisées de façon antiparallèle avec une colonne HR2 équivalente appartenant à des oligomères Fzo1 de la membrane opposée. Cela serait cohérent avec les interactions antiparallèles HR2 observées en 2004 pour MFN1 [135]. De plus, la distance de liaison de 5,5 à 9,5 nm imposée par ce modèle d'oligomérisation Fzo1 est conforme à la séparation d'environ 8 nm entre les membranes extérieures observée expérimentalement douze ans plus tard [37]. La cis-oligomérisation tête-à-tête à ouverture étroite suivie d'une association en trans antiparallèle dos-à-dos permet donc de spéculer sur le devenir du complexe de fusion. Après l'hydrolyse du GTP, les charnières 1a et 1b pourraient rétablir le déplacement du domaine GTPase vers le tronc. Cela contribuerait à réduire la distance entre les membranes extérieures à 2-3 nm, comme observé précédemment [37]. Ce cycle se réitèrerait à proximité de cette région initiale de contact minimal jusqu'à atteindre le stade d'amarrage caractérisé par une zone étendue d'apposition membranaire délimitée par le complexe d'amarrage mitochondrial (MDC) en forme d'anneau. Il est tentant de proposer que ce MDC, qui se trouve dans des régions où la distance entre les membranes extérieures atteint 6 à 8 nm [37], correspond à un ensemble macromoléculaire d'oligomères Fzo1 rappelant ceux décrits dans cette étude. Nos modèles sont suffisamment détaillés pour proposer des hypothèses vérifiables expérimentalement dans de futurs travaux, par exemple en termes de réticulations possibles et de résidus exposés en surface. De même, nos modèles pourraient être encore améliorés sur la base de nouvelles observations expérimentales. Indépendamment de ces possibilités, nous espérons que ces hypothèses donneront matière à réflexion avant d'élucider la structure complète des mitofusines et de leurs cofacteurs, connus ou encore à décrire, ainsi que leurs mécanismes précis d'oligomérisation.

7.3 Partie 2: Ajouter l'hydrodynamique aux lipides

L'étude des membranes biologiques, fondée sur des approches expérimentales, théoriques et informatiques, occupe une place centrale dans la biophysique moderne. Plus que de simples barrières de protection, les membranes forment un environnement très dynamique qui accueille de nombreux processus métaboliques et de signalisation. En outre, des bicouches lipidiques qui imitent les membranes cellulaires sont utilisées par les bioingénieurs comme base pour de nouvelles voies d'administration de médicaments et la construction d'organites artificiels [41, 100]. Ce n'est que récemment que nous avons commencé à saisir toute la complexité de ces environnements lipidiques, donnant naissance au domaine de la lipidomique [226].

De nos jours, une grande variété de techniques expérimentales permet de caractériser de nombreux aspects de la structure et de la dynamique des biomembranes, comme l'extension des domaines membranaires [212], la diffusion lipidique [95], la localisation des protéines [144], la déformation des membranes et la fusion [227]. De plus, les constructions membranaires telles que les vésicules sont étudiées expérimentalement pour leurs capacités de transport et de libération de contenu [195, 248]. Leur comportement dans un contexte d'écoulement a été étudié pour évaluer les déformations [71] et les effets des contraintes induites par les forces mécaniques [132].

La modélisation et la simulation par ordinateur constituent un support solide pour la compréhension des systèmes membranaires. Au cours des 30 dernières années, le développement des champs de forces tout-atome, combiné à la disponibilité d'une puissance informatique massive, a stimulé l'étude de systèmes de taille et de complexité croissantes [155]. Cette évolution a permis de caractériser des processus tels que la perméation membranaire de petits solutés [228], la formation de pores [130] ou la dynamique de tout un complexe respiratoire [222]. Bien que la popularité des simulations atomistiques de systèmes membranaires bénéficie de l'amélioration des algorithmes et de l'efficacité des programmes informatiques, ainsi que de l'utilisation de techniques d'échantillonnage amélioré [118, 179], les échelles de taille et de temps associées à de nombreux phénomènes liés aux membranes nécessitent l'utilisation de modèles simplifiés, dites à "gros-grain" (coarse-grained, CG). Ces modèles reposent sur la réduction substantielle du nombre de degrés de liberté (nombre de particules et d'interactions) pour traiter le calcul. En outre, la fluidité du paysage énergétique CG permet l'utilisation de pas de temps d'intégration significativement plus longs. À cet égard, les modèles implicites de solvants [10, 59, 63, 237, 237] sont particulièrement intéressants, mais, par défaut, ne peuvent pas décrire l'effet des corrélations dynamiques médiées par le solvant (phénomènes hydrodynamiques). Pourtant, ces corrélations jouent non seulement un rôle important dans le processus d'agrégation des lipides [7] et d'autres biomolécules

[54], mais peuvent également influencer de manière critique la migration des biomolécules à proximité des membranes [202] ainsi que la mobilité des vésicules et leurs rencontres [209].

Dans cette partie de la thèse, nous présentons un couplage de Dry Martini [10], un modèle gros-grain de lipides à solvant implicite, avec la technique LBMD (Lattice Boltzmann Molecular Dynamics) [2, 3, 20, 239] afin d'inclure les interactions hydrodynamiques (IH) dans la simulation de membranes en solvant implicite. Ce travail représente une alternative au couplage récemment publié de Dry Martini à la méthode de collision multiparticulaire [274] et prolonge des travaux antérieurs où le modèle en solvant implicite OPEP était couplé aux IH [51, 54, 55, 143, 239, 240]. Cette technique permettra des interactions synergiques avec des expériences où des systèmes lipidiques sont manipulés dans des écoulements fluides [11, 168]. En effet, le caractère microscopique de Dry Martini et l'efficacité ainsi que la nature multi-échelle de LBMD pour traiter le fluide font de leur combinaison une méthode attrayante. Nous avons d'abord validé l'implémentation du champ de force dans le logiciel de simulation multi-échelle MUPHY [26]. Ensuite, nous avons exploré différentes situations où le couplage avec le solvant est essentiel pour étudier le comportement d'un système lipidique. Par exemple, nous avons examiné l'effet de taille dans l'estimation de la diffusion lipidique. À cet égard, la technique LBMD basée sur Dry Martini pourrait être intéressante du point de vue informatique par rapport aux modèles explicites de solvants afin d'approcher la limite infinie nécessaire pour calculer les propriétés de transport des lipides ou protéines dans les systèmes membranaires bidimensionnels [45, 261, 262]. Puisque la technique permet de simuler une grande variété de conditions d'écoulement de fluide, nous avons utilisé cette méthodologie pour explorer le comportement d'une vésicule soumise à un écoulement cisaillant, reproduisant ainsi des travaux expérimentaux [25, 71, 265].

Enfin, inspirés par des travaux expérimentaux récents [138], nous avons exploré la possibilité de simuler un nanoréacteur afin d'étudier l'effet combiné du confinement, de l'encombrement (*crowding*) et de l'hydrodynamique sur les rencontres substrat/enzyme. Le couplage présenté peut être encore étendu. Toutes nos simulations ont été réalisées dans l'ensemble canonique à volume et température fixes. Cependant, pour un système membranaire, il est bien connu qu'il est important de contrôler la pression latérale pour reproduire correctement le comportement structurel et dynamique, en particulier pour les systèmes latéralement périodiques [192, 250]. Dans notre schéma, il est possible d'implémenter un algorithme efficace pour la dynamique des particules afin d'échantillonner l'ensemble isobare. La cellule de simulation de taille variable serait associée à un nombre fixe de points sur le réseau utilisé pour la résolution de la dynamique des fluides.

Un deuxième aspect à mentionner concerne l'application de maillages multirésolution pour augmenter l'efficacité de la LBMD. Notre schéma est adapté à l'étude des propriétés de

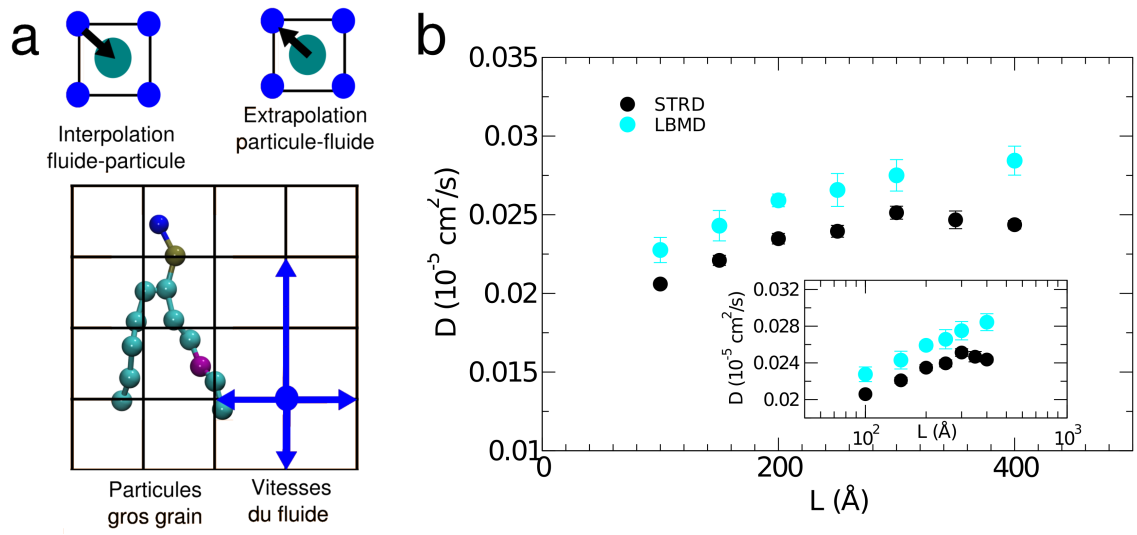

Fig. 7.2 a:Représentation schématique du couplage entre le système moléculaire (CG, MD) et le fluide (Lattice Boltzmann). Les flèches bleues représentent les populations de vitesses du fluide évoluant sur le réseau. Nous montrons également une description des stratégies d'interpolation et d'extrapolation utilisées pour communiquer les forces entre la dynamique des particules et celle du fluide. *Modifié de [239, 246]* b: Effet de la taille de la boîte sur la diffusion lipidique. Coefficient de diffusion du lipide POPC extrait d'une simulation de 800 ns d'une membrane plane dans différentes boîtes de taille $L_x = L_y = L$ et $L_z = 240$ Å(cercle cyan). Le coefficient de diffusion D a été obtenu à partir d'un ajustement linéaire du MSD calculé pendant les 800 dernières ns de la simulation et en utilisant une fenêtre temporelle de 40 à 70 ns pour l'ajustement. Dans la figure (cercles noirs), nous présentons également les données de la Réf. [274] basées sur des simulations STRD de la membrane planaire lipidique POPC avec une séparation inter-membrane de 200 Å. Encart : Diagramme log-linéaire des coefficients de diffusion.

transport dans des systèmes membranaires bidimensionnels dans des conditions de limites périodiques où, pour approcher la limite infinie, la dimension transversale doit être énormément étendue. Le coût de calcul augmente linéairement avec la dimension transversale avec un seul espacement de grille, mais un gain énorme peut être réalisé en utilisant une approche multirésolution, où seulement une partie de l'espace entourant la membrane est résolue à l'échelle fine tandis que le reste du volume est traité par une grille plus grossière. De plus, le calcul de la dynamique des particules peut être amélioré par la mise en œuvre de nouveaux algorithmes pour exploiter la capacité vectorielle des CPU modernes [1] et des approches de parallélisme pour traiter plus efficacement la nature éparse de certains systèmes lipidiques comme les grandes vésicules.

En conclusion, notre approche nous permet maintenant d'explorer une grande variété de processus biophysiques impliquant des systèmes lipidiques à l'échelle mésoscopique où les interactions médiées par solvant sont importantes. Notre approche peut complémenter la bio-ingénierie et l'investigation expérimentale en biotechnologie des systèmes lipidiques tels que les vésicules dans des conditions d'écoulement de fluide [11, 168].

7.4 Partie 3: Couplage lipide-protéine

Après avoir mis en œuvre et validé l'approche multi-échelles pour l'étude des systèmes lipidiques, nous nous sommes intéressés à l'inclusion des protéines. Préalablement à ces travaux, une approche LBMD était déjà disponible pour la simulation multi-échelle de protéines. Nous avons donc décidé d'étendre ces travaux antérieurs réalisés au laboratoire, où le modèle gros grain de protéines OPEP à solvant implicite était couplé aux interactions hydrodynamiques [51, 54, 55, 143, 239], afin d'inclure les lipides. Pour ce faire, il est nécessaire de concevoir un couplage compatible entre lipides et protéines. Ainsi, dans cette partie de la thèse, nous avons travaillé sur une première tentative systématique de couplage d'un champ de force gros grain pour les lipides (Dry Martini) avec le champ de force OPEP pour peptides.

Le champ de force gros grain OPEP pour protéines a été conçu afin d'étudier le repliement des protéines et possède une description très détaillée du squelette peptidique, tout en ne conservant qu'un seul pseudo-atome pour la chaîne latérale de chaque acide aminé (sauf la glycine). Il serait ainsi particulièrement intéressant d'étudier les changements conformationnels et le repliement de peptides ou de protéines dans un environnement membranaire, puisque le champ de forces OPEP est en mesure de traiter ces changements. D'un point de vue plus technique, l'idée de coupler le champ de force Dry Martini pour les lipides avec le champ de force OPEP pour les protéines est motivée par deux caractéristiques principales :

(1) les deux sont des modèles en solvant implicite et (2) le champ de force lipidique de Dry Martini a une résolution comparable à celle utilisée dans OPEP pour modéliser les chaînes latérales d'acides aminés.

Les champs de force de Dry Martini et OPEP ayant été paramétrés avec des stratégies différentes, nous avons d'abord comparé l'échelle d'énergie entre ces modèles, en examinant les interactions de type Lennard-Jones (LJ) présentes dans un lipide typique (POPC) et en les comparant aux pseudo-atomes analogues dans le champ de force OPEP. En particulier, le potentiel énergétique non lié décrivant l'interaction entre deux pseudo-atomes hydrophobes lipidiques (C1-C1) a une énergie minimale beaucoup plus élevée que son analogue protéique. Pour éviter la possible rupture de la membrane liée à la présence d'une protéine, nous avons décidé d'utiliser l'échelle énergétique de Dry Martini comme référence pour l'interaction croisée entre la protéine OPEP et les lipides Dry Martini [10]. Un autre point clé que nous avons dû aborder concerne le cas particulier de certaines chaînes latérales d'acides aminés qui sont représentées par plus d'un pseudo-atome dans Dry Martini, alors qu'elles en ont un seul dans OPEP. Pour ces acides aminés, nous avons dérivé un nouveau potentiel effectif (de pseudo-atome unique) pour décrire les interactions entre la chaîne latérale des acides aminés et les pseudo-atomes lipidiques. Pour cette optimisation du champ de force Dry Martini, nous avons opté pour une approche thermodynamique cohérente avec celle utilisée à l'origine dans le développement du champ de force de protéines Martini. Tout d'abord, pour chacune des chaînes latérales d'acides aminés à multiples pseudo-atomes, nous avons calculé le potentiel de force moyenne (PMF) par *umbrella sampling* pour extraire une chaîne latérale d'acide aminé du centre d'une membrane POPC dans la direction orthogonale à la surface de la bicouche et en utilisant le champ de force Dry Martini. Cela nous a fourni une référence, PMF^{ref}, pour ajuster le nouveau jeu de paramètres. Afin de générer le PMF pour les chaînes latérales de simple pseudo-atomes, une première estimation des interactions croisées avec les lipides était nécessaire. Pour ce faire, nous avons conçu deux stratégies : (1) Nous avons identifié l'interaction principale des pseudo-atomes qui contrôle la forme du PMF^{ref}, et (2) nous avons calculé un potentiel effectif entre la chaîne latérale et chacun des pseudo-atomes lipidiques pour obtenir un ensemble de paramètres de LJ. Le protocole suivi pour chaque acide aminé paramétré peut être résumé avec l'algorithme suivant :

Obtention des paramètres LJ initiaux pour le couplage protéine-lipide, avec la stratégie 1 ou 2. Calcul du PMF de l'extraction des acides aminés de la bicouche avec les paramètres LJ initiaux et comparaison avec le PMF de référence PMF^{ref} obtenu en utilisant le champ de force Dry Martini. Nouveau calcul du PMF avec différentes valeurs d'epsilon pour le LJ (valeur initiale d'epsilon multipliée par un facteur). Cette étape a été répétée jusqu'à ce que le PMF corresponde au profil de référence, en particulier aux valeurs z les plus élevées.

Contrôle de la fonction de distribution radiale (RDF) du pseudo-atome de la chaîne latérale et des pseudo-atomes du noyau hydrophobe de la membrane pour s'assurer que l'emballage local dans la membrane est préservé. Fixation d'une nouvelle valeur de sigma si nécessaire. Nouveau calcul du PMF avec la nouvelle valeur sigma. Si nécessaire, cette étape a été répétée avec sigma fixé et changement d'epsilon jusqu'à ce que le PMF corresponde au profil de référence.

Voir figure 7.3 comme exemple pour le processus complet dans le cas de la tyrosine.

Le choix des paramètres a été fait de manière à ce que les PMF résultants correspondent aux PMF de référence générés avec les paramètres standards de Dry Martini (version 2.1) dans Gromacs. De plus, les fonctions de distribution radiale entre les pseudo-atomes de la chaîne latérale des acides aminés et les différents pseudo-atomes lipidiques ont été vérifiées pour assurer une distribution géométrique correcte de l'acide aminé dans la bicouche. Les fonctions de distribution radiale entre la chaîne latérale et les pseudo-atomes qui forment la queue lipidique hydrophobe (C1, C3) montrent une excellente concordance, en particulier pour la position du premier pic.

Enfin, ces paramètres ainsi que ceux des acides aminés dont les topologies comportent un seul pseudo-atome de chaîne latérale ont été utilisés pour réaliser des simulations avec OPEP/Dry Martini dans MUPHY, le moteur de simulation où OPEP est déjà implémenté. Nous avons testé les paramètres sur des hélices transmembranaires bien connues pour être stables et intégrées à la membrane. Nous avons réalisé des simulations de 3 systèmes différents : polyleucine (L10), WALP23 et KALP23. Les différents systèmes ont été étudiés avec OPEP/Dry Martini avec MUPHY et avec les modèles CHARMM36m (tout-atome), Dry Martini ou Martini (Wet Martini) avec Gromacs. Leur comportement a été évalué en tenant compte de caractéristiques pertinentes : la distance bout à bout du peptide, la structure secondaire (pour les simulations tout-atome et OPEP/Dry Martini), le positionnement dans la membrane et l'angle d'inclinaison.

Dans tous les systèmes, les simulations OPEP/Dry Martini maintiennent une structure secondaire stable et la distance bout à bout des peptides. Ils ont le même comportement qualitatif que la simulation tout atome de référence. Fait intéressant, le peptide L10 peut se déplacer librement dans la membrane et explorer différentes orientations. Cependant, l'angle d'inclinaison des peptides flanqués WALP23 et KALP23 s'est stabilisé pendant la simulation et dans le cas de KALP23 il se stabilise autour de 10°, montrant la même tendance que la simulation atomistique et comparablement aux valeurs précédemment rapportées ($7° - 17°$ [98]). D'autre part, l'angle d'inclinaison du peptide WALP23 obtenu dans les simulations OPEP/Dry Martini, bien que stable, s'est montré très différent de l'angle d'inclinaison obtenu dans les simulations Martini et tout-atome, ainsi que des angles rapportés expérimentalement.

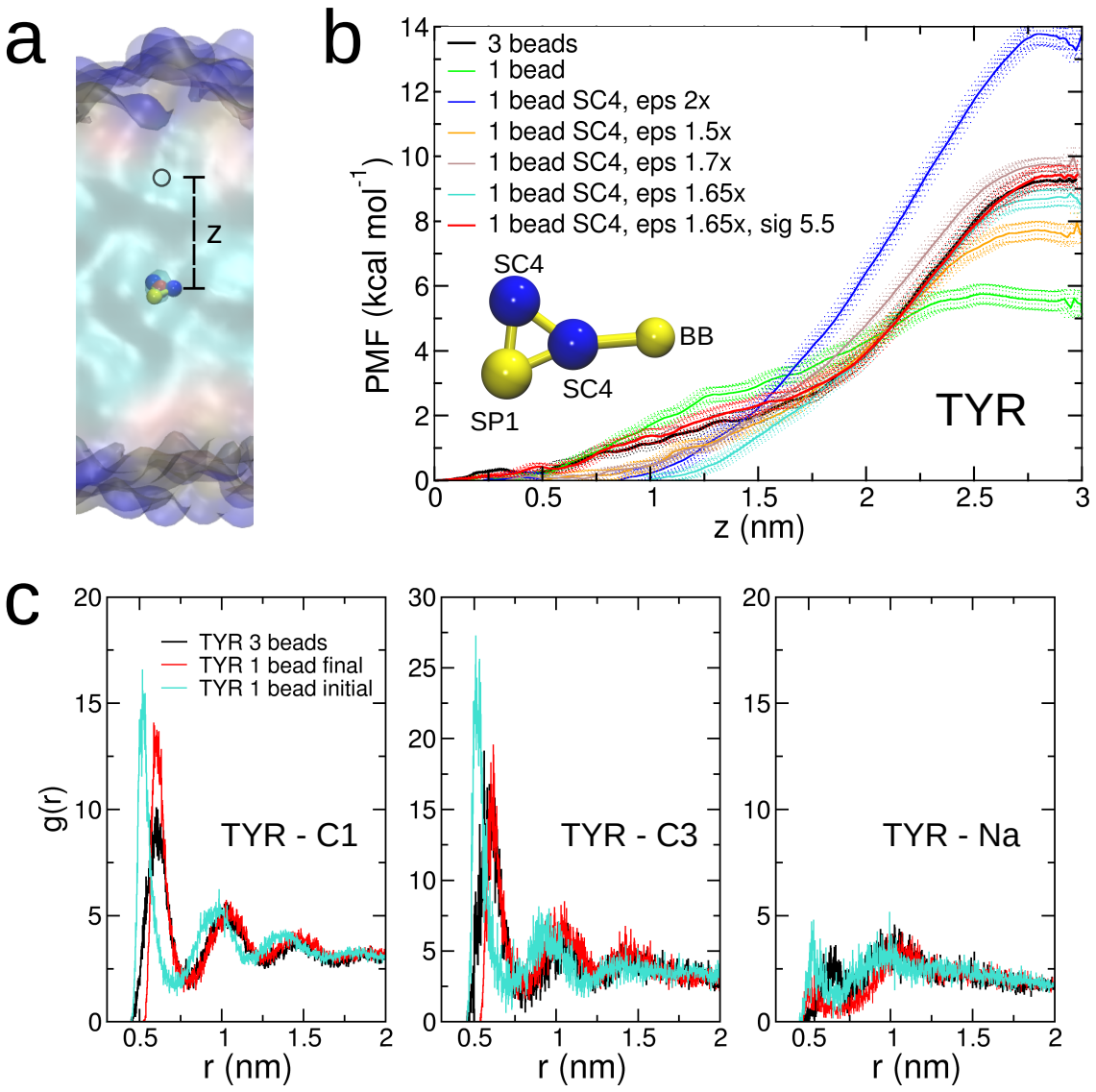

Fig. 7.3 Exemple de l'ensemble du processus de paramétrisation dans le cas de la tyrosine (TYR) comme un acide aminé représentatif. a : Configuration pour les calculs de PMF montrant les coordonnées initiales de la fenêtre correspondant à z = 0 nm (zoom sur la région entourant la chaîne latérale des acides aminés). b : Exemple de différents PMF obtenus à partir de la simulation de référence et de différentes suppositions LJ soit avec la stratégie 1 (un pseudo-atome SC4) ou la stratégie 2 (un pseudo-atome deviné). Pour la première stratégie, plusieurs tests sont présentés avec différentes valeurs ε (eps) et σ (sig). c : Fonctions de distribution radiale entre la chaîne latérale et les pseudo-atomes hydrophobes présents dans le lipide, pour le système de référence (noir), meilleur PMF après ajustement lors de la modification des valeurs ε seulement (turquoise), et résultat final (rouge)

Il montre néanmoins la même tendance que Dry Martini, les deux étant éloignés de 50° des valeurs rapportées dans la littérature. Cela pourrait indiquer la nécessité d'améliorer les paramètres de l'interaction croisée lipide-protéine dans OPEP/Dry Martini, mais aussi la nécessité d'améliorer les paramètres du tryptophane dans Dry Martini lui-même.

Dans cette partie, nous avons décrit la stratégie et guidé le lecteur à travers la recherche et la mise en œuvre d'un ensemble de paramètres pour coupler le champ de force lipidique Dry Martini au champ de force peptidique OPEP. Nous avons utilisé l'échelle d'énergie Dry Martini comme base pour le couplage. Après avoir obtenu un premier ensemble de paramètres croisés, nous les avons mis à l'essai pour une série de peptides transmembranaires connus, dans le but de comparer les tendances et les comportements qualitatifs des propriétés des peptides lorsqu'ils sont intégrés dans une membrane. Nous avons montré qu'il est possible d'effectuer des simulations de peptides transmembranaires sans les altérer (par exemple, ni la longueur du peptide ni la structure secondaire n'ont été modifiées de façon significative). Malheureusement, dans le cas particulier du peptide WALP23, nous n'avons pas pu reproduire l'angle d'inclinaison prévu et l'orientation transmembranaire relative avec cette version des paramètres. Un nouvel ensemble de paramètres de Lennard-Jones pour le tryptophane devrait probablement être dérivé. D'autre part, les résultats de la simulation avec KALP23 ont été très prometteurs, et nous souhaitons étendre ces travaux à l'étude de la dimérisation de ce monomère dans la bicouche. En guise de conclusion générale, étant donné que les paramètres de Dry Martini pour la protéine sont encore en cours de développement, il serait sage d'affiner cet ensemble de paramètres dans un deuxième temps en utilisant les PMF atomistiques comme référence plus fiable afin de générer un ensemble universel de paramètres pour l'interaction croisée.

7.5 Conclusion finale et perspectives

Dans les travaux présentés dans cette thèse de doctorat, nous avons couvert plusieurs aspects qui sont d'intérêt pour l'étude des protéines membranaires. Tout d'abord, nous avons exploré les premières étapes de l'ancrage membranaire dans le contexte de la fusion mitochondriale chez la levure par des modèles gros grains déjà disponibles (chapitre 3). Nous avons construit trois modèles indépendants de dimères en cis de Fzo1 qui ont donné naissance à trois modèles indépendants de trans-oligomères de mitofusine. Ensuite, ces modèles ont donné trois scénarios distincts d'oligomérisation de Fzo1. Chaque scénario implique deux composantes principales de l'auto-oligomérisation de la mitofusine : l'interface du domaine GTPase [46, 271] et l'interface du domaine « tronc »[135]. Nous avons montré que le mécanisme d'ouverture du complexe tête-à-tête « étroit » pourrait employer une stratégie de

trans-oligomérisation qui s'avère conforme à la littérature passée et actuelle. L'interaction antiparallèle en trans de HR2 proposée dans le complexe Fzo1 est en accord avec une interaction antiparallèle similaire observée pour les répétitions heptades dans MFN1 [135]. En outre, les distances entre les membranes extérieures atteindraient 6 à 8 nm [37], ce qui correspond à un assemblage macromoléculaire cohérent avec ceux décrits dans la présente étude. De plus, nos modèles fournissent suffisamment de détails pour proposer des hypothèses vérifiables expérimentalement, par *cross-linking* et analyse des résidus exposés en surface. Ces modèles pourraient être encore améliorés sur la base de nouvelles observations expérimentales. Plus généralement, en mettant ces modèles (dimères et tétramètres) à la disposition de la communauté scientifique, nous proposons des hypothèses structurelles réalistes sur les mitofusines entières et leurs cofacteurs, ainsi que sur leurs propriétés d'oligomérisation.
Deuxièmement, nous avons mis en place un nouveau couplage entre un modèle gros grain à solvant implicite pour les lipides et l'hydrodynamique (chapitre 4). Cette implémentation a également été présentée avec un ensemble d'applications, qui peuvent être utiles pour résoudre d'importantes questions biologiques.
Dans ce travail, nous avons étendu la capacité opérationnelle de la technique multi-échelle de la dynamique moléculaire de Boltzmann sur réseau (Lattice Boltzmann Molecular Dynamics) développée dans le laboratoire [54, 55, 143, 239]. Grâce à cette approche, nous sommes maintenant en mesure d'explorer une grande variété de processus biophysiques impliquant des systèmes lipidiques à l'échelle mésoscopique, où les interactions portées par le solvant sont importantes. Plus intéressant encore, notre technique peut explorer des problèmes de bio-ingénierie et complémenter les études expérimentales en biotechnologie sur des systèmes lipidiques comme des vésicules dans des écoulements fluides [11, 168]. La perspective d'étudier des systèmes lipidiques comportant également des protéines a conduit au développement d'un couplage effectif rigoureux entre le champ de force de Dry Martini pour les lipides et le modèle gros grain pour protéines OPEP, qui est aussi un champ de force avec solvant implicite (chapitre 5). Pour concevoir ce couplage, il a fallu adopter des stratégies pour traiter les interactions croisées des lipides et des protéines, et en particulier un nouveau schéma de résolution a dû être mis en place.
Au stade actuel du projet, il est possible de simuler des protéines avec le champ de force OPEP dans un environnement membranaire. Cependant, le couplage OPEP/Dry Martini est dans sa première version et doit encore être affiné. Davantage de simulations sont en cours de collecte pour les différents systèmes peptides/membranes afin de valider le couplage avec de meilleures statistiques. Depuis que cette implémentation a été faite pour le moteur de simulation MUPHY, et que les deux modèles de protéines et de lipides ont été testés dans des simulations avec interactions hydrodynamiques, il est facile de les combiner au

sein d'un système membrane/protéine. Nous sommes conscients qu'il existe déjà de nombreux champs de force à gros grain qui permettent de simuler les protéines et les lipides ensemble à une résolution quasi-atomistique [15, 22, 170]. Cependant, aucun d'entre eux ne dispose de modèle protéique avec une description détaillée du squelette peptidique comme c'est le cas d'OPEP. Nous pensons que ce modèle pourrait être très utile pour étudier les interactions peptide-membrane, car elles impliquent généralement des changements dans le repliement du peptide. De plus, comme le comportement cinétique du solvant peut également être inclus explicitement via des interactions hydrodynamiques, il pourrait être d'intérêt biotechnologique d'étudier, par exemple, le cas de l'agrégation peptidique en présence ou non d'un écoulement du fluide environnant. En résumé, à l'étape finale de ce projet, nous disposons d'un modèle lipidique qui peut être utilisé pour des simulations dans MUPHY avec ou sans hydrodynamique. Ce modèle a fait l'objet de nombreux essais et il est opérationnel. D'autre part, le jeu de paramètres obtenu pour le couplage protéine-lipide en est à sa première version (v.0), il faudra donc l'affiner encore pour obtenir une meilleure concordance avec les résultats expérimentaux, en particulier dans le cas du tryptophane. En revanche, l'étude de la fusion mitochondriale avec des modèles gros grain "standard" a donné un premier aperçu de l'organisation possible du complexe de fixation dans des conditions très simples (Fzo1 seul et membrane) mais suffisamment pertinentes pour proposer une description structurelle à la communauté travaillant sur la dynamique mitochondriale. Le plan à court terme est de continuer avec une représentation plus détaillée du système choisi (le tétramère «trans-antiparallèle, textitcis tête-à-tête parallèle») ainsi que la modélisation des répétitions heptades seules et leur oligomérisation en fragments dans un contexte membranaire (travaux en cours par nos collaborateurs expérimentateurs). Dans l'ensemble, d'un point de vue méthodologique, nous avons présenté une nouvelle approche pour effectuer des simulations multi-échelles de systèmes lipidiques à gros grains avec une représentation cinétique hydrodynamique explicite au moyen de Lattice Boltzmann. Nous avons également mis au point une première version du couplage des champs de force OPEP et Dry Martini, permettant d'étendre les systèmes qui peuvent être utilisés avec OPEP et d'inclure les lipides. Enfin, nous avons contribué à caractériser les premières étapes de la fusion mitochondriale et à démystifier les modèles théoriques existant qui ont été largement répandus dans la littérature sans aucune considération structurelle détaillée.

References

[1] Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25.

[2] Ahlrichs, P. and Dünweg, B. (1998). Lattice-Boltzmann Simulation of Polymer-Solvent Systems. *Int. J. Mod. Phy. C*, 9(8):1429–1438.

[3] Ahlrichs, P. and Dünweg, B. (1999). Simulation of a single polymer chain in solution by combining lattice Boltzmann and molecular dynamics. *J. Chem. Phys.*, 111(17):8225–8239.

[4] Aldridge, A. C., Benson, L. P., Siegenthaler, M. M., Whigham, B. T., Stowers, R. S., and Hales, K. G. (2007). Roles for drp1, a dynamin-related protein, and milton, a kinesin-associated protein, in mitochondrial segregation, unfurling and elongation during drosophila spermatogenesis. *Fly (Austin)*, 1(1):38–46.

[5] Altmann, K., Frank, M., Neumann, D., Jakobs, S., and Westermann, B. (2008). The class V myosin motor protein, Myo2, plays a major role in mitochondrial motility in Saccharomyces cerevisiae. *J. Cell Biol.*, 181(1):119–130.

[6] Ando, T. and Skolnick, J. (2010). Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion. *Proc. Natl. Acad. Sci. {USA}*, 107:18457–18462.

[7] Ando, T. and Skolnick, J. (2013). On the Importance of Hydrodynamic Interactions in Lipid Membrane Formation. *Biophys. J.*, 104(1):96–105.

[8] Anton, F., Fres, J. M., Schauss, A., Pinson, B., Praefcke, G. J. K., Langer, T., and Escobar-Henriques, M. (2011). Ugo1 and Mdm30 act sequentially during Fzo1-mediated mitochondrial outer membrane fusion. *J. Cell. Sci.*, 124(Pt 7):1126–1135.

[9] Arnarez, C., Marrink, S. J., Periole, X., Arnarez, X., Marrink, S. J., Periole, X., Arnarez, C., Marrink, S. J., and Periole, X. (2016). Molecular mechanism of cardiolipin-mediated assembly of respiratory chain supercomplexes. *Chem. Sci.*, 7(7):4435–4443.

[10] Arnarez, C., Uusitalo, J. J., Masman, M. F., Ingólfsson, H. I., De Jong, D. H., Melo, M. N., Periole, X., De Vries, A. H., and Marrink, S. J. (2015). Dry martini, a coarse-grained force field for lipid membrane simulations with implicit solvent. *Journal of Chemical Theory and Computation*, 11(1):260–275.

[11] Aydin, F., Chi, X., Greenstein, J., and Dutt, M. (2018). A review on phospholips vesicles flowing through channels. *MRS Communications*, 8:718–726.

[12] Balaban, R. S., Nemoto, S., and Finkel, T. (2005). Mitochondria, oxidants, and aging. *Cell*, 120(4):483–495.

[13] Ban, T., Ishihara, T., Kohno, H., Saita, S., Ichimura, A., Maenaka, K., Oka, T., Mihara, K., and Ishihara, N. (2017). Molecular basis of selective mitochondrial fusion by heterotypic action between OPA1 and cardiolipin. *Nat. Cell Biol.*, 19(7):856–863.

[14] Baoukina, S. and Tieleman, D. P. (2010). Direct simulation of protein-mediated vesicle fusion: lung surfactant protein B. *Biophysical journal*, 99(7):2134–2142.

[15] Barrera, E. E., Machado, M. R., and Pantano, S. (2019). Fat SIRAH: Coarse-Grained Phospholipids To Explore Membrane–Protein Dynamics. *Journal of Chemical Theory and Computation*, 15(10):5674–5688.

[16] Bauer, M. and Pelkmans, L. (2006). A new paradigm for membrane-organizing and -shaping scaffolds. *FEBS Letters*, 580(23):5559–5564.

[17] Baumann, P., Spulber, M., Fischer, O., Car, A., and Meier, W. (2017). Investigation of Horseradish Peroxidase Kinetics in an "Organelle-Like" Environment. *Small*, 13(17):10–12.

[18] Bausewein, T., Mills, D. J., Langer, J. D., Nitschke, B., Nussberger, S., and Kühlbrandt, W. (2017). Cryo-EM Structure of the TOM Core Complex from Neurospora crassa. *Cell*, 170(4):693 – 700.e7.

[19] Bennun, S. V., Hoopes, M. I., Xing, C., and Faller, R. (2009). Coarse-grained modeling of lipids. *Chemistry and Physics of Lipids*, 159(2):59–66.

[20] Benzi, R., Succi, S., and Vergassola, M. (1992). The lattice Boltzmann equation: theory and applications. *Physics Reports*, 222(3):145–197.

[21] Bereau, T. and Kremer, K. (2015). Automated Parametrization of the Coarse-Grained Martini Force Field for Small Organic Molecules. *Journal of Chemical Theory and Computation*, 11(6):2783–2791.

[22] Bereau, T., Wang, Z. J., and Deserno, M. (2014). More than the sum of its parts: Coarse-grained peptide-lipid interactions from a simple cross-parametrization. *Journal of Chemical Physics*, 140(11).

[23] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690.

[24] Berman, H., Henrick, K., and Nakamura, H. (2003). Announcing the worldwide Protein Data Bank. *Nat. Struct. Bio.*, 10(12):980.

[25] Bernard, A.-L., Guedeau-Boudeville, M.-A., Marchi-Artzner, V., Gulik-Krzywicki, T., di Meglio, J.-M., and Jullien, L. (2005). Shear-induced permeation and fusion of lipid vesicles. *J. Coll. Int. Sci.*, 287:298–306.

[26] Bernaschi, M., Melchionna, S., Succi, S., Fyta, M., Kaxiras, E., and Sircar, J. K. (2009). MUPHY: A parallel MUlti PHYsics/scale code for high performance bio-fluidic simulations. *Comput. Phys. Comm.*, 180:1495–1502.

[27] Bertholet, A. M., Delerue, T., Millet, A. M., Moulis, M. F., David, C., Daloyau, M., Arnauné-Pelloquin, L., Davezac, N., Mils, V., Miquel, M. C., Rojo, M., and Belenguer, P. (2016). Mitochondrial fusion/fission dynamics in neurodegeneration and neuronal plasticity. *Neurobiol. Dis.*, 90:3–19.

[28] Bettolo-Marconi, U. M. and Melchionna, S. (2012). Charge Transport in Nanochannels: A Molecular Theory. *Langmuir*, 28:13727–13740.

[29] Bhaskara, R. M., Linker, S. M., Vögele, M., Köfinger, J., and Hummer, G. (2017). Carbon Nanotubes Mediate Fusion of Lipid Vesicles. *ACS Nano*, 11(2):1273–1280.

[30] Bian, X., Klemm, R. W., Liu, T. Y., Zhang, M., Sun, S., Sui, X., Liu, X., Rapoport, T. A., and Hu, J. (2011). Structures of the atlastin GTPase provide insight into homotypic fusion of endoplasmic reticulum membranes. *Proc. Natl. Acad. Sci. U.S.A.*, 108(10):3976–3981.

[31] Bidone, T. C., Polley, A., Jin, J., Driscoll, T., Iwamoto, D. V., Calderwood, D. A., Schwartz, M. A., and Voth, G. A. (2019). Coarse-Grained Simulation of Full-Length Integrin Activation. *Biophysical Journal*, 116(6):1000–1010.

[32] Botu, V., Batra, R., Chapman, J., and Ramprasad, R. (2017). Machine Learning Force Fields: Construction, Validation, and Outlook. *The Journal of Physical Chemistry C*, 121(1):511–522.

[33] Bourne, H. R., Sanders, D. A., and McCormick, F. (1991). The GTPase superfamily: conserved structure and molecular mechanism. *Nature*, 349(6305):117–127.

[34] Brandner, A., De Vecchis, D., Baaden, M., Cohen, M. M., and Taly, A. (2019a). Physics-based oligomeric models of the yeast mitofusin Fzo1 at the molecular scale in the context of membrane docking. *Mitochondrion*.

[35] Brandner, A., De Vecchis, D., Baaden, M., Cohen, M. M., and Taly, A. (2019b). Structural dataset from microsecond-long simulations of yeast mitofusin Fzo1 in the context of membrane docking. *Data in Brief*, page 104460.

[36] Brandner, A., Timr, S., Melchionna, S., Derreumaux, P., Baaden, M., and Sterpone, F. (2019c). Modelling lipid systems in fluid with Lattice Boltzmann Molecular Dynamics simulations and hydrodynamics. *Scientific Reports*, (Under Revission).

[37] Brandt, T., Cavellini, L., Kühlbrandt, W., and Cohen, M. M. (2016). A mitofusin-dependent docking ring complex triggers mitochondrial fusion in vitro. *Elife*, 5.

[38] Braun, R. J. and Westermann, B. (2011). Mitochondrial dynamics in yeast cell death and aging. *Biochemical Society Transactions*, 39(5):1520–1526.

[39] Brouillette, C. G., Segrest, J. P., Ng, T. C., and Jones, J. L. (1982). Minimal size phosphatidylcholine vesicles: effects of radius of curvature on head group packing and conformation. *Biochemistry*, 21:194569–4575.

[40] Bujacz, A. (2012). Structures of bovine, equine and leporine serum albumin. *Acta Crystallogr. Sect. D: Struct. Bio.*, 68(10):1278–1289.

[41] Bunker, A., Magarkar, A., and Viitala, T. (2016). Rational design of liposomal drug delivery systems, a review: Combined experimental and computational studies of lipid membranes, liposomes and their PEGylation. *Biochim. Biophys. Acta. Biomembrane*, 1858(10):2334–2352.

[42] Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):14101.

[43] Bustillo-Zabalbeitia, I., Montessuit, S., Raemy, E., Basañez, G., Terrones, O., and Martinou, J.-C. (2014). Specific interaction with cardiolipin triggers functional activation of Dynamin-Related Protein 1. *PloS one*, 9(7):e102738.

[44] Byrnes, L. J., Singh, A., Szeto, K., Benvin, N. M., O'Donnell, J. P., Zipfel, W. R., and Sondermann, H. (2013). Structural basis for conformational switching and GTP loading of the large G protein atlastin. *EMBO J.*, 32(3):369–384.

[45] Camley, B. A., Lerner, M. G., Pastor, R. W., and Brown, F. L. (2015). Strong influence of periodic boundary conditions on lateral diffusion in lipid bilayer membranes. *Journal of Chemical Physics*, 143(24):1–12.

[46] Cao, Y.-L. L., Meng, S., Chen, Y., Feng, J.-X. X., Gu, D.-D. D., Yu, B., Li, Y.-J. J., Yang, J.-Y. Y., Liao, S., Chan, D. C., and Gao, S. (2017). MFN1 structures reveal nucleotide-triggered dimerization critical for mitochondrial fusion. *Nature*, 542(7641):372–376.

[47] Cao, Z. and Voth, G. A. (2015). The multiscale coarse-graining method. XI. Accurate interactions based on the centers of charge of coarse-grained sites. *Journal of Chemical Physics*, 143(24).

[48] Chan, D. C. (2012). Fusion and fission: interlinked processes critical for mitochondrial health. *Annu. Rev. Genet.*, 46:265–287.

[49] Chappie, J. S., Acharya, S., Leonard, M., Schmid, S. L., and Dyda, F. (2010). G domain dimerization controls dynamin's assembly-stimulated GTPase activity. *Nature*, 465(7297):435–440.

[50] Chavent, M., Reddy, T., Goose, J., Dahl, A. C. E., Stone, J. E., Jobard, B., and Sansom, M. S. (2014). Methodologies for the analysis of instantaneous lipid diffusion in md simulations of large membrane systems. *Faraday Discussions*, 169:455–475.

[51] Chebaro, Y., Pasquali, S., and Derreumaux, P. (2012). The coarse-grained OPEP force field for non-amyloid and amyloid proteins. *Journal of Physical Chemistry B*, 116(30):8741–8752.

[52] Chen, H., Chomyn, A., and Chan, D. C. (2005). Disruption of fusion results in mitochondrial heterogeneity and dysfunction. *J. Biol. Chem.*, 280(28):26185–26192.

[53] Chen, Q., Schönherr, H., and Vancso, G. J. (2009). Block-copolymer vesicles as nanoreactors for enzymatic reactions. *Small*, 5(12):1436–1445.

[54] Chiricotto, M., Melchionna, S., Derreumaux, P., and Sterpone, F. (2016a). Hydrodynamic Effects on b-Amyloid (16-22) Peptide Aggregation. *J. Chem. Phys.*, 145:35102.

[55] Chiricotto, M., Sterpone, F., Derreumaux, P., and Melchionna, S. (2016b). Multiscale Simulation of Molecular Processes in Cellular Environments. *Philos. Trans. A*, 374:20160225.

[56] Chu, X., Yu, X., Greenstein, J., Aydin, F., Uppaladadium, G., and Dutt, M. (2017). Flow-Induced Shape Reconfiguration, Phase Separation, and Rupture of Bio-Inspired Vesicles. *ACS Nano*, 11:6661-6671.

[57] Cohen, M. M., Amiott, E. A., Day, A. R., Leboucher, G. P., Pryce, E. N., Glickman, M. H., McCaffery, J. M., Shaw, J. M., and Weissman, A. M. (2011). Sequential requirements for the GTPase domain of the mitofusin Fzo1 and the ubiquitin ligase SCFMdm30 in mitochondrial outer membrane fusion. *J. Cell. Sci.*, 124(Pt 9):1403–1410.

[58] Cohen, M. M. and Tareste, D. (2018). Recent insights into the structure and function of Mitofusins in mitochondrial fusion. *F1000Res*, 7.

[59] Cooke, I. R. and Deserno, M. (2005). Solvent-free model for self-assembling fluid bilayer membranes: Stabilization of the fluid phase based on broad attractive tail potentials. *Journal of Chemical Physics*, 123(22):1–14.

[60] Coonrod, E. M., Karren, M. A., and Shaw, J. M. (2007). Ugo1p is a multipass transmembrane protein with a single carrier domain required for mitochondrial fusion. *Traffic (Copenhagen, Denmark)*, 8(5):500–511.

[61] Corradi, V., Mendez-Villuendas, E., Ingólfsson, H. I., Gu, R.-X., Siuda, I., Melo, M. N., Moussatova, A., DeGagné, L. J., Sejdiu, B. I., Singh, G., Wassenaar, T. A., Delgado Magnero, K., Marrink, S. J., and Tieleman, D. P. (2018). Lipid-Protein Interactions Are Unique Fingerprints for Membrane Proteins. *ACS central science*, 4(6):709–717.

[62] Corradi, V., Sejdiu, B. I., Mesa-Galloso, H., Abdizadeh, H., Noskov, S. Y., Marrink, S. J., and Tieleman, D. P. (2019). Emerging Diversity in Lipid-Protein Interactions. *Chem. Rev.*

[63] Curtis, E. M. and Hall, C. K. (2013). Molecular dynamics simulations of DPPC bilayers using "lIME", a new coarse-grained model. *Journal of Physical Chemistry B*, 117(17):5019–5030.

[64] Darré, L., Machado, M., Brandner, A., González, H., Ferreira, S., and Pantano, S. (2015). SIRAH: A structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics. *Journal of Chemical Theory and Computation*, 11(2).

[65] Datar, A. V., Fyta, M., Bettolo-Marconi, U. M., and Melchionna, S. (2017). Electrokinetic Lattice Boltzmann Solver Coupled to Molecular Dynamics: Application to Polymer Translocation. *Langmuir*, 33:11635–11645.

[66] Daumke, O. and Praefcke, G. J. K. (2016). Invited review: Mechanisms of GTP hydrolysis and conformational transitions in the dynamin superfamily. *Biopolymers*, 105(8):580–593.

[67] de Jong, D. H., Singh, G., Bennett, W. F. D., Arnarez, C., Wassenaar, T. A., Schäfer, L. V., Periole, X., Tieleman, D. P., and Marrink, S. J. (2013). Improved Parameters for the Martini Coarse-Grained Protein Force Field. *Journal of Chemical Theory and Computation*, 9(1):687–697.

[68] De Planque, M. R. and Killian, J. A. (2003). Protein-lipid interactions studied with designed transmembrane peptides: Role of hydrophobic matching and interfacial anchoring (Review). *Molecular Membrane Biology*, 20(4):271–284.

[69] De Vecchis, D., Brandner, A., Baaden, M., Cohen, M. M., and Taly, A. (2019). A Molecular Perspective on Mitochondrial Membrane Fusion: From the Key Players to Oligomerization and Tethering of Mitofusin. *The Journal of Membrane Biology*.

[70] De Vecchis, D., Cavellini, L., Baaden, M., Hénin, J., Cohen, M. M., and Taly, A. (2017). A membrane-inserted structural model of the yeast mitofusin Fzo1. *Sci Rep*, 7(1):10217.

[71] Deschamps, J., Kantsler, V., Segre, E., and Steinberg, V. (2009). Dynamics of a vesicle in general flow. *Proceedings of the National Academy of Sciences*, 106(28):11444–11447.

[72] Dever, T. E., Glynias, M. J., and Merrick, W. C. (1987). GTP-binding domain: three consensus sequence elements with distinct spacing. *Proc. Natl. Acad. Sci. U.S.A.*, 84(7):1814–1818.

[73] Dohm, J. A., Lee, S. J., Hardwick, J. M., Hill, R. B., and Gittis, A. G. (2004). Cytosolic domain of the human mitochondrial fission protein fis1 adopts a TPR fold. *Proteins*, 54(1):153–156.

[74] Dumas, F., Lebrun, M. C., and Tocanne, J. F. (1999). Is the protein/lipid hydrophobic matching principle relevant to membrane organization and functions? *FEBS letters*, 458(3):271–277.

[75] Edidin, M. (2003). Lipids on the frontier: A century of cell-membrane bilayers. *Nature Reviews Molecular Cell Biology*, 4(5):414–418.

[76] Ermak, D. L. and McCammon, J. A. (1978). Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.*, 69:1352–1360.

[77] Eura, Y., Ishihara, N., Yokota, S., and Mihara, K. (2003). Two mitofusin proteins, mammalian homologues of FZO, with distinct functions are both required for mitochondrial fusion. *J. Biochem.*, 134(3):333–344.

[78] Ewald, P. P. (1921). Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik*, 369(3):253–287.

[79] Faelber, K., Dietrich, L., Noel, J. K., Wollweber, F., Pfitzner, A.-K., Mühleip, A., Sánchez, R., Kudryashev, M., Chiaruttini, N., Lilie, H., Schlegel, J., Rosenbaum, E., Hessenberger, M., Matthaeus, C., Kunz, S., von der Malsburg, A., Noé, F., Roux, A., van der Laan, M., Kühlbrandt, W., and Daumke, O. (2019). Structure and assembly of the mitochondrial membrane remodelling GTPase Mgm1. *Nature*, 571(7765):429–433.

[80] Faelber, K., Gao, S., Held, M., Posor, Y., Haucke, V., Noé, F., and Daumke, O. (2013). Oligomerization of dynamin superfamily proteins in health and disease. *Prog Mol Biol Transl Sci*, 117:411–443.

[81] Ferguson, S. M. and De Camilli, P. (2012). Dynamin, a membrane-remodelling GTPase. *Nat. Rev. Mol. Cell Biol.*, 13(2):75–88.

[82] Fiser, A., Do, R. K., and Sali, A. (2000). Modeling of loops in protein structures. *Protein science : a publication of the Protein Society*, 9(9):1753–1773.

[83] Fonseca, T. B., Sánchez-Guerrero, Á., Milosevic, I., and Raimundo, N. (2019). Mitochondrial fission requires DRP1 but not dynamins. *Nature*, 570(7761):E34—-E42.

[84] Franco, A., Kitsis, R. N., Fleischer, J. A., Gavathiotis, E., Kornfeld, O. S., Gong, G., Biris, N., Benz, A., Qvit, N., Donnelly, S. K., Chen, Y., Mennerick, S., Hodgson, L., Mochly-Rosen, D., and Dorn, G. W. (2016). Correcting mitochondrial fusion by manipulating mitofusin conformations. *Nature*, 540(7631):74–79.

[85] Frembgen-Kesner, T. and Elcock, A. H. (2009). Striking Effects of Hydrodynamic Interactions on the Simulated Diffusion and Folding of Proteins. *J. Chem. Theory Comput.*, 5(2):242–256.

[86] Frembgen-Kesner, T. and Elcock, A. H. (2013). Computer Simulations of the Bacterial Cytoplasm. *Biophys. Rev.*, 5:109–119.

[87] Frezza, C., Cipolat, S., de Brito, O., Micaroni, M., Beznoussenko, G. V., Rudka, T., Bartoli, D., Polishuck, R. S., Danial, N. N., De Strooper, B., and Scorrano, L. (2006). OPA1 controls apoptotic cristae remodeling independently from mitochondrial fusion. *Cell*, 126(1):177–189.

[88] Friedman, J. R., Lackner, L. L., West, M., DiBenedetto, J. R., Nunnari, J., and Voeltz, G. K. (2011). ER tubules mark sites of mitochondrial division. *Science*, 334(6054):358–362.

[89] Friedman, J. R. and Nunnari, J. (2014). Mitochondrial form and function. *Nature*, 505(7483):335–343.

[90] Friedman, R., Khalid, S., Aponte-Santamaría, C., Arutyunova, E., Becker, M., Boyd, K. J., Christensen, M., Coimbra, J. T., Concilio, S., Daday, C., van Eerden, F. J., Fernandes, P. A., Gräter, F., Hakobyan, D., Heuer, A., Karathanou, K., Keller, F., Lemieux, M. J., Marrink, S. J., May, E. R., Mazumdar, A., Naftalin, R., Pickholz, M., Piotto, S., Pohl, P., Quinn, P., Ramos, M. J., Schiøtt, B., Sengupta, D., Sessa, L., Vanni, S., Zeppelin, T., Zoni, V., Bondar, A. N., and Domene, C. (2018). Understanding Conformational Dynamics of Complex Lipid Mixtures Relevant to Biology. *Journal of Membrane Biology*, 251(5-6):609–631.

[91] Frisch, U., Hasslacher, B., and Pomeau, Y. (1986). Lattice-Gas Automata for the Navier-Stokes Equation. *Phys. Rev. Lett.*, 56(14):1505–1508.

[92] Frishman, D. and Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins*, 23(4):566–579.

[93] Fritz, S., Rapaport, D., Klanner, E., Neupert, W., and Westermann, B. (2001). Connection of the mitochondrial outer and inner membranes by Fzo1 is critical for organellar fusion. *J. Cell Biol.*, 152(4):683–692.

[94] Fröhlich, C., Grabiger, S., Schwefel, D., Faelber, K., Rosenbaum, E., Mears, J., Rocks, O., and Daumke, O. (2013). Structural insights into oligomerization and mitochondrial remodelling of dynamin 1-like protein. *EMBO J.*, 32(9):1280–1292.

[95] Gaede, H. C. and Gawrisch, K. (2003). Lateral diffusion rates of lipid, water, and a hydrophobic drug in a multilamellar liposome. *Biophysical Journal*, 85(3):1734–1740.

[96] Gallas, M. R., Dienhart, M. K., Stuart, R. A., and Long, R. M. (2006). Characterization of Mmp37p, a Saccharomyces cerevisiae mitochondrial matrix protein with a role in mitochondrial protein import. *Molecular biology of the cell*, 17(9):4051–4062.

[97] Gardner, J. M. and Abrams, C. F. (2018). Lipid flip-flop vs. lateral diffusion in the relaxation of hemifusion diaphragms. *Biochimica et Biophysica Acta - Biomembranes*, 1860(7):1452–1459.

[98] Genheden, S. and Essex, J. W. (2015). A Simple and Transferable All-Atom/Coarse-Grained Hybrid Model to Study Membrane Processes. *Journal of Chemical Theory and Computation*, 11(10):4749–4759.

[99] Georgakopoulos, N. D., Wells, G., and Campanella, M. (2017). The pharmacological regulation of cellular mitophagy. *Nature Chemical Biology*, 13(2):136–146.

[100] Godoy-Gallardo, M., York-Duran, M. J., and Hosta-Rigau, L. (2017). Recent Progress in Micro/Nanoreactors toward the Creation of Artificial Organelles. *Adv. Healthcare Mater.*, 1700917:1700917.

[101] Griffin, E. E. and Chan, D. C. (2006). Domain interactions within Fzo1 oligomers are essential for mitochondrial fusion. *J. Biol. Chem.*, 281(24):16599–16606.

[102] Griffin, E. E., Graumann, J., and Chan, D. C. (2005). The WD40 protein Caf4p is a component of the mitochondrial fission machinery and recruits Dnm1p to mitochondria. *J. Cell Biol.*, 170(2):237–248.

[103] Guex, N. and Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 18(15):2714–2723.

[104] Guidotti, G. (1972). DISCUSSION PAPER: MEMBRANE PROTEINS .

[105] Guy, H. R. and Seetharamulu, P. (1986). Molecular model of the action potential sodium channel. *Proceedings of the National Academy of Sciences of the United States of America*, 83(2):508–512.

[106] Hadley, K. R. and McCabe, C. (2010). A coarse-grained model for amorphous and crystalline fatty acids. *The Journal of chemical physics*, 132(13):134505.

[107] Hales, K. G. and Fuller, M. T. (1997). Developmentally regulated mitochondrial fusion mediated by a conserved, novel, predicted GTPase. *Cell*, 90(1):121–129.

[108] Hay, J. C. (2001). SNARE Complex Structure and Function. *Experimental Cell Research*, 271(1):10–21.

[109] He, S. and Scheraga, H. A. (1998). Brownian dynamics simulations of protein folding. *The Journal of Chemical Physics*, 108(1):287–300.

[110] Hermann, G. J., Thatcher, J. W., Mills, J. P., Hales, K. G., Fuller, M. T., Nunnari, J., and Shaw, J. M. (1998). Mitochondrial fusion in yeast requires the transmembrane GTPase Fzo1p. *J. Cell Biol.*, 143(2):359–373.

[111] Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.*, 18(12):1463–1472.

[112] Holt, A., Koehorst, R. B., Rutters-Meijneke, T., Gelb, M. H., Rijkers, D. T., Hemminga, M. A., and Killian, J. A. (2009). Tilt and rotation angles of a transmembrane model peptide as studied by fluorescence spectroscopy. *Biophysical Journal*, 97(8):2258–2266.

[113] Honda, S., Aihara, T., Hontani, M., Okubo, K., and Hirose, S. (2005). Mutational analysis of action of mitochondrial fusion factor mitofusin-2. *J. Cell. Sci.*, 118(Pt 14):3153–3161.

[114] Hoppins, S., Horner, J., Song, C., McCaffery, J. M., and Nunnari, J. (2009). Mitochondrial outer and inner membrane fusion requires a modified carrier protein. *J. Cell Biol.*, 184(4):569–581.

[115] Horvath, S. E. and Daum, G. (2013). Lipids of mitochondria. *Progress in Lipid Research*, 52(4):590–614.

[116] Hu, J., Shibata, Y., Zhu, P.-P., Voss, C., Rismanchi, N., Prinz, W. A., Rapoport, T. A., and Blackstone, C. (2009). A class of dynamin-like GTPases involved in the generation of the tubular ER network. *Cell*, 138(3):549–561.

[117] Huang, J. and MacKerell, A. D. J. (2013). CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of computational chemistry*, 34(25):2135–2145.

[118] Huang, K. and García, A. E. (2014). Acceleration of Lateral Equilibration in Mixed Lipid Bilayers Using Replica Exchange with Solute Tempering. *J. Chem. Theory Comput.*, 10(10):4264–4272.

[119] Hub, J. S., De Groot, B. L., and Van Der Spoel, D. (2010). G-wham-a free Weighted Histogram Analysis implementation including robust error and autocorrelation estimates. *Journal of Chemical Theory and Computation*, 6(12):3713–3720.

[120] Ishihara, N., Eura, Y., and Mihara, K. (2004). Mitofusin 1 and 2 play distinct roles in mitochondrial fusion reactions via GTPase activity. *J. Cell. Sci.*, 117(Pt 26):6535–6546.

[121] Ishihara, N., Jofuku, A., Eura, Y., and Mihara, K. (2003). Regulation of mitochondrial morphology by membrane potential, and DRP1-dependent division and FZO1-dependent fusion reaction in mammalian cells. *Biochem. Biophys. Res. Commun.*, 301(4):891–898.

[122] Ishihara, N., Nomura, M., Jofuku, A., Kato, H., Suzuki, S. O., Masuda, K., Otera, H., Nakanishi, Y., Nonaka, I., Goto, Y.-I., Taguchi, N., Morinaga, H., Maeda, M., Takayanagi, R., Yokota, S., and Mihara, K. (2009). Mitochondrial fission factor Drp1 is essential for embryonic development and synapse formation in mice. *Nat. Cell Biol.*, 11(8):958–966.

[123] Itoh, K., Nakamura, K., Iijima, M., and Sesaki, H. (2013). Mitochondrial dynamics in neurodegeneration. *Trends Cell Biol.*, 23(2):64–71.

[124] Jilly, R., Khan, N. Z., Aronsson, H., and Schneider, D. (2018). Dynamin-like proteins are potentially involved in membrane dynamics within chloroplasts and cyanobacteria. *Frontiers in Plant Science*, 9(February):1–13.

[125] Karplus, M. and Kuriyan, J. (2005). Molecular dynamics and protein function. *Proceedings of the National Academy of Sciences*, 102(19):6679–6685.

[126] Khaminets, A., Behl, C., and Dikic, I. (2016). Ubiquitin-Dependent And Independent Signals In Selective Autophagy. *Trends in Cell Biology*, 26(1):6–16.

[127] Killian, J. A. (1998). Hydrophobic mismatch between proteins and lipids in membranes. *Biochimica et biophysica acta*, 1376(3):401–415.

[128] Kim, K. T., Moon, Y., Jang, Y., Lee, K. T., Lee, C., Jun, Y., and Lee, S. (2017). Molecular mechanisms of atlastin-mediated ER membrane fusion revealed by a FRET-based single-vesicle fusion assay. *Sci Rep*, 7(1):8700.

[129] Kim, T. and Im, W. (2010). Revisiting hydrophobic mismatch with free energy simulation studies of transmembrane helix tilt and rotation. *Biophysical Journal*, 99(1):175–183.

[130] Kirsch, S. A. and Böckmann, R. A. (2016). Membrane pore formation in atomistic and coarse-grained simulations. *Biochim. Biophys. Acta. Biomembranes*, 1858(10):2266–2277.

[131] Knott, A. B., Perkins, G., Schwarzenbacher, R., and Bossy-Wetzel, E. (2008). Mitochondrial fragmentation in neurodegeneration. *Nat. Rev. Neurosci.*, 9(7):505–518.

[132] Kogan, M., Feng, B., Nordén, B., Rocha, S., and Beke-Somfai, T. (2014). Shear-induced membrane fusion in viscous solutions. *Langmuir*, 30(17):4875–4878.

[133] Koirala, S., Guo, Q., Kalia, R., Bui, H. T., Eckert, D. M., Frost, A., and Shaw, J. M. (2013). Interchangeable adaptors regulate mitochondrial dynamin assembly for membrane scission. *Proceedings of the National Academy of Sciences of the United States of America*, 110(15):E1342–51.

[134] Kong, L., Sochacki, K. A., Wang, H., Fang, S., Canagarajah, B., Kehr, A. D., Rice, W. J., Strub, M.-P., Taraska, J. W., and Hinshaw, J. E. (2018). Cryo-EM of the dynamin polymer assembled on lipid membrane. *Nature*, 560(7717):258–262.

[135] Koshiba, T., Detmer, S. A., Kaiser, J. T., Chen, H., McCaffery, J. M., and Chan, D. C. (2004). Structural basis of mitochondrial tethering by mitofusin complexes. *Science*, 305(5685):858–862.

[136] Kraus, F. and Ryan, M. T. (2017). The constriction and scission machineries involved in mitochondrial fission. *J Cell Sci*, 130(18):2953–2960.

[137] Kučerka, N., Nieh, M. P., and Katsaras, J. (2011). Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochimica et Biophysica Acta - Biomembranes*, 1808(11):2761–2771.

[138] Küchler, A., Yoshimoto, M., Luginbühl, S., Mavelli, F., and Walde, P. (2016). Enzymatic reactions in confined environments. *Nature Nanotech.*, 11(5):409–420.

[139] Kutik, S., Rissler, M., Guan, X. L., Guiard, B., Shui, G., Gebert, N., Heacock, P. N., Rehling, P., Dowhan, W., Wenk, M. R., Pfanner, N., and Wiedemann, N. (2008). The translocator maintenance protein Tam41 is required for mitochondrial cardiolipin biosynthesis. *The Journal of cell biology*, 183(7):1213–1221.

[140] Kuznetsov, A. V., Hermann, M., Saks, V., Hengster, P., and Margreiter, R. (2009). The cell-type specificity of mitochondrial dynamics. *Int. J. Biochem. Cell Biol.*, 41(10):1928–1939.

[141] Labrousse, A. M., Zappaterra, M. D., Rube, D. A., and van der Bliek, A. M. (1999). C. elegans dynamin-related protein DRP-1 controls severing of the mitochondrial outer membrane. *Mol. Cell*, 4(5):815–826.

[142] Lackner, L. L., Horner, J. S., and Nunnari, J. (2009). Mechanistic Analysis of a Dynamin Effector. *Science*, 325(5942):874–877.

[143] Languin-Cattoën, O., Melchionna, S., Derreumaux, P., Stirnemann, G., and Sterpone, F. (2018). Three Weaknesses for Three Perturbations: Comparing Protein Unfolding Under Shear, Force, and Thermal Stresses. *J. Phys. Chem. B*, 122:11922–11930.

[144] Lee, I.-H., Kai, H., Carlson, L.-A., Groves, J. T., and Hurley, J. H. (2015). Negative membrane curvature catalyzes nucleation of endosomal sorting complex required for transport (ESCRT)-III assembly. *Proc. Natl. Acad. Sci. {USA}*, 112(52):15892–15897.

[145] Lee, J. E., Westrate, L. M., Wu, H., Page, C., and Voeltz, G. K. (2016). Multiple dynamin family members collaborate to drive mitochondrial division. *Nature*, 540(7631):139–143.

[146] Lee, M., Moon, Y., Lee, S., Lee, C., and Jun, Y. (2019). Ergosterol interacts with Sey1p to promote atlastin-mediated endoplasmic reticulum membrane fusion in Saccharomyces cerevisiae. *FASEB J.*, 33(3):3590–3600.

[147] Leonarski, F., Trovato, F., Tozzini, V., and Trylska, J. (2011). Genetic algorithm optimization of force field parameters: Application to a coarse-grained model of RNA. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6623 LNCS:147–152.

[148] Li, Z., Okamoto, K.-I., Hayashi, Y., and Sheng, M. (2004). The importance of dendritic mitochondria in the morphogenesis and plasticity of spines and synapses. *Cell*, 119(6):873–887.

[149] Liesa, M. and Shirihai, O. S. (2013). Mitochondrial dynamics in the regulation of nutrient utilization and energy expenditure. *Cell Metab.*, 17(4):491–506.

[150] Linder, T., de Groot, B. L., and Stary-Weinzinger, A. (2013). Probing the Energy Landscape of Activation Gating of the Bacterial Potassium Channel KcsA. *PLOS Computational Biology*, 9(5):1–9.

[151] Logan, D. C. (2010). The dynamic plant chondriome. *Semin. Cell Dev. Biol.*, 21(6):550–557.

[152] Low, H. H. and Löwe, J. (2006). A bacterial dynamin-like protein. *Nature*, 444(7120):766–769.

[153] Low, H. H., Sachse, C., Amos, L. A., and Löwe, J. (2009). Structure of a bacterial dynamin-like protein lipid tube provides a mechanism for assembly and membrane curving. *Cell*, 139(7):1342–1352.

[154] Lyman, E., Hsieh, C. L., and Eggeling, C. (2018). From Dynamics to Membrane Organization: Experimental Breakthroughs Occasion a "Modeling Manifesto". *Biophysical Journal*, 115(4):595–604.

[155] Lyubartseva, A. P. and Rabinovich, A. L. (2016). Force Field Development for Lipid Membrane Simulations. *Biochim. Biophys. Acta. Biomembranes*, 1858:2483–2497.

[156] Maiese, K. (2016). The bright side of reactive oxygen species: lifespan extension without cellular demise. *J Transl Sci*, 2(3):185–187.

[157] Malevanets, A. and Kapral, R. (2000). Solute molecular dynamics in a mesoscale solvent. *J. Chem. Phys.*, 112:7260–7269.

[158] Marrink, S. J., Corradi, V., Souza, P. C., Ingólfsson, H. I., Tieleman, D. P., and Sansom, M. S. (2019). Computational Modeling of Realistic Cell Membranes. *Chemical Reviews*, 119(9):6184–6226.

[159] Marrink, S. J., de Vries, A. H., and Mark, A. E. (2004). Coarse Grained Model for Semiquantitative Lipid Simulations. *The Journal of Physical Chemistry B*, 108(2):750–760.

[160] Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., and De Vries, A. H. (2007). The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B*, 111(27):7812–7824.

[161] Martinez, L., Andrade, R., Birgin, E. G., and Martínez, J. M. (2009). PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.*, 30(13):2157–2164.

[162] Maupetit, J., Tuffery, P., and Derreumaux, P. (2007). A coarse-grained protein force field for folding and structure prediction. *Proteins: Structure, Function and Genetics*, 69(2):394–408.

[163] Mcdonagh, J. L., Shkurti, A., Bray, D. J., Anderson, R. L., and Pyzer-knapp, E. O. (2019). Utilizing Machine Learning for Efficient Parameterization of Coarse Grained Molecular Force Fields. (1):1–33.

[164] McNamara, G. R. and Zanetti, G. (1988). Use of the Boltzmann Equation to Simulate Lattice-Gas Automata. *Physical Review Letters*, 61(20):2332–2335.

[165] Mears, J. A., Lackner, L. L., Fang, S., Ingerman, E., Nunnari, J., and Hinshaw, J. E. (2011). Conformational changes in Dnm1 support a contractile mechanism for mitochondrial fission. *Nat. Struct. Mol. Biol.*, 18(1):20–26.

[166] Merz, S. and Westermann, B. (2009). Genome-wide deletion mutant analysis reveals genes required for respiratory growth, mitochondrial genome maintenance and mitochondrial protein synthesis in Saccharomyces cerevisiae. *Genome Biol.*, 10(9):R95.

[167] Mikhailov, A. S. and Kapral, R. (2015). Hydrodynamic collective effects of active protein machines in solution and lipid bilayers. *Proc. Natl. Acad. Sci. {USA}*, 112(28):E3639—-E3644.

[168] Mills, Z. G., Mao, W., and Alexeev, A. (2013). Mesoscale modeling: solving complex flows in biology and biotechnology. *Trends Biotechnol*, 31:426–434.

[169] Milo, R. and Phillips, R. (2016). *Cell biology by the numbers*.

[170] Monticelli, L., Kandasamy, S. K., Periole, X., Larson, R. G., Tieleman, D. P., and Marrink, S.-J. J. (2008). The MARTINI Coarse-Grained Force Field: Extension to Proteins. *Journal of Chemical Theory and Computation*, 4(5):819–834.

[171] Monticelli, L., Tieleman, D. P., and Fuchs, P. F. (2010). Interpretation of 2H-NMR experiments on the orientation of the transmembrane helix WALP23 by computer simulations. *Biophysical Journal*, 99(5):1455–1464.

[172] Moss, T. J., Andreazza, C., Verma, A., Daga, A., and McNew, J. A. (2011). Membrane fusion by the GTPase atlastin requires a conserved C-terminal cytoplasmic tail and dimerization through the middle domain. *Proceedings of the National Academy of Sciences of the United States of America*, 108(27):11133–11138.

[173] Mouritsen, O. G. and Bloom, M. (1984). Mattress model of lipid-protein interactions in membranes. *Biophysical Journal*, 46(2):141–153.

[174] Mouritsen, O. G. and Bloom, M. (1993). Models of Lipid-Protein Interactions in Membranes. *Annual Review of Biophysics and Biomolecular Structure*, 22(1):145–171.

[175] Mozdy, A. D., McCaffery, J. M., and Shaw, J. M. (2000). Dnm1p GTPase-mediated mitochondrial fission is a multi-step process requiring the novel integral membrane component Fis1p. *J. Cell Biol.*, 151(2):367–380.

[176] Murtola, T., Falck, E., Patra, M., Karttunen, M., and Vattulainen, I. (2004). Coarse-grained model for phospholipid/cholesterol bilayer. *Journal of Chemical Physics*, 121(18):9156–9165.

[177] M.Venable, Ingólfsson, H., Lerner, M. G., Perrin, S., Camley, B. A., Marrink, S. J., Brown, F. L. H., and Pastor, R. W. (2016). Lipid and Peptide Diffusion in Bilayers: The Saffman-Delbruck Model and Periodic Boundary Conditions. *J. Phys. Chem. B*, 121:3443–3457.

[178] Nakamura, K., Omori, T., and Ishikawa, T. (2019). Shear-induced migration of a transmembrane protein within a vesicle. *Biophy. J.*, 116:1483–1494.

[179] Neale, C., Madill, C., Rauscher, S., and Pomès, R. (2013). Accelerating Convergence in Molecular Dynamics Simulations of Solutes in Lipid Membranes by Conducting a Random Walk along the Bilayer Normal. *J. Chem. Theory Comput.*, 9(8):3686–3703.

[180] Noid, W. G. (2013). Perspective: Coarse-grained models for biomolecular systems. *The Journal of Chemical Physics*, 139(9):90901.

[181] Nury, H., Poitevin, F., Van Renterghem, C., Changeux, J. P., Corringer, P. J., Delarue, M., and Baaden, M. (2010). One-microsecond molecular dynamics simulation of channel gating in a nicotinic receptor homologue. *Proceedings of the National Academy of Sciences of the United States of America*, 107(14):6275–6280.

[182] Okamoto, K. and Shaw, J. M. (2005). Mitochondrial morphology and dynamics in yeast and multicellular eukaryotes. *Annu. Rev. Genet.*, 39:503–536.

[183] Ono, T., Isobe, K., Nakada, K., and Hayashi, J. I. (2001). Human cells are protected from mitochondrial dysfunction by complementation of DNA products in fused mitochondria. *Nat. Genet.*, 28(3):272–275.

[184] Orso, G., Pendin, D., Liu, S., Tosetto, J., Moss, T. J., Faust, J. E., Micaroni, M., Egorova, A., Martinuzzi, A., McNew, J. A., and Daga, A. (2009). Homotypic fusion of ER membranes requires the dynamin-like GTPase atlastin. *Nature*, 460(7258):978–983.

[185] Osellame, L. D., Singh, A. P., Stroud, D. A., Palmer, C. S., Stojanovski, D., Ramachandran, R., and Ryan, M. T. (2016). Cooperative and independent roles of Drp1 adaptors Mff and MiD49/51 in mitochondrial fission. *J. Cell. Sci.*

[186] Otera, H., Wang, C., Cleland, M. M., Setoguchi, K., Yokota, S., Youle, R. J., and Mihara, K. (2010). Mff is an essential factor for mitochondrial recruitment of Drp1 during mitochondrial fission in mammalian cells. *J. Cell Biol.*, 191(6):1141–1158.

[187] Otsuga, D., Keegan, B. R., Brisch, E., Thatcher, J. W., Hermann, G. J., Bleazard, W., and Shaw, J. M. (1998). The dynamin-related GTPase, Dnm1p, controls mitochondrial morphology in yeast. *J. Cell Biol.*, 143(2):333–349.

[188] Ozturk, S. B. and Kinzy, T. G. (2008). Guanine Nucleotide Exchange Factor Independence of the G-protein eEF1A through Novel Mutant Forms and Biochemical Properties. *J Biol Chem*, 283(34):23244–23253.

[189] Palmer, C. S., Osellame, L. D., Laine, D., Koutsopoulos, O. S., Frazier, A. E., and Ryan, M. T. (2011). MiD49 and MiD51, new components of the mitochondrial fission machinery. *EMBO Rep.*, 12(6):565–573.

[190] Parrinello, M. and Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190.

[191] Pasi, M., Lavery, R., and Ceres, N. (2013). PaLaCe: A Coarse-Grain Protein Model for Studying Mechanical Properties. *Journal of Chemical Theory and Computation*, 9(1):785–793.

[192] Patel, R. Y. and Balaji, P. V. (2005). Effect of the Choice of the Pressure Coupling Method on the Spontaneous Aggregation of DPPC Molecules. *J. Phys. Chem. B*, 109:14667–14674.

[193] Perez, A. and Orozco, M. (2010). Real-Time Atomistic Description of DNA Unfolding. *Angewandte Chemie International Edition*, 49(28):4805–4808.

[194] Periole, X., Cavalli, M., Marrink, S.-j., and Ceruso, M. A. (2009). Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J Chem Theory Comput*, 5(9):1–7.

[195] Peschka, R., Dennehy, C., and Szoka, F. C. (1998). A simple in vitro model to study the release kinetics of liposome encapsulated material. *Journal of Controlled Release*, 56(1-3):41–51.

[196] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612.

[197] Praefcke, G. J. K. and McMahon, H. T. (2004). The dynamin superfamily: universal membrane tubulation and fission molecules? *Nat. Rev. Mol. Cell Biol.*, 5(2):133–147.

[198] Qi, Y., Ingólfsson, H. I., Cheng, X., Lee, J., Marrink, S. J., and Im, W. (2015). CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *J. Chem. Theory Comput.*, 11(9):4486–4494.

[199] Qi, Y., Yan, L., Yu, C., Guo, X., Zhou, X., Hu, X., Huang, X., Rao, Z., Lou, Z., and Hu, J. (2016). Structures of human mitofusin 1 provide insight into mitochondrial tethering. *Journal of Cell Biology*, 215(5):621–629.

[200] Qian, Y. H., D Humières, D., and Lallemand, P. (1992). Lattice BGK Models for Navier-Stokes Equation. *Europhysics Letters (EPL)*, 17(6):479–484.

[201] Rahaman, O., Melchionna, S., Laage, D., and Sterpone, F. (2013). The Effect of Protein Composition on Hydration Dynamics. *Phys. Chem. Chem. Phys.*, 15(10):3570–3576.

[202] Rallabandi, B., Oppenheimer, N., Zion, M. Y. B., and Stone, H. A. (2018). Membrane-induced hydroelastic migration of a particle surfing its own wave. *Nat. Physics*, 14:1211–1215.

[203] Ramachandran, R. (2018). Mitochondrial dynamics: The dynamin superfamily and execution by collusion. *Seminars in Cell & Developmental Biology*, 76:201–212.

[204] Rapaport, D., Brunner, M., Neupert, W., and Westermann, B. (1998). Fzo1p is a mitochondrial outer membrane protein essential for the biogenesis of functional mitochondria in Saccharomyces cerevisiae. *J. Biol. Chem.*, 273(32):20150–20155.

[205] Rasmussen, S. G., Devree, B. T., Zou, Y., Kruse, A. C., Chung, K. Y., Kobilka, T. S., Thian, F. S., Chae, P. S., Pardon, E., Calinski, D., Mathiesen, J. M., Shah, S. T., Lyons, J. A., Caffrey, M., Gellman, S. H., Steyaert, J., Skiniotis, G., Weis, W. I., Sunahara, R. K., and Kobilka, B. K. (2011). Crystal structure of the β 2 adrenergic receptor-Gs protein complex. *Nature*, 477(7366):549–557.

[206] Risselada, H. J., Bubnis, G., and Grubmüller, H. (2014). Expansion of the fusion stalk and its implication for biological membrane fusion. *Proceedings of the National Academy of Sciences of the United States of America*, 111(30):11043–11048.

[207] Rojo, M., Legros, F., Chateau, D., and Lombès, A. (2002). Membrane topology and mitochondrial targeting of mitofusins, ubiquitous mammalian homologs of the transmembrane GTPase Fzo. *J. Cell. Sci.*, 115(Pt 8):1663–1674.

[208] Roosen-Runge, F., Hennig, M., Zhang, F., Jacobs, R. M. J., Sztucki, M., Schober, H., Seydel, T., and Schreiber, F. (2011). Protein self-diffusion in crowded solutions. *Proc. Natl. Acad. Sci. {USA}*, 108:11815–11820.

[209] Rothman, J. S., Kocsis, L., Herzog, E., Nusser, Z., and Silver, R. A. (2016). Physical determinants of vesicle mobility and supply at a central synapse. *eLife*, 5:e15133.

[210] Rouse, P. E. (1953). A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers. *J. Chem. Phys.*, 21:1272–1280.

[211] Roux, B. (1997). Influence of the membrane potential on the free energy of an intrinsic protein. *Biophysical journal*, 73(6):2980–2989.

[212] Šachl, R., Amaro, M., Aydogan, G., Koukalová, A., Mikhalyov, I. I., Boldyrev, I. A., Humpolíčková, J., and Hof, M. (2015). On multivalent receptor activity of GM1 in cholesterol containing membranes. *Biochim. Biophys. Acta. Mol. Cell. Res.*, 1853(4):850–857.

[213] Sahoo, B. R., Genjo, T., Cox, S. J., Stoddard, A. K., Anantharamaiah, G. M., Fierke, C., and Ramamoorthy, A. (2018). Nanodisc-Forming Scaffold Protein Promoted Retardation of Amyloid-Beta Aggregation. *Journal of Molecular Biology*, 430(21):4230–4244.

[214] Santel, A., Frank, S., Gaume, B., Herrler, M., Youle, R. J., and Fuller, M. T. (2003). Mitofusin-1 protein is a generally expressed mediator of mitochondrial fusion in mammalian cells. *J. Cell. Sci.*, 116(Pt 13):2763–2774.

[215] Santel, A. and Fuller, M. T. (2001). Control of mitochondrial morphology by a human mitofusin. *J. Cell. Sci.*, 114(Pt 5):867–874.

[216] Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990). The P-loop–a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.*, 15(11):430–434.

[217] Sato, H. and Feix, J. B. (2006). Peptide-membrane interactions and mechanisms of membrane destruction by amphipathic α-helical antimicrobial peptides. *Biochimica et Biophysica Acta - Biomembranes*, 1758(9):1245–1256.

[218] Sengupta, D. and Chattopadhyay, A. (2012). Identification of Cholesterol Binding Sites in the Serotonin1A Receptor. *The Journal of Physical Chemistry B*, 116(43):12991–12996.

[219] Sesaki, H. and Jensen, R. E. (1999). Division versus Fusion: Dnm1p and Fzo1p Antagonistically Regulate Mitochondrial Shape. *J Cell Biol*, 147(4):699–706.

[220] Sesaki, H. and Jensen, R. E. (2001). UGO1 encodes an outer membrane protein required for mitochondrial fusion. *J. Cell Biol.*, 152(6):1123–1134.

[221] Sesaki, H. and Jensen, R. E. (2004). Ugo1p links the Fzo1p and Mgm1p GTPases for mitochondrial fusion. *J. Biol. Chem.*, 279(27):28298–28303.

[222] Sharma, V., Belevich, G., Gamiz-Hernandez, A. P., Róg, T., Vattulainen, I., Verkhovskaya, M. L., Wikström, M., Hummer, G., and Kaila, V. R. I. (2015). Redox-induced activation of the proton pump in the respiratory complex I. *Proc. Natl. Acad. Sci. {USA}*, 112(37):11571–11576.

[223] Sharp, L., Salari, R., and Brannigan, G. (2019). Boundary lipids of the nicotinic acetylcholine receptor: Spontaneous partitioning via coarse-grained molecular dynamics simulation. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1861(4):887–896.

[224] Shen, M.-Y. and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein science : a publication of the Protein Society*, 15(11):2507–2524.

[225] Shen, Z., Ye, H., Kroger, M., and Li, Y. (2017). Self-assembled core–polyethylene glycol–lipid shell nanoparticles demonstrate high stability in shear flow. *Phys. Chem. Chem. Phys.*, 19:13294.

[226] Shevchenko, A. and Simons, K. (2010). Lipidomics: coming to grips with lipid diversity. *Nat. Rev. Mol. Cell Biol.*, 11:593–598.

[227] Shim, S.-H., Babcock, H. P., Zhuang, X., Huang, B., Zhong, G., Xu, C., Vaughan, J. C., Bi, G.-Q., Wang, X., and Xia, C. (2012). Super-resolution fluorescence imaging of organelles in live cells with photoswitchable membrane probes. *Proc. Natl. Acad. Sci. {USA}*, 109(35):13978–13983.

[228] Shinoda, W. (2016). Permeability across lipid membranes. *Biochim. Biophys. Acta. Biomembranes*, 1858(10):2254–2265.

[229] Shrivastava, I. H. and Sansom, M. S. P. (2000). Simulations of Ion Permeation Through a Potassium Channel: Molecular Dynamics of KcsA in a Phospholipid Bilayer. *Biophysical Journal*, 78(2):557–570.

[230] Shutt, T., Geoffrion, M., Milne, R., and McBride, H. M. (2012). The intracellular redox state is a core determinant of mitochondrial fusion. *EMBO Rep*, 13(10):909–915.

[231] Silva, J. R., Pan, H., Wu, D., Nekouzadeh, A., Decker, K. F., Cui, J., Baker, N. A., Sept, D., and Rudy, Y. (2009). A multiscale model linking ion-channel molecular dynamics and electrostatics to the cardiac action potential. *Proceedings of the National Academy of Sciences of the United States of America*, 106(27):11102–11106.

[232] Singer, S. and Nicolson, G. (1972). The Fluid Mosaic Model of the Structure of Cell Membranes. *Science*, 175(4023):720–731.

[233] Skulachev, V. P. (2001). Mitochondrial filaments and clusters as intracellular power-transmitting cables. *Trends Biochem. Sci.*, 26(1):23–29.

[234] Smirnova, E., Griparic, L., Shurland, D. L., and van der Bliek, A. M. (2001). Dynamin-related protein Drp1 is required for mitochondrial division in mammalian cells. *Mol. Biol. Cell*, 12(8):2245–2256.

[235] Smirnova, E., Shurland, D. L., Ryazantsev, S. N., and van der Bliek, A. M. (1998). A human dynamin-related protein controls the distribution of mitochondria. *J. Cell Biol.*, 143(2):351–358.

[236] Smirnova, Y. G., Risselada, H. J., and Müller, M. (2019). Thermodynamically reversible paths of the first fusion intermediate reveal an important role for membrane anchors of fusion proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 116(7):2571–2576.

[237] Sodt, A. J. and Head-Gordon, T. (2010). An implicit solvent coarse-grained lipid model with correct stress profile. *Journal of Chemical Physics*, 132(20):1–9.

[238] Steck, T. L. (1974). The Organization of proteins in the human red blood cell membrane. *Journal of cell biology*, 62:1–19.

[239] Sterpone, F., Derreumaux, P., and Melchionna, S. (2015). Protein simulations in fluids: Coupling the OPEP coarse-grained force field with hydrodynamics. *Journal of Chemical Theory and Computation*, 11(4):1843–1853.

[240] Sterpone, F., Derreumaux, P., and Melchionna, S. (2018). Molecular Mechanism of Protein Unfolding under Shear: A Lattice Boltzmann Molecular Dynamics Study. *J. Phys. Chem. B*, 122:1573–1579.

[241] Sterpone, F., Doutreligne, S., Tran, T. T., Melchionna, S., Baaden, M., Nguyen, P. H., and Derreumaux, P. (2017). Multi-Scale Simulations of Biological Systems Using the OPEP Coarse-Grained Model. *Bioch. Biophys. Res. Comm.*, 498:296–304.

[242] Sterpone, F., Melchionna, S., Tuffery, P., Pasquali, S., Mousseau, N., Cragnolini, T., Chebaro, Y., St-Pierre, J.-F., Kalimeri, M., Barducci, A., Laurin, Y., Tek, A., Baaden, M., Nguyen, P. H., and Derreumaux, P. (2014). The opep protein model: from single molecules, amyloid formation, crowding and hydrodynamics to dna/rna systems. *Chem. Soc. Rev.*, 43:4871–4893.

[243] Stone, M. B., Shelby, S. A., and Veatch, S. L. (2017). Super-Resolution Microscopy: Shedding Light on the Cellular Plasma Membrane. *Chem. Rev.*, 117(11):7457–7477.

[244] Strandberg, E., Özdirekcan, S., Rijkers, D. T., Van Der Wel, P. C., Koeppe, R. E., Liskamp, R. M., and Killian, J. A. (2004). Tilt angles of transmembrane model peptides in oriented and non-oriented lipid bilayers as determined by 2H solid-state NMR. *Biophysical Journal*, 86(6):3709–3721.

[245] Succi, S. (2001). Lattice Boltzmann Equation. 240:221–228.

[246] Succi, S., Amati, G., Bernaschi, M., Falcucci, G., Lauricella, M., and Montessori, A. (2019). Towards Exascale Lattice Boltzmann computing. *Computers and Fluids*, 181:107–115.

[247] Sun, F., Huo, X., Zhai, Y., Wang, A., Xu, J., Su, D., Bartlam, M., and Rao, Z. (2005). Crystal Structure of Mitochondrial Respiratory Membrane Protein Complex II. *Cell*, 121(7):1043–1057.

[248] Ta, T. and Porter, T. M. (2013). Thermosensitive liposomes for localized delivery and triggered release of chemotherapy. *Journal of Controlled Release*, 169(1-2):112–125.

[249] Taguchi, N., Ishihara, N., Jofuku, A., Oka, T., and Mihara, K. (2007). Mitotic phosphorylation of dynamin-related GTPase Drp1 participates in mitochondrial fission. *J. Biol. Chem.*, 282(15):11521–11529.

[250] Tieleman, D. P. and Berendsen, H. J. C. (1996). Molecular dynamics simulations of a fully hydrated dipalmitoylphosphatidylcholine bilayer with different macroscopic boundary conditions and parameters. *J. Chem. Phys.*, 105:4871–4880.

[251] Tieu, Q., Okreglak, V., Naylor, K., and Nunnari, J. (2002). The WD repeat protein, Mdv1p, functions as a molecular adaptor by interacting with Dnm1p and Fis1p during mitochondrial fission. *J. Cell Biol.*, 158(3):445–452.

[252] Tilokani, L., Nagashima, S., Paupe, V., and Prudent, J. (2018). Mitochondrial dynamics: overview of molecular mechanisms. *Essays In Biochemistry*, 62(3):341–360.

[253] Torre, G. D. L., Huertas, J. M. L., and Carrasco, B. (2000). Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.*, 78:719–730.

[254] Trick, J. L., Chelvaniththilan, S., Klesse, G., Aryal, P., Wallace, E. J., Tucker, S. J., and Sansom, M. S. (2016). Functional Annotation of Ion Channel Structures by Molecular Simulation. *Structure*, 24(12):2207–2216.

[255] van der Ploeg, P. and Berendsen, H. (1982). Molecular dynamics simulation of a bilayer membrane.

[256] Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. *J. Comput. Chem.*, 26:1701–1718.

[257] van Eerden, F. J., van den Berg, T., Frederix, P. W. J. M., de Jong, D. H., Periole, X., and Marrink, S. J. (2017). Molecular Dynamics of Photosystem II Embedded in the Thylakoid Membrane. *The Journal of Physical Chemistry B*, 121(15):3237–3249.

[258] van Meer D. R. Voelker, G. and Feigenson, G. W. (2009). Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell. Biol.*, 9:112–124.

[259] Vetter, I. R. and Wittinghofer, A. (2001). The guanine nucleotide-binding switch in three dimensions. *Science*, 294(5545):1299–1304.

[260] Visscher, K. M., Medeiros-Silva, J., Mance, D., Rodrigues, J. P., Daniëls, M., Bonvin, A. M., Baldus, M., and Weingarth, M. (2017). Supramolecular Organization and Functional Implications of K+ Channel Clusters in Membranes. *Angewandte Chemie - International Edition*, 56(43):13222–13227.

[261] Vögele, M. and Hummer, G. (2016). Divergent Diffusion Coefficients in Simulations of Fluids and Lipid Membranes. *J. Phys. Chem. B*, 120:8722–8732.

[262] Vögele, M., Köfinger, J., and Hummer, G. (2018). Hydrodynamics of Diffusion in Lipid Membrane Simulations. *Physical Review Letters*, 120(26):1–16.

[263] Wakabayashi, M. and Matsuzaki, K. (2007). Formation of amyloids by Abeta-(1-42) on NGF-differentiated PC12 cells: roles of gangliosides and cholesterol. *Journal of molecular biology*, 371(4):924–933.

[264] Wan, J., Ristenpart, W. D., and Stone, H. A. (2009). Dynamics of shear-induced ATP release from red blood cells. *Proc. Natl. Acad. Sci. USA*, 105:16432–16437.

[265] Wang, J., Kaplan, J. A., Colson, Y. L., and Grinstaff, M. W. (2017). Mechanoresponsive Materials for Drug Delivery: Harnessing Forces for Controlled Release. *Adv. Drug Deliv. Rev.*, 108:68–82.

[266] Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P., Marrink, S. J., of Chemical Theory, S. J. M. J., and Computation 2015 11 (5), .-. (2015). Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *Journal of Chemical Theory and Computation*, 11(5):2144–2155.

[267] Westermann, B. (2010). Mitochondrial fusion and fission in cell life and death. *Nat. Rev. Mol. Cell Biol.*, 11(12):872–884.

[268] Wong, E. D., Wagner, J. A., Gorsich, S. W., McCaffery, J. M., Shaw, J. M., and Nunnari, J. (2000). The dynamin-related GTPase, Mgm1p, is an intermembrane space protein required for maintenance of fusion competent mitochondria. *J. Cell Biol.*, 151(2):341–352.

[269] Wong, E. D., Wagner, J. A., Scott, S. V., Okreglak, V., Holewinske, T. J., Cassidy-Stone, A., and Nunnari, J. (2003). The intramitochondrial dynamin-related GTPase, Mgm1p, is a component of a protein complex that mediates mitochondrial fusion. *J. Cell Biol.*, 160(3):303–311.

[270] Yan, L., Sun, S., Wang, W., Shi, J., Hu, X., Wang, S., Su, D., Rao, Z., Hu, J., and Lou, Z. (2015). Structures of the yeast dynamin-like GTPase Sey1p provide insight into homotypic ER fusion. *J Cell Biol*, 210(6):961–972.

[271] Yan, L., Yu, C., Rao, Z., Lou, Z., Qi, Y., Huang, X., Guo, X., Lan, L., and Hu, J. (2018). Structural basis for GTP hydrolysis and conformational change of MFN1 in mediating membrane fusion. *Nature Structural and Molecular Biology*, 25(3):233–243.

[272] Yu, R., Jin, S., Lendahl, U., Nistér, M., and Zhao, J. (2019). Human Fis1 regulates mitochondrial dynamics through inhibition of the fusion machinery. *EMBO J*, 38(8).

[273] Zabusky, N. J., Segre, E., Deschamps, J., Kantsler, V., and Steinberg, V. (2011). Dynamics of vesicles in shear and rotational flows: Modal dynamics and phase diagram. *Phys. Fluids*, 23:41905.

[274] Zgorski, A. and Lyman, E. (2016). Toward Hydrodynamics with Solvent Free Lipid Models: STRD Martini. *Biophys. J.*, 111(12):2689–2697.

[275] Zhao, G., Perilla, J. R., Yufenyuy, E. L., Meng, X., Chen, B., Ning, J., Ahn, J., Gronenborn, A. M., Schulten, K., Aiken, C., and Zhang, P. (2013). Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature*, 497(7451):643–646.

[276] Zhong, Q., Jiang, Q., Moore, P. B., Newns, D. M., and Klein, M. L. (1998). Molecular Dynamics Simulation of a Synthetic Ion Channel. *Biophysical Journal*, 74(1):3–10.

[277] Zimm, B. H. (1956). Dynamics of Polymer Molecules in Dilute Solution: Viscoelasticity, Flow Birefringence and Dielectric Loss. *J. Chem. Phys.*, 24:269–278.

Appendix A

Peer-reviewed publications derived from this thesis

1. **Physics-based oligomeric models of the yeast mitofusin Fzo1 at the molecular scale in the context of membrane docking.** A Brandner, D De Vecchis, M Baaden, MM Cohen, A Taly. Mitochondrion (2019) *In press*

2. **Structural dataset from microsecond-long simulations of yeast mitofusin Fzo1 in the context of membrane docking.** A Brandner, D De Vecchis, M Baaden, MM Cohen, A Taly Data in Brief (2019), 104460

3. **A Molecular Perspective on Mitochondrial Membrane Fusion: From the Key Players to Oligomerization and Tethering of Mitofusin.** D De Vecchis, A Brandner, M Baaden, MM Cohen, A Taly. J Membrane Biol (2019). https://doi.org/10.1007/s00232-019-00089-y

4. **Modelling lipid systems in fluid with Lattice Boltzmann Molecular Dynamics simulations and hydrodynamics.** AF Brandner, S Timr, S Melchionna, P Derreumaux, M Baaden, P. Derreumaux, F Sterpone. Scientific Reports (2019).*Under Revision*