Imperial College London
Faculty of Natural Sciences
Department of Chemistry

Molecular insight into the barrier function of sebaceous and skin lipids and the interactions of lipid membranes with antimicrobial peptides and therapeutic nanoparticles

Anna Sofia Tascini

A thesis presented for the degree of Doctor of Philosophy,
London, September 2018

Declaration

I declare that all the work presented in this thesis is my own, unless clearly stated and referenced otherwise.

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

Sebaceous and skin lipids from the outermost layer of the epidermis, the stratum corneum (SC), play a vital role in regulating the hydration level of our skin. Skin lipids constitute the first layer of defence against toxic particles entering the body, and the first barrier to overcome when cosmetic and therapeutic creams are applied to the skin.

This thesis gives insights into the properties of the skin surface at the microscopic level by introducing a model of the sebum, the oily layer covering the SC. Molecular dynamics simulations are used to investigate the structural and thermodynamic properties of the sebum in bulk and at air and water interfaces, and to study its interactions with water and SC lipids. The sebum model accurately reproduces experimental results, such as density, surface tensions and surface pressure-area isotherms. Calculating the free energy barrier to water penetration indicates that the sebum is permeable to water diffusion via a network of clustered triglyceride glycerol groups, suggesting that the sebum may play a significant role in the water regulation function of the skin.

The model is also used to obtain a molecular understanding of SC lipid-sebum interactions. It is shown that sebum triglycerides interact strongly with epidermic lipids and extract them from the bilayer. Using microsecond simulations, the mechanisms associated with the extraction processes are identified and the corresponding free energies are quantified. Further, it is shown that the skin surface has a low affinity for water (contact angle: $\sim 100^{\circ}$), highlighting a striking difference between SC lipids and phospholipids.

Heat transfer from a nanoparticle and the conformations of an antimicrobial peptide are also studied, with the eventual, future goal of combining these with the sebum-SC model. Finally, an outreach activity to improve the general public's understanding of lipid molecules and their amphiphilic structure or "split-personality", is described.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors Prof. Fernando Bresme, Prof. John Seddon and Dr Rongjun Chen for their support throughout the development of the scientific research and the outreach activities presented in this thesis. In particular, a special thanks to Prof. Fernando Bresme who believed I could succeed doing research in a field not closely related to my previous background and motivated me to pursue a research career at a doctoral level. Despite the initial difficulties, during these years I learnt that with passion, motivation and hard work, also the most difficult challenges can be faced.

I am truly grateful to my PhD examiners, Dr Rebecca Notman and Prof. Oscar Ces, for their thoughtful and accurate revisions and positive feedback on my work.

My sincere thanks also go to Dr Massimo Noro and Dr Ian Tucker who provided me with the opportunity to visit the Port-Sunlight Research Lab at Unilever, and for the useful feedback on my research project.

I am also thankful to Prof. Pietro Asinari, Dr Eliodoro Chiavazzo and Dr. Matteo Fasano for developing together the research about nanoparticle interfaces and for helping me to create the connection with Imperial College, which has been the perfect environment to do research and develop my carrier.

I feel honoured I had the possibility of doing research within the multidisciplinary environment of the EU international training network, SNAL. Thanks to everyone in the network for the discussions, insightful comments and encouragement, which incited me to widen my research from various and different perspectives. One of the greatest achievement of the SNAL project has also been all the friendships I have formed. A special mention to Daniela Ciumac, who I have enjoyed working together with on the peptide project (I can't wait to tell her "Yes, Dana, I made it, too!"), to Berardo Mario Manzi (and Fabiola) for their hospitality during my secondment in Tarragona and all the useful and enjoyable talks, to Chanfei Su for the great time we spent together at the conference in Helsinki.

Thanks to everyone I met in Fernando Bresme's group, all inspiring researchers and excellent friends: James Carter, Silvia Di Lecce, Gøran Svaland, Juan Olarte Plata, Ollie Robotham, Oliver Gittus, Miguel Angel Gonzales, Oscar Fajardo, Irene Iriarte Carratero, Niall Jackson, Wendy Wang, Jeff Armstrong. Thanks for the advises, lunch breaks, Friday pints and all the

activities organised together. In particular, I would like to express my deepest gratitude to James Carter for proofreading the English of this thesis, my paper drafts, important emails, for the code–sharing, for his support in these years and all the great jazz concerts we have been to together, which have inspired my thoughts.

Thanks to all the volunteers helping out with the outreach activity "Fats' split personality" for sharing my enthusiasm, giving suggestions and making things more relaxed and more enjoyable. In alphabetical order: Zainab Ahmed, Marie Bachelet, James Carter, Yu Han Chang, Siyuan Chen, Silvia Di Lecce, Gøran Svatland, Christian Jürgensen, Michal Kopytynski, Berardo Mario Manzi, Enrica Mazzon, Suhuur Osman, Oliver Robotham and Shiqi Wang.

Last but not least, I am really grateful to my family and friends for their crucial support. A special mention to the "Romeo's girls", Vale, Sonja, Andrea and Yu-Han, my incredible sisters, who are my main strength and motivation to do better, my best friend Elena who is thought of as a sister, and my aunty Simonetta Battistoni because she always understands my needs and makes me feel confident.

Dedication

To my family who always supports and motivates me

To the next generations of scientists, hope this work would inspire you

"Just like you can't judge a book by the cover
We all gotta be careful
How we treat one another
Skin Deep
Skin Deep
Underneath we're all the same"

Buddy Guy, "Skin Deep"

Contents

Declaration **3**

Abstract **5**

Acknowledgements **7**

List of Tables **18**

List of Figures **25**

List of Abbreviations **28**

1 Introduction **29**

1.1 Motivation 30

1.2 Current status of the field 34

1.2.1 Stratum Corneum models 34

1.2.2 The unique lipids at the skin surface 35

1.2.3 Review of the current state of SC simulations and gaps in the literature . 42

1.3 Thesis outline 44

1.4 My publications 46

Bibliography to chapter 1 47

2 Background Theory and Methods **53**

2.1 The basics of molecular dynamics 54

2.1.1 Thermostatting and barostatting 57

2.2 Forcefields for biomolecular simulations 60

2.3 Extracting information from MD simulations 65

2.3.1 Free energy difference between two states 66

2.3.2 Enhanced conformational sampling 67

2.3.3 Simulating rare events 69

2.4 MD simulation *vs.* experiment 71

2.4.1 NMR and MD: structure and dynamics 72

2.4.2 Structural features from diffraction – solving the phase problem with MD 74

2.5 Data analysis 75

2.5.1 Lipid properties 75

2.5.2 Peptide properties 80

Bibliography to chapter 2 82

3 On the top of our skin: the sebum oily layer **88**

3.1 Introduction 89

3.2 Methods 92

3.2.1 Sebum model: structure and forcefield 92

3.2.2 Simulation details 93

3.2.3 Data analysis 97

3.3 Results 101

3.3.1 TG-bulk 101

3.3.2 TG-air . . . 107

3.3.3 TG-water . . . 110

3.3.4 TG monolayer . . . 114

3.4 Impact of the charge distribution of the GLY group . . . 120

3.5 Conclusion and Future Work . . . 122

Bibliography to chapter 3 . . . 124

4 How the sebum oil interacts with SC lipids 129

4.1 Introduction . . . 130

4.2 Methods . . . 132

4.2.1 Input configurations and simulation details . . . 133

4.2.2 Multicomponent bilayer analysis . . . 134

4.2.3 Free energy pathway of lipid extractions . . . 138

4.3 Results . . . 139

4.3.1 Mechanisms of interdigitation of TG and SC . . . 139

4.3.2 Bilayer Stability . . . 148

4.4 Conclusions . . . 157

4.5 CPP5 Outreach . . . 157

Bibliography to chapter 4 . . . 159

5 How wettable is the skin surface? 162

5.1 Introduction . . . 163

5.2 Methods . . . 165

5.2.1 Input configurations and simulations details . . . 165

5.2.2 Contact angle computations . . . 166

5.3 Results . . . 170

5.3.1 Characterising the water adsorption . . . 170

5.3.2 Structure of the SC bilayer . . . 174

5.3.3 Comparison between SC and DPPC . . . 177

5.3.4 The wettability of the SC-TG interface . . . 180

5.4 Conclusions and future work . . . 183

Bibliography to chapter 5 . . . 185

6 Nanoparticle interfaces **187**

6.1 Introduction . . . 188

6.2 Methods . . . 190

6.2.1 Input configuration: nanoparticle model . . . 190

6.2.2 Non-Equilibrium Molecular Dynamics . . . 192

6.2.3 Data analysis . . . 193

6.3 Results . . . 200

6.3.1 Temperature profiles around particle . . . 200

6.3.2 Thermal conductance - dependence on **F** . . . 203

6.3.3 Thermal conductance - dependence on curvature and coupling effect . . . 209

6.4 Conclusions . . . 213

6.5 Future work: nanoparticles interacting with SC-lipid bilayers . . . 215

Bibliography to chapter 6 . . . 215

7 Antimicrobial peptides **220**

7.1 Introduction . . . 221

7.2 Methods . . . 223

7.2.1 G4 in aqueous solutions . . . 223
7.2.2 G4 interaction with DPPC and DPPG monolayers . . . 226
7.2.3 Ramachandran plot . . . 228
7.3 G4 in water . . . 230
7.3.1 Acceleration techniques . . . 234
7.4 G4 interacting with lipid membranes - Lipid charge effect . . . 238
7.5 Conclusion and future work . . . 242
Bibliography to chapter 7 . . . 243

8 Fats' split personality **247**

8.1 Introduction . . . 248
8.2 Activity preparation . . . 249
8.2.1 Materials . . . 249
8.2.2 Preparation before the activity . . . 250
8.3 The activity . . . 251
8.4 Computer simulations for an insight at the molecular level . . . 254
8.4.1 Method . . . 254
8.4.2 Results . . . 257
8.5 Discussion . . . 258
8.6 Conclusion . . . 262
Bibliography to chapter 8 . . . 263

9 Conclusion **265**

9.1 Summary of thesis conclusions . . . 266
9.1.1 Research outcomes . . . 266

9.1.2 Science communication 269

9.2 Future challenges 269

Bibliography to chapter 9 271

List of Tables

1.1 Sebum composition. 41

3.1 Main simulation parameters for the different input configurations. 94

3.2 Probability of finding different shapes for molecules in bulk-TG. 107

3.3 Different surface tension values for a fatty acid and three triglyceride structures. 109

3.4 Probability of finding different shapes for interfacial TG in TG-water configurations. 112

3.5 Probability of finding different shapes for bulk TG in TG-water configurations. . 112

4.1 Mole fraction of SC lipids (CHO and FFA) extracted and dispersed in the TG phase. 140

4.2 Mean value and standard deviation of A_L measurements for the TG-SC system. 151

4.3 Number of hydrogen bonds for SC lipid type 152

5.1 Mean value and standard deviation of A_L measurements for the TG-SC-water system. 175

5.2 Contact angles described by the two spherical caps used to characterise the water droplet. 180

6.1 Number of bins (n_{BIN}) corresponding to different particle sizes, R^*. 194

6.2 Thermal conductances and thermal conductivities of the nanoparticle and liquid as a function of the interaction strength. 202

6.3 Dependence of the thermal conductance with the nanoparticle-fluid interaction strength 211

7.1 List of peptide simulations 225

7.2 Probabilities of G4 secondary structures 233

7.3 Comparison of parameters between standard and modified solvent masses 235

7.4 Probabilities of G4 secondary structures when interacting with DPPC or DPPG lipids 241

8.1 List of events 258

List of Figures

1.1 Title page figure for chapter 1. 29

1.2 a) Skin layers and appendages. b) Water content across the epidermis 30

1.3 Transdermal delivery routes 31

1.4 ITN-SNAL poster 32

1.5 Schematic representation of the evolution of SC models over time. 36

1.6 Ceramide structures 37

1.7 Models for the SC intercellular lipid matrix 39

1.8 Graphical abstract of the content of chapters 3, 4 and 5 compared to the current literature 45

2.1 Title page figure for chapter 2. 53

2.2 Berni Alder, Mary Ann Mansigh and Tom Wainwright. 54

2.3 MD scheme and periodic boundary conditions 57

2.4 Forcefield terms definition and use 61

2.5 REMD methodology 68

2.6 Umbrella sampling methodology 70

2.7 Examples of Voronoi tessellation in 2D and 3D. 77

2.8 a) Pressure tensor components as a function of the Coulombic cutoff radius. b) Ceramide 2 pressure profile 80

2.9 Backbone dihedral angles definition and Ramachandran plot 81

3.1 Title page figure for chapter 3. 88

3.2 Structure of the sebum triglyceride with atom types and partial charge distribution of the glycerol group. 93

3.3 Input configurations 94

3.4 Left) Vectors describing the orientations of the TG acyl chains. Right) Four classes of TG shape. 97

3.5 Subdivision of the angles phase space into four TG conformations. 99

3.6 Density of TG *vs.* time. 102

3.7 RDF of the central carbon atom of the glycerol group. 103

3.8 a) TG nearest neighbours. b) Histograms of TG acyl chain lengths 104

3.9 Number of TG head group clusters *vs.* time 105

3.10 Analysis of the intramolecular TG-conformations. 106

3.11 3D distributions of TG conformations overlapped to TG shape regions. 107

3.12 Results for the TG-air interface (left) and TG-water interface (right) 108

3.13 Analysis of the intra-molecular TG-conformations for bulk TG (a) and interfacial TG (b) at the TG-water interface. 111

3.14 3D distributions of TG conformations (interfacial on the left and bulk on the right) overlapped to TG shape regions. 112

3.15 Top) Density profiles and free energy of water insertion in the TG-phase. Bottom) water molecules distribution within the system 113

3.16 Computed surface pressure-area isotherm at $T = 310\,\mathrm{K}$. 115

3.17 Lateral RDF for different values of surface pressure 117

3.18 TG monolayer - hysteresis effect. 118

3.19 Area per lipid of TG-monolayers *vs.* time 119

3.20 Comparison between the c2 RDFs computed using FF1 and FF2 set of charges. 120

3.21 Comparison of the interaction between TG and water for FF1 (left) and FF2 (right) set of charges. 121

4.1 Title page figure for chapter 4. 129

4.2 a) Schematic representation of the top of the skin surface. b) Chemical structure of TG, CR2, CHO and FFA. c) System used. 132

4.3 Construction of the input configuration using Packmol. 133

4.4 CR2, CHO and FFA molecules with labels. 135

4.5 a) Histogram of lipid subgroups. b) 2D Voronoi tesselation method on a bilayer slice. 136

4.6 Assignment of TG and CR2 atoms to groups. 137

4.7 Representative snapshots illustrating the main mechanisms of lipid interdigitation. 139

4.8 a) Number of TG molecules anchored to the bilayer *vs.* time. b) 3D distribution of TG conformations. 141

4.9 a) Number of CR2 splitting *vs.* time. b) Examples of C24 splitting dynamics. . . 143

4.10 a) Distribution of CHO head groups and free energy profile. b) Representative snapshot showing TG clusters and CHO molecules. 144

4.11 a) CHO molecule with labels. b) 2D Free energy landscape for CHO extraction as a function of θ_T and z. 145

4.12 FFA with the chain tilted in the head-group region. 146

4.13 a) Distribution of FFA head groups and free energy profile. b) Representative snapshot showing TG clusters and FFA molecules. 147

4.14 a) FFA molecule with labels. b) 2D Free energy landscape for FFA extraction as a function of θ_T and z. 148

4.15 Density profile of TG and SC lipids. 149

4.16 Percentage of lipid remaining in the bilayer and number of extraction event (in the inset) *vs.* time 150

4.17 Area per lipid *vs.* time for SC and sebaceus lipids for the TG-SC system. 151

4.18 Number of hydrogen bonding per lipid *vs.* time (left) and individual lipid contribution (right). 153

4.19 Order parameter of the lipid tails *vs.* atom index. 155

4.20 Lateral pressure profile for the TG-SC system. 156

4.21 Text and image entry for CCP5 Outreach library and contest. 158

5.1 Title page figure for chapter 5 162

5.2 a) Water content within the layers of the epidermis. b) Chemical structures of TG, CR2, CHO and FFA. c) Initial configuration. 165

5.3 Liquid droplet above a rigid (left) and a liquid (right) substrates. 167

5.4 Left) Density profile of the water lens in cylindrical coordinates. Right) Zoom on the two spherical caps describing the shape of the water lens. 169

5.5 Volume of the water droplet (top) and number of water molecules at the lower TG-SC interface (bottom) *vs.* time. 171

5.6 TG clusters over two representative snapshots from the simulations at 1 µs (left) and 5 µs (right). 173

5.7 a) Percentage of lipids remaining in the bilayer and the number of extraction events (in inset) *vs.* time for the TG-SC-water system. b) Number of anchored TG molecules *vs.* time. 174

5.8 Properties of the SC lipid bilayer. 176

5.9 Footprint of the water droplet on the bilayer surface *vs.* time (on a logarithmic scale) for the SC-TG (top) and TG-DPPC (bottom) interfaces. 178

5.10 Representative snapshots from the simulations and density profiles in cylindrical coordinates for TG-SC and TG-DPPC systems. 179

5.11 Surface tensions and angles relationships represented with the Neumann's triangle construction. 181

6.1 Title page figure for chapter 6. 187

6.2 Snapshot of a representative nanoparticle configuration. 192

6.3 Radial heat flux (a) and temperature profile (b) for a nanoparticle of radius = 5. 195

6.4 Left) Average density profile of the nanoparticle and the liquid-vapor interface. Right) Dependence of the nanoparticle contact angle on the interaction strength. 200

6.5 Temperature profiles (a) and thermal conductivities (b) as a function of the interaction strength. 201

6.6 Impact of the cutoff radius on the temperature profiles. 203

6.7 Left) Dependence of the thermal conductance on the interaction strength (left) and on the free energy of adsorption (right). 205

6.8 a) Density profiles of the solvent around a nanoparticle of radius $R^* = 5$. b) Density of states. 206

6.9 Thermal conductance of the nanoparticles as a function of the particle curvature and interactions strength. b) Density profiles of the fluid around nanoparticles of different radii. 209

6.10 Dependence of the parameter δ on the thermal conductance for nanoparticles of radius 2 nm, this work *vs.* the literature. 213

6.11 Initial configuration for the study of the interaction of nanoparticles and SC lipids 216

7.1 Title page figure for chapter 7. 220

7.2 Antimicrobial resistance in EU 221

7.3 Peptide G4 initial configurations 224

7.4 DPPC/DPPG mutation 227

7.5 Ramachandran plots 229

7.6 RMSD, Radius of gyration, end to end distance and hydrogen bonds *vs.* time . . 231

7.7 Ramachandran diagrams for G4 peptides. 233

7.8 Secondary structure probability for G4 residues 234

7.9 G4 conformations 235

7.10 RMSD, Radius of gyration, end to end distance and hydrogen bonds *vs.* time with modified solvent masses 236

7.11 Secondary structure probability for G4 residues with modified solvent masses . . 237

7.12 (Left) REMD potential histogram overlapping, (Right) replica index *vs.* time . . 237

7.13 REMD: Left) Ramachandran plot, Right) Probability of secondary structure for the different G4 resides 238

7.14 (a) RDF (b) Area per lipids *vs.* time. 239

7.15 Simulation snapshots for G4 interacting with DPPC and DPPG 240

7.16 Ramachandran plots for G4 interacting with DPPC and DPPG monolayer . . . 241

7.17 Secondary structure probability for G4 residues when interacting with DPPC and DPPG monolayer 242

8.1 Title page figure for chapter 8. 247

8.2 Different set-ups of the stand for the outreach activity 250

8.3 Informative material to introduce the "Fats' split personality" and its applications to the general public 252

8.4 The 4-steps of the hands-on experiment 253

8.5 Flyer with experimental instruction 254

8.6 Representations of the system components in the MARTINI forcefield: soap, water and milk triglycerides. 256

8.7 Three representative snapshots from the MD trajectory showing the interaction between milk and soap lipids 257

8.8 General public's engagement with the outreach activity 260

8.9 Feedback of the general public 260

8.10 Image competition 261

8.11 Outreach activity at Platanos College 262

9.1 Title page figure for chapter 9. Word cloud realised using the words in this chapter. 265

List of Abbreviations

AA	All Atom
AMP	Antimicrobial Peptide
CG	Coarse Grained
CHO	Cholesterol
CR2	Ceramide 2
DPPC	1,2-dipalmitoyl-sn-glycero-3-phosphocholine
DPPG	1,2-dipalmitoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
FFA	Free Fatty Acid: Lignoceric acid C24:0
G4	Peptide $G(IIKK)_4-I-NH_2$
MC	Monte Carlo
MD	Molecular Dynamics
$N\gamma T$	Constant number of particles, surface tension and temperature
NP	Nanoparticle
NPT	Constant number of particles, pressure and temperature
NVT	Constant number of particles, volume and temperature
RDF	Radial distribution function
REMD	Replica Exchange Molecular Dynamics
RP	Ramachandran Plot
SC	Stratum Corneum
T_m	Phase transition temperature
TG	Tri-cis-6-hexadecenoin
UA	United Atom

Chapter 1

Introduction

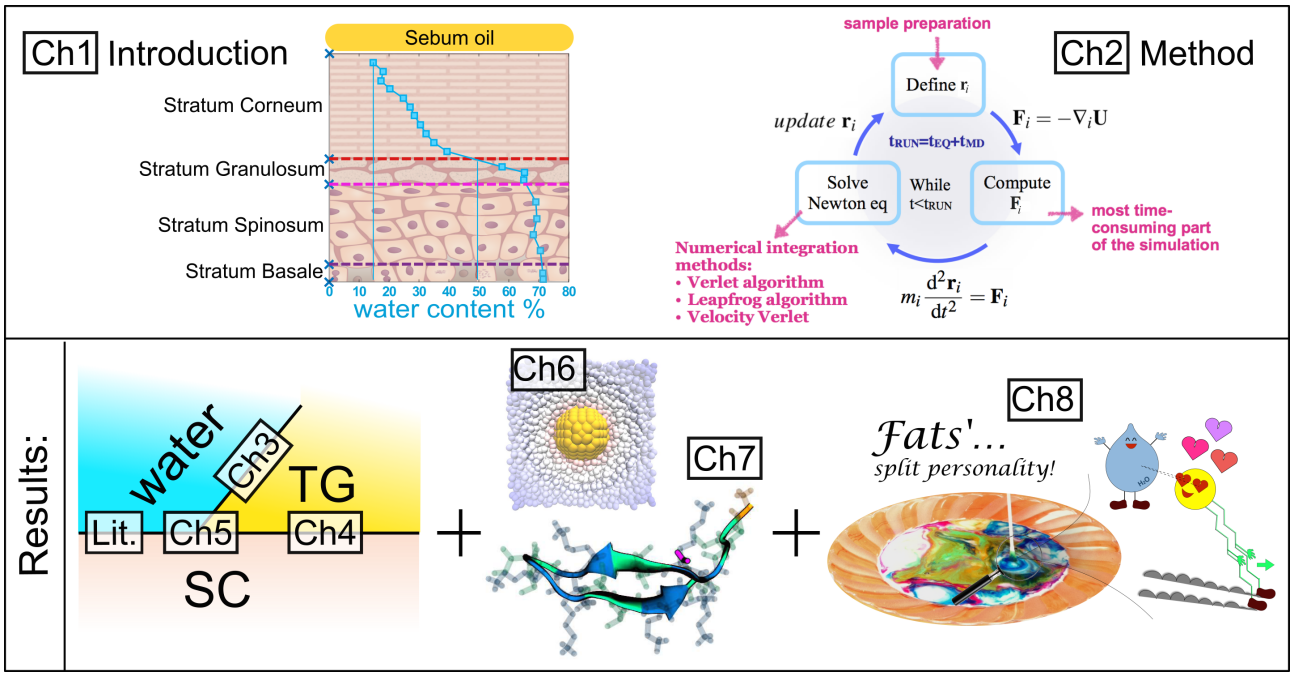
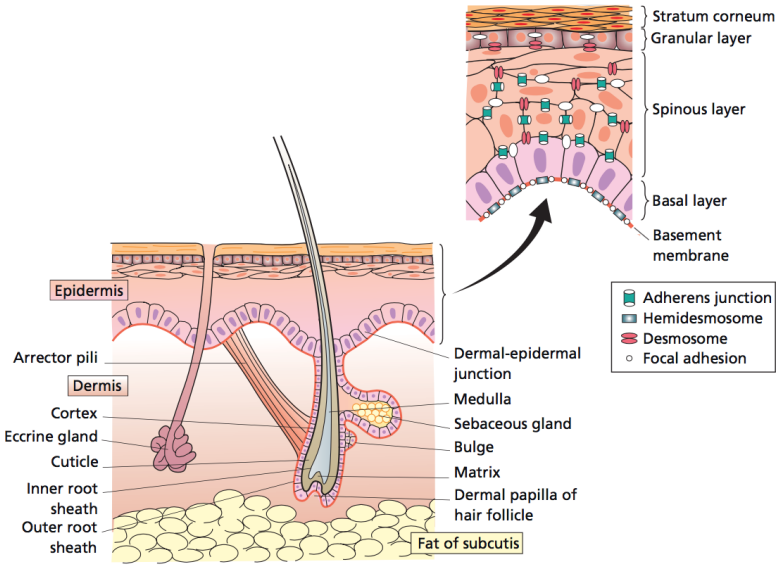


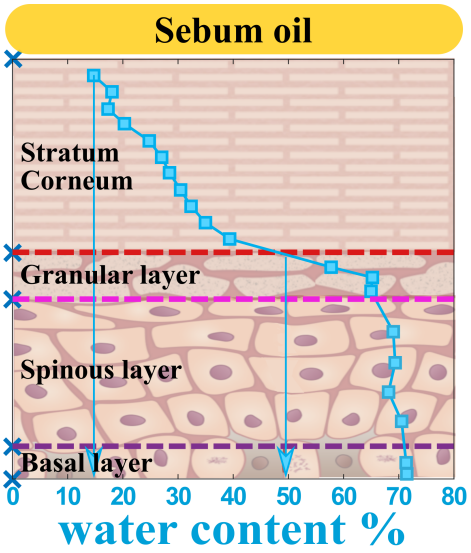
Figure 1.1: Title page figure for chapter 1. The imagines presented in this figure summarise the contents of the chapters of this thesis. "Ch" stands for chapter, "Lit." for Literature.

1.1 Motivation

The skin is the largest organ of the human body and anatomically one of the most heterogeneous. It represents the first barrier against the external environment, shielding us against the invasion of foreign pathogens and regulating water loss from the body[1]. The skin has a rather elaborate structure, organised in several layers (see Figure 1.2a), the outermost of which is called the stratum corneum (SC). This layer has a "brick and mortar" like structure, with skin cell "bricks" embedded in a lipid "mortar". It is the main layer responsible for the low permeability and barrier function of the skin[2], and is especially significant in regulating hydration and avoiding water loss for the body. This function results in a very steep water concentration gradient across this layer (from 50% to 15% grams of water per total grams of wet tissue[3] - as displayed in Figure 1.2b), and a very low water concentration at the skin surface.



(a) Skin layers and appendages.



(b) Water % across the epidermis

Figure 1.2: (a) Skin layers: moving in from the external surface, there is the SC (∼ 10–20 µm), the viable epidermis (∼ 50–100 µm, subdivided into granular layer (stratum granolosum), Spinous layer (stratum spinosum) and basal layer (stratum basale)) and the dermis (∼ 1–2 mm thick). This picture was originally published in J. A. McGrath, R. A. J. Eady and F. M. Pope; Chapter 3: Anatomy and Organization of Human Skin; in *Rook's Textbook of Dermatology*, Volume 1, Eighth Edition Edited by T. Burns, S. Breathnach, N. Cox and C. Griffiths © 2010 Blackwell Publishing Ltd. ISBN: 978-1-405-16169-5[4]. (b) The layers of the epidermis and the water content within them (the data for the water content are taken from ref. 3).

The surface of the SC is covered by an oily layer, called sebum. The presence of this layer is responsible for the oily touch of certain parts of the body, which are rich in sebum, such as

the forehead. The sebum oil has often been associated with the onset of the disease acne[5, 6]. However, in more recent years, it has also been shown to contribute to the regulation of the hydration and waterproofing properties of the skin[7], confer antimicrobial properties[8] to the skin and assist in the transport of antioxidants, such as vitamin E[9, 10]. The effect of the sebum on the underlying SC has been the subject of a few studies[11, 12], which have found evidence of stratum coneum lipids being present in the sebum oil and alteration of the SC properties due to the sebum. Despite these promising studies, a detailed understanding of the roles and functions of the sebum, and its interactions with water and SC lipids at a molecular level is still missing, and represents one of the primary objectives of this thesis.

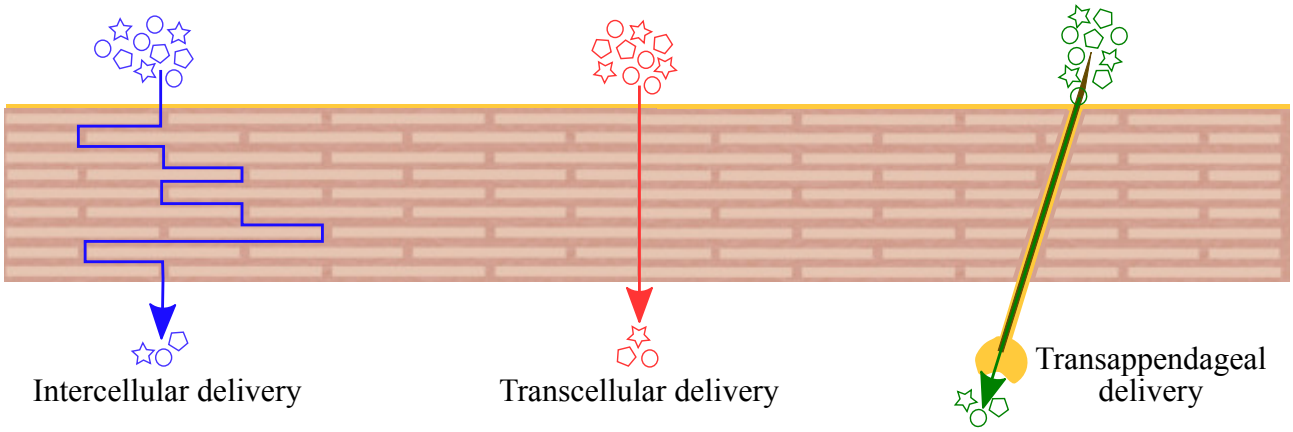


Figure 1.3: There are three pathways to penetrate the SC: intercellular devilery (through the skin lipid mixture), transcellular delivery (crossing the SC cells), transappendageal delivery (via the hair follicles). The SC is represented by the "brick and mortar" model, with corneocyte cells (light pink) surrounded by the intercellular lipid mixture (dark pink). The sebum (yellow) and a hair follicle penetrating the SC layer is also shown.

As the most external layers of our body, the sebum and the SC are the first line of defence against toxic particles entering the body, but also the first barrier to overcome when cosmetic and therapeutic creams and ointments are applied to the skin surface. As shown in Figure 1.3, there are three pathways to penetrate the skin barrier[13]: transcellular delivery (crossing the SC cells), intercellular delivery (through the skin lipid mixture, the “mortar”), and via the hair follicles (which are surrounded by sebum[14]). In each of these routes external agents have to interact and pass through either sebum or SC or both. A better understanding of the properties of these layers will open the doors to designing better drug delivery procedures and to further studies of interactions of drug molecules and nano-objects with the skin surface.

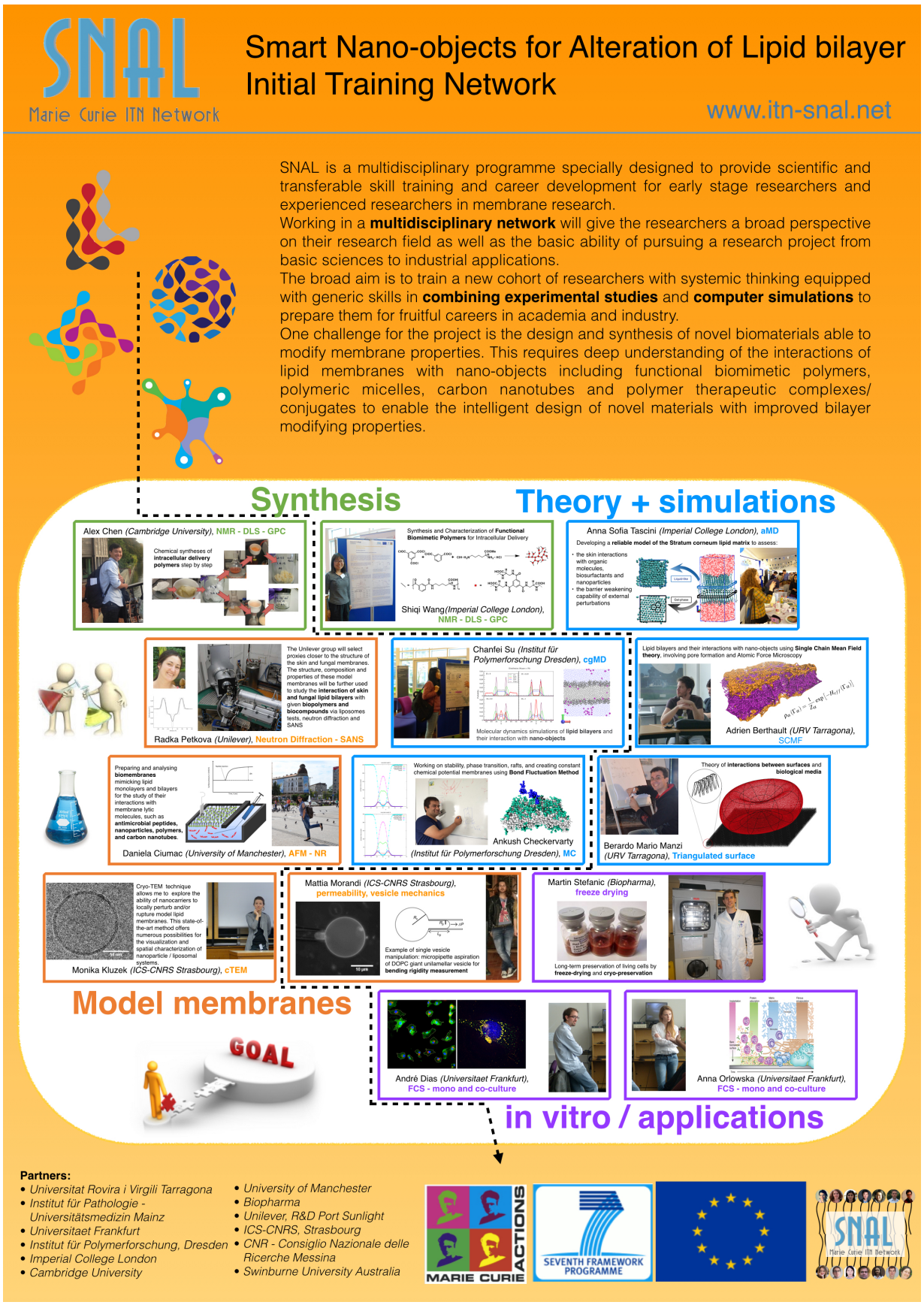


Figure 1.4: ITN-SNAL poster. I designed this poster in collaboration with Shiqi Wang to promote the goals of the network at international conferences and outreach events.

Nanomaterials are currently being used in biomedicine, cosmetics, chemical coatings, electronic devices and as drug carriers, diagnostic imaging tools and catalysts[15, 16, 17]. The promising technological applications of nano-objects and nanomaterials has led to a striking increase in the interest of the scientific community towards a better understanding of their properties and potential uses. An example of this is the substantial support from the European Commission in promoting research in nanoscience, nanotechnology and nanomedicine through its Framework Programmes for Research and Development (FP7) and the current Horizon 2020 programme. A total of 116 projects on nanotechnologies for healthcare applications were funded within the FP7 framework (corresponding to a total EU financial contribution of $\sim$ 550M €)[18].

The EU-funded international training network, SNAL[19] (Smart Nano-objects for Alteration of Lipid bilayers), which funded my PhD project, had as its main challenge the design and synthesis of novel biomaterials able to modify membrane properties. This aim was carried out by exploiting a multidisciplinary approach, gaining a deep understanding of how biological membranes work and interact with nano-objects, including functional biomimetic polymers, antimicrobial peptides, polymeric micelles, carbon nanotubes and therapeutic polymer complexes/conjugates. Alongside developing high-quality research, a main objective of SNAL was to train a new cohort of researchers equipped with generic skills in combining experimental studies and computer simulations to prepare them for fruitful careers in academia and industry. The poster used to introduce SNAL and promote the action of the network at outreach events and conferences is presented in Figure 1.4. As shown in the poster, my role, within the network, has been to use computational methods, specifically molecular dynamics (MD) simulations, to obtain a molecular insight into the barrier function of sebaceous and epidermal lipids and the interactions of lipid membranes with antimicrobial peptides and therapeutic nanoparticles.

Molecular dynamics simulation methods have evolved significantly in the last fifty years reaching a high level of sophistication that now permits the investigation of complex biological molecules and processes. MD is now a discipline in its own right, that both complements experimental measurement and acts as a unique predictive tool, providing much needed insight at molecular scales, where experiments are not always possible or are difficult. The development of multiscale

models, which was recognised with the recent award of the 2013 Nobel Prize in Chemistry, has contributed to the increasing use of MD simulations in biology[20].

In this thesis, the current understanding of the properties of the skin surface is expanded by introducing a model of the sebum layer and studying its interactions with water and SC lipids. This model, is then used to address the question: "how wettable is the skin surface?". In addition, the thermal properties of nanoparticle interfaces are characterised, as well as the conformations of an antimicrobial peptide. Alongside scientific research, science communication is vital to address misinformation, inspire new generations towards STEM (Science, technology, engineering, and mathematics), inform the general public and impact decision making processes, guiding political and ethical thinking. The last chapter of this thesis describes an outreach activity I have designed during my PhD to promote a deeper understanding among the general public of the structures and vital functions of lipid molecules and how their study can help in designing new opportunities for drug delivery.

1.2 Current status of the field

1.2.1 Stratum Corneum models

According to the current understanding of the SC, it exhibits a dynamic structure and acts as a metabolically-active bio-sensor, able to quickly adapt to different climates and hygrometric conditions, mechanical stresses, pH variations and psychological states. The physical behaviour of the SC is the result of a complex interplay between these stimuli.

However, $\sim$ 50 years ago, due to artefacts in the imaging techniques, microscopy studies of the SC described it as a static "basket weave" structure. In this model, a specific barrier function was not attributed to the SC, but was instead believed to be provided by lower layers of the epidermis (*e.g.* the spinous layer). The breakthrough studies of Kligman and co-workers[21, 22], in 1964, showed that the SC layer was a "coherent, plastic membrane, which serves as the barrier against inward and outward movement of all substances including water", and propa-

gated the concept of the the SC as a single uniform membrane, which became known as the "plastic wrap" analogy. This model was later replaced with the two-compartment "bricks and mortar" structure, introduced by Elias and co-workers[23]. This model is still currently adopted as an analogy of the complex SC organisation. It includes two fundamental components: the corneocytes (the "bricks"), *i.e.* the pancake-like dead cells typical of this layer, which are surrounded by a lipid matrix (the "mortar") composed of more than three hundred different lipid components. However, the simplicity of this model can not capture some critical features of the SC. For example (i) the role of the corneodesmosomes, *i.e.* a class of proteins which serve as junctions to hold the corneocytes together, in regulating the desquamation process[24]; (ii) the heterogeneity of the corneocytes structure and, in particular, the role of the lipid monolayer which surrounds them known as corneocyte-bound lipid envelope[25]; (iii) the existence of localised spaces within the intercellular region which might contain water at high hydration level[26]. Figure 1.5 summarises the evolution of SC models over time, highlighting the main breakthrough discoveries and studies and presenting schematic representations of the main skin models.

1.2.2 The unique lipids at the skin surface

SC intercellular lipids

Lipid molecules are a critical component of all cells. Thanks to their amphiphilic nature, they can assemble into organised structures and provide cells and organelles with thin boundaries, *i.e.* lipid membranes, which protect the cell or organelle from its external environment and enable bidirectional selective permeation of molecules. Lipids also provide compartments where protein and enzymatic sequences can organise, and where specific chemical environments can be maintained, for example, a means to bring enzymes to the conformation necessary for them to carry out their functions. The most abundant class of lipids in mammalian cells are phospholipids and sterols.

The composition of the SC lipid mixture is very different to most other cell membrane compo-

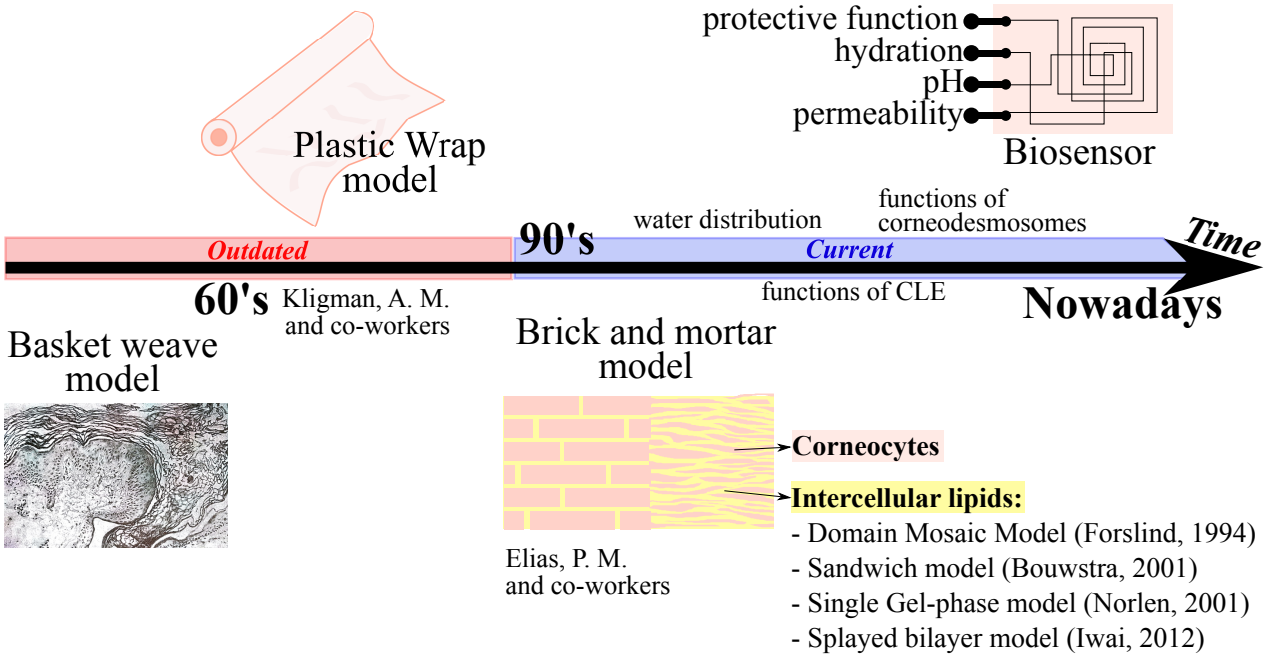
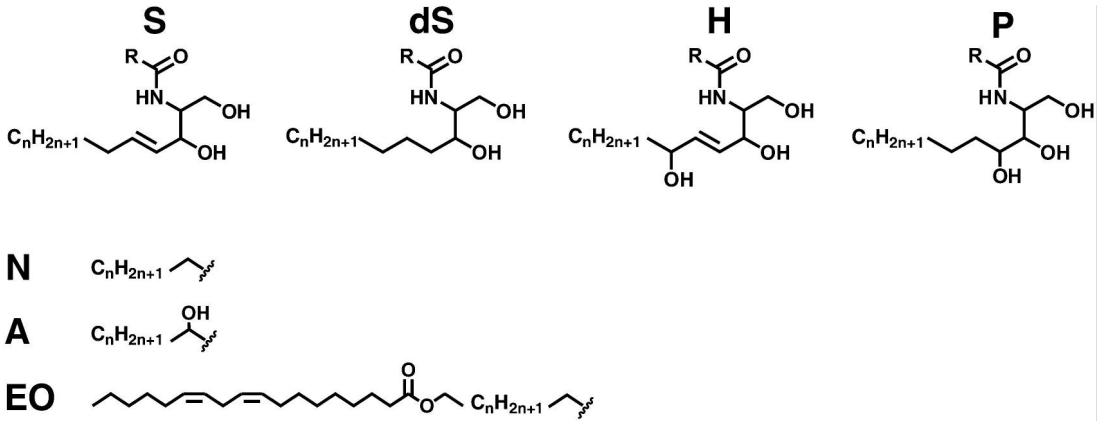


Figure 1.5: Schematic representation of the evolution of SC models over time. The picture of the basket-weave model was originally published in ref. 27, and shows the SC as a loose criss-cross pattern. According to the plastic wrap model, the SC is imagined as a uniform plastic membrane. Both these two models are considered obsolete. The importance of corneocyte cells and intercellular lipids is underlined by the SC brick and mortar model. The organisation of the lipids in the intercellular space is itself the subject of several models, listed in the picture. Currently the SC is thought of as a biosensor which responds to several inputs (*e.g.* hydration level and pH) in a complex way to perform the SC functions (*e.g.* providing a barrier and controlling permeability). On the time axis, along with the models, some of the breakthrough discoveries and challenges are also reported.

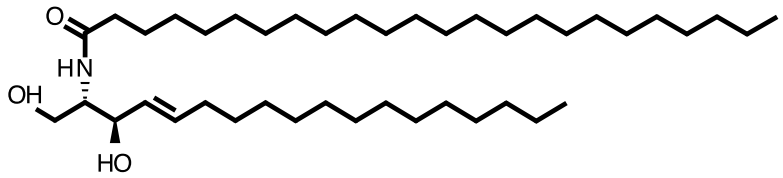
sitions and also to the lipid population of the viable epidermis, the layers below the SC, where the phospholipid composition is much higher. The principal components of the SC are, instead, a mixture of ceramides, cholesterol and long chain free fatty acids[28, 29].

The ceramides' structure is made up of an amide-linked fatty acid and a long-chain amino alcohol (*i.e.* sphingosine backbone). The molecular heterogeneity of ceramides is due to either the chain length or the hydroxylation of the fatty acid motif and the sphingoid base. This heterogeneity results in 11 different ceramide subclasses (see Figure 1.6a) and more than 300 different ceramide species present in the SC lipid mixture[30, 31]. Despite this heterogeneity, ceramide 2 (also known as ceramide NS) is the most abundant, accounting for $\sim$ 30% of the total ceramide composition, as measured by high-performance thin-layer chromatography on human skin[32]. Ceramide 2 captures two critical features of skin ceramides, namely low head group polarity and strong asymmetry in the length of the two tails. It has often been used as

a model of SC ceramides in simulation studies of skin lipids (*e.g.* see ref. 33, 34, 35, 36, 37, 38) and it is also used in this thesis as a representative SC ceramide (see chapters 4 and 5). Its structure is shown in Figure 1.6b



(a) Ceramide structures



(b) Ceramide NS

Figure 1.6: (a) Ceramide structures. Four possible sphingosine related chains (S, sphingosine; dS, dihydrosphingosine; H, 6-hydroxy-sphingosine; P, phyto-sphingosine) are linked via an amide bond (at position R) to one of three different fatty acid components (N, nonhydroxy fatty acid; A, α-hydroxy fatty acid; EO, esterified ω-hydroxy fatty acid) resulting in, theoretically, 12 different ceramide subclasses. The literature has reported the presence of all but EOdS in human SC. This picture was originally published in J. van Smeden, L. Hoppel, R. van der Heijden, T. Hankemeier, R. J. Vreeken and J. A. Bouwstra. LC-MS analysis of SC lipids: ceramide profiling and discovery, *Journal of Lipid Research*, 2011; 52,1211–1221. © the American Society for Biochemistry and Molecular Biology[30]. (b) Ceramide 2 (or ceramide NS) structure.

The role of ceramide molecules in the SC mixture is vital to the barrier function of the skin, due to their tight lateral packing and capability of forming hydrogen bonds[33, 39]. Ceramide rich domains are highly ordered and feature a very high gel-fluid phase transition temperature ($\sim 70\,^{\circ}\mathrm{C}$)[40]. These properties result in reduced lateral diffusion and confer a very low permeability to the SC lipid matrix. A lack of ceramide in SC lipid mixtures has been associated with skin diseases, such as atopic dermatitis, psoriasis and ichthyosis, and the application of ceramide to the skin has been studied for symptoms mitigation, with excellent results in barrier regeneration[41]. However, the studies of Elias and co-workers[42, 43] suggest that a mixture of ceramides, cholesterol and free fatty acids is more efficient in recovering the barrier function of

the SC, providing evidence of the vital roles of cholesterol and free fatty acids for the proper functioning of the SC.

Infrared spectroscopy studies on model binary mixtures of SC lipids have helped to explain the roles of the different components: binary mixtures of ceramide 2 and cholesterol show a looser chain packing, while binary mixtures of ceramide and free fatty acids appear very brittle and rigid[44]. Mojumdar *et al.*[45] and Norlén[46] demonstrated the importance of cholesterol in the formation of dense lipid organisation (*e.g.* formation of a lamellar orthorhombic phase in multicomponent SC model membranes - see ref. 45), in broadening the phase transition temperature range and in promoting lamellar structures and stabilising the gel phase (see ref. 46). Free fatty acids have been shown to increase the density of the hydrocarbon chain packing, but also disrupt ceramide head group packing by interrupting their hydrogen bond networks[39]. In general, these studies highlight the complexity of understanding the roles of SC lipid components and the substantial differences from other biological membranes mainly composed of phospholipids.

A clear understanding of the arrangement of the lipids within the SC lipid matrix is still an open question. There are many factors which make the intercellular SC characterisation challenging, such as (i) the lack of precise information on the presence or absence of water between lipid lamellae, (ii) the possible existence of different chain packings with different tilt angles within stacks of lamellar structures, (iii) the capability of ceramide molecules to assume hairpin (*i.e.* both chains pointing in the same direction) or splayed (*i.e.* chains pointing in opposite directions) conformations, and (iv) the capability of long acyl chains to interdigitate in the space between lipid leaflets. Over the years, four main models have been proposed (listed in Figure 1.5 and illustrated in Figure 1.7):

- The mosaic model proposed by Forslind[50], in which the SC lipids form segregated crystalline domains surrounded by fluid-crystalline phase borders. Phase separations explain the response of the skin to mechanical stress and allow controlled diffusion of water, lipid and hydrophobic molecules through the fluid borders. The barrier function, against water loss and external pathogens, is instead carried out by the crystalline domains[51].

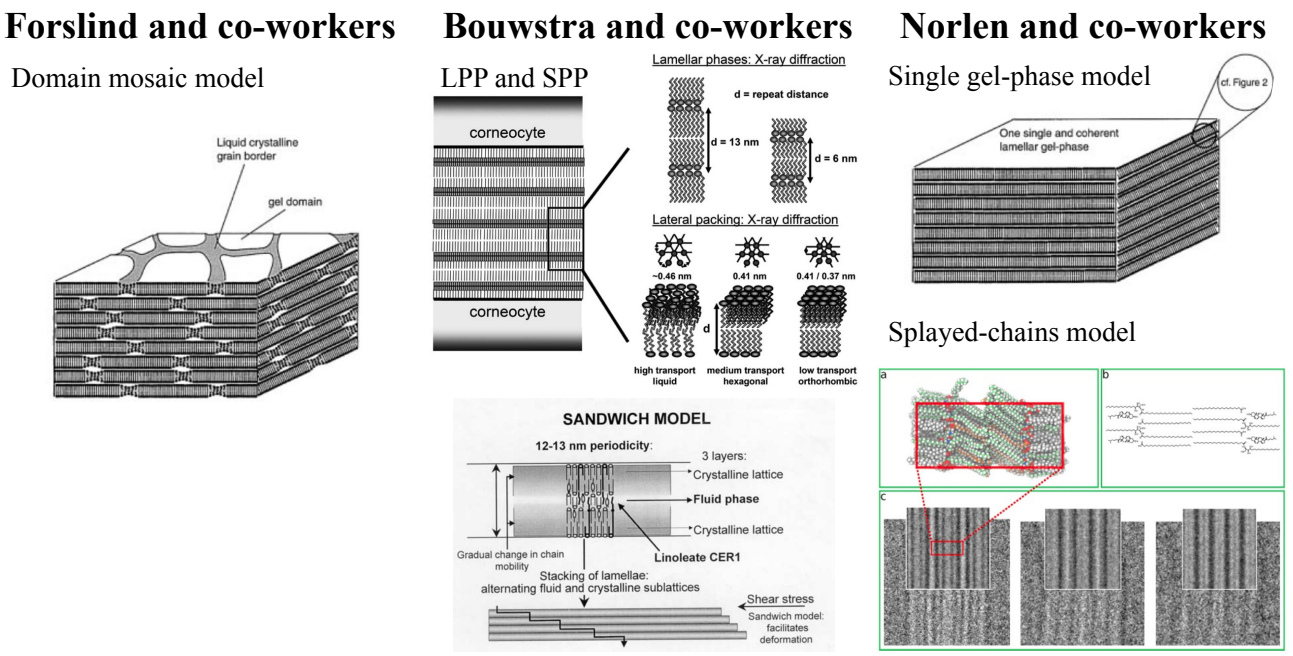


Figure 1.7: Models for the SC intercellular lipid matrix. The figures for the domain mosaic model and the single gel phase model were originally published in L. Norlén, Skin Barrier Structure and Function: The Single Gel Phase Model, *Journal of Investigative Dermatology*, Elsevier, 2001, Copyright © 2001 The Society for Investigative Dermatology, Inc. All rights reserved[46]. The figure showing the long periodicity and short periodicity phases (LPP and SPP) was originally published in J. A. Bouwstra and G. S. Gooris, The Lipid Organisation in Human SC and Model Systems, *The Open Dermatology Journal*, 2010, 4(2), 10–13.[47], under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0). The figure showing the sandwich model phase is reprinted by permission from Springer Nature and Copyright Clearance Center: Springer Nature, *Journal of Biological Physics*, The Lipid Organisation of the Skin Barrier: Liquid and Crystalline Domains Coexist in Lamellar Phases, J. A. Bouwstra, G. S. Gooris, M. Ponec, Copyright © 2002, Kluwer Academic Publishers (2002)[48]. The figure showing the splayed bilayer model was originally published in M. Lundborg *et al.*, Human skin barrier structure and function analyzed by cryo-EM and molecular dynamics simulation, *Journal of Structural Biology*, 2018, `https://doi.org/10.1016/j.jsb.2018.04.005`[49], under the terms of the Creative Commons Attribution-NonCommercial-No Derivatives License (CC BY NC ND).

- Bouwstra and co-workers suggested that the SC lipids, between corneocytes, are organised in two lamellar phases with repeat distances of 6 nm (short periodicity phase, SPP) and 13 nm (long periodicity phase, LPP). The SC lipids are organised in either a liquid phase, a hexagonal phase or an orthorhombic lateral packing[47]. They proposed an alternative distribution of fluid-crystalline and gel-phase domains, known as the sandwich model[48, 52]. According to this model, the liquid domains are located in a narrow central part of the repeating long periodicity phase (LPP) unit. To cross the SC barrier via the intercellular lipid phase substances have, therefore, to permeate alternatively through two densely packed layers and one less densely packed layer.

- In contrast to the previous models, Norlén[46] denied the existence of a fluid-crystalline phase and proposed the single gel-phase model, where the chain length of the SC lipids ($\sim 20-24$ carbon atoms), the presence cholesterol (30 mol%), and the compositional impurity promote the stability of the gel-phase. Norlén suggests that this model probably does not apply to the superficial levels of the SC, where the lower water content, the contamination by sebaceous lipids and the desquamation process might generate liquid-crystalline domains. Recently, Norlén and co-workers proposed a refinement of this model, which builds on results from high magnification cryo-electron microscopy of vitreous skin sections, molecular modelling and cryo-EM simulations[49, 53]. They proposed a lipid configuration where ceramides in stacks of lipid bilayers assume a splayed chain configuration with most of the cholesterol placed near to the sphingosine backbone and long chain free fatty acids associated with the amino-linked fatty acid chain.

Sebaceous Lipids

A thin oily layer called sebum (from the Latin meaning fat or tallow) covers the surface of our skin. Unlike the majority of other lipids in humans and animals, which carry out internal functions, the sebaceous lipids are generated in the sebaceous gland to be excreted externally, by diffusing through hair follicles onto the surface of our skin. The distribution of the sebaceous glands in the human body varies greatly depending on the region of the body, resulting in sebum-rich areas (*e.g.* the forehead) and sebum-poor areas (*e.g.* the forearm).

The sebum is made up of a mixture of lipids unique to the human body and, particularly, to this lipid layer. Table 1.1 shows the average composition of the human sebum. Triacylglycerols consist of three fatty acid chains esterified to glycerol and are the most abundant component of the human sebum. Despite being a commonly occurring class of lipids (*e.g.* in the human body triacylglycerides are found in fat present in the bloodstream[54] and in the tear film[55]), the uniqueness of the specific lipids in the human sebum can be appreciated by considering the unusual patterns of unsaturation and the broad range of acyl chain lengths[56]. An example is the sebum's most abundant fatty acid, namely sapienic acid[5], which is an isomer of palmitoleic

Lipid class	Range weight [%]	Mean weight [%]
Triglycerides	20–60	45
Wax esters	23–29	25
Squalene	10–14	12
Free fatty acids	5–40	10
Cholesterol and sterol esers	1–5	4
Diglycerides	1–2	2

Table 1.1: Sebum composition. The data in this table was originally published in A. Pappas, Epidermal surface lipids, *Dermato-Endocrinology*, 2009, 1:2, 1(2), 72–76. ©2009 Landes Bioscience[5].

acid, with a double bond in position 6. This double bond position is very uncommon in nature and makes this structure unique to humans, hence the name sapienic (from *Homo sapiens*).

For a long time, the sebum was considered a "remnant of our mammalian heritage"[8], without a critical role in human skin biology. However, in the last decade, the development of *in vitro* models of the human sebaceous glands[57] has inspired many new studies and questions about the function of the sebum. Some of the possible roles of the sebum oil are reported below. Most of these roles are linked to the properties of the individual sebum lipids. Further research on this subject is still needed.

- Onset of the disease acne. Many early studies have shown a link between an increase in sebum production and the pathophysiology of the disease acne vulgaris[6, 58]. Acne patients also show a change in sebum composition, probably due to the presence of the bacterium *P. acnes* which hydrolyses triglycerides to fatty acids[59]. Differences in sebum composition can, therefore, be used to diagnose and monitor the development of the disease.

- Skin hydration and waterproofing properties. Studies from Fluhr *et al.*[9] have highlighted the importance of glycerol generation from triglycerides in sebaceous glands for ensuring optimal SC hydration. Furthermore, the role of the sebum on the wettability properties of the skin surface in contact with a number of different solvents has been discussed in several studies[60, 61, 62, 63], suggesting that sebum-rich areas feature lower contact angles (*i.e.* better wettability) than sebum-poor areas.

- Antimicrobial property. The sebum lipids composition contributes to the natural antimi-

crobial activity of the skin[64]. Specifically, the antibacterial activity is associated with two specific fatty acids derived from sebaceous triglycerides, sapienic acid and lauric acid. For example, Drake *et al.*[65] found that sapienic acid in combination with ethanol at low concentration is very effective against methicillin-resistant *Staphylococcus aureus*. Lauric acid, instead, has a potential application as a therapeutic for the treatment of acne[66].

- Transport of antioxidants. Thiele *et al.*[10] compared the amount of antioxidant in sebum-rich and sebum-poor areas, finding the former richer in vitamin E. They concluded that vitamin E is secreted in the sebaceous gland and delivered to upper layers of skin via the sebum. This mechanism may serve to protect skin surface lipids and the upper SC from harmful oxidation.

- Drug transport. Hydrophilic molecules were found to permeate the SC four times more efficiently in sebum-rich areas than in sebum-poor areas[67]. In addition, the permeability of the forearm (*i.e.* a sebum-poor area) could be increased threefold by sebum treatment. Tsai *et al.* attribute this finding to a disordering of the SC lipid molecules induced by the sebum. Furthermore, the sebum fills the gap between the hair follicle and the skin and is considered the main transport route of the follicular pathway[14] (see the transappendageal pathway in Figure 1.3). The high permeability of follicles has promoted the development of transdermal drug delivery approaches which target this pathway[68, 69, 70, 71]

1.2.3 Review of the current state of SC simulations and gaps in the literature

As is discussed further in chapter 2, molecular dynamics (MD) simulations represent a promising tool to characterise the properties of biological systems and to unravel the mechanisms of interaction between biological complexes, with atomistic resolution[20]. However, the full complexity of the SC layer is not accessible to the current length and time scales of MD simulations, and simplified models are required to capture the most important features of SC organisation and to give insights into SC behaviour at the molecular level. Several atomistic descriptions of

SC lipids have already been presented in the literature and will be discussed below. The detail of these descriptions depends on the main aspects and questions the authors want to focus on, and on the available knowledge and benchmarking data from *in vitro* and *in vivo* experiments, on which the simulation set up can be based and the reliability of the results assessed.

Atomistic simulations have previously been used to describe the structure and phase behaviour of a single component SC lipid bilayer in water[33, 35]. Ceramide 2 was used in this study as a representative SC lipid. In addition, cholesterol and free fatty acids have been added to mixtures to understand the contribution of the different lipids to the SC properties and the relationship between the polydispersity of the lipid composition and the overall packing and fluidity of the resultant membrane[35, 72]. Further simulation studies have explored the effect of additive compounds such as ethanol[73], DMSO[33] and oils[38, 74] and have attempted to characterise the permeation of small molecules through homogeneous skin lipid membranes[39]. Reference 39 presents an extensive review on the state of the art of SC lipid simulations.

Despite the experimental evidence for interactions between the SC and the sebum oil[67], little attention has been given to the impact that the sebum oil layer might have on the properties of the skin surface. This is one of the main objective of this thesis. In addition, in all previous simulation studies, the SC membrane is modelled as a fully hydrated bilayer, neglecting the small amount of water actually present in the SC layer and its distribution. A recent publication by Del Regno and Notman[75] represents a first step in reproducing the level of hydration measured experimentally (2 water molecules per lipid[76]) in computer simulations. Interestingly, despite the low amount of water between two SC bilayers, they found that the layer of water molecules placed at the interface tends to dewet from the bilayer interplane region and accumulate into water pools. The location of water in the SC layer raises questions on the wettability of the SC lipid matrix. This question will be addressed in this thesis.

1.3 Thesis outline

The main aim of this thesis is to build on existing models of the SC by introducing the sebum and studying its properties and interactions with water and SC lipids. Properties of nanoparticles and peptides are also studied, with a view to combining these with the sebum-SC system in the future.

This thesis is sub-divided into nine chapters:

- This first chapter (**chapter 1**) presents an introduction to the field and an overview of the questions addressed in this thesis. It also includes a literature review of the current understanding of the structure and functionality of the skin and the use of computer simulations to give insight into the properties and behaviour of the outermost layer of the skin, the SC.
- **Chapter 2** introduces the theory and practicalities associated with classical molecular dynamics (MD) simulations, along with the tools and methods employed in this thesis to analyse the simulation data.
- Chapters 3, 4 and 5 are focused on understanding the functionality of the skin as a barrier and giving insight into the interaction between sebaceous and skin lipids. Current literature studies only investigate SC in water, however, in this thesis the presence of sebum is introduced. To do this, the interactions of sebum with water and with SC need to investigated separately before all three components can be combined (this scheme is illustrated in Figure 1.8).
  - In **chapter 3**, a model for the sebum layer is developed by representing it by its most abundant component, the triglyceride tri-cis-6-hexadecenoin (TG). The bulk and interfacial structure and properties of this sebum triglyceride are studied with a focus on the interaction between the sebum triglyceride and water and the possible role that the sebum may play in regulating water loss.

– **Chapter 4** explores the interaction between the sebum TG and a multicomponent model of the skin, which incorporates the three most abundant constituents of the SC: ceramide 2, cholesterol and lignoceric acid. The mechanisms of TG penetration into the SC membrane and three different mechanisms for SC lipid extraction are elucidated. These extraction mechanisms may explain the presence of SC lipids in the sebum films coating the skin surface, previously reported in *in vivo* experiments on human skin[11].

– In **chapter 5**, molecular dynamics simulations of a water droplet placed at the TG-SC lipid interface are performed to microscopically investigate the wettability properties of the skin surface.

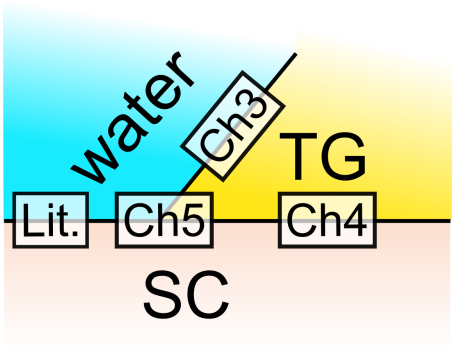


Figure 1.8: Graphical abstract of the content of chapters 3, 4 and 5 compared to the current literature (Lit.).

- Chapter 6 and 7 are focused on the study of nanoparticles and antimicrobial peptides.

  – Specifically, in **chapter 6** the thermal properties of a heated nanoparticle are investigated, with the future aim of combining this with the sebum-SC model in future work. Heated gold nanoparticles have been suggested for skin cancer treatment[77] and nanoparticles are also currently used in sunscreen creams to absorb sunlight radiation[78] so a better understanding of the properties of these particles and their interactions with the skin surface is vital.

  – The properties of the antimicrobial peptide, $G(IIKK)_4–I–NH_2$, synthesised and characterised by J. R. Lu and co-workers[79, 80], are the focus of **chapter 7**. The

conformation of the peptide in water and in proximity to DPPC and DPPG monolayers is analysed and compared to experimental data.

- **Chapter 8** presents an outreach activity called "Lipids split personality", that I have designed, developed and presented in several science festivals, to explain the main features of lipid molecules to students and the general public. The activity includes the use of a simple, hands-on experiment alongside simple cartoon illustrations and the use of MD simulations to illustrate lipid interactions at the molecular level.

- Finally, **chapter 9** presents a summary of the main conclusions of this thesis and highlights potential future work.

1.4 My publications

So far, the work presented in this thesis has been reported in the following publications (reported in chronological order):

- The contents of chapter 6 were published in: **A. S. Tascini**, J. Armstrong, E. Chiavazzo, M. Fasano, P. Asinari and F. Bresme, *Physical Chemistry Chemical Physics*, 2017, **51**, 13725-13730[81].

- Some of the contents of chapter 2 were published in: J. W. Carter, **A. S. Tascini**, J. M. Seddon and F. Bresme, *Computational Tools for Chemical Biology*, The Royal Society of Chemistry, 2018, ch 2, pp. 39-68[20].

- The contents of chapter 3 were published in: **A. S. Tascini**, M. G. Noro, R. Chen, J. M. Seddon and F. Bresme, *Physical Chemistry Chemical Physics*, 2018, **20**, 1848-1860[82].

- Part of the contents of chapter 4 have been submitted for publication as: **A. S. Tascini**, M. G. Noro, J. M. Seddon, R. Chen and F. Bresme, *Physical Chemistry Chemical Physics*, 2018, submitted.

- The content of chapter 8 has been submitted for publication as: **A. S. Tascini**, S. Wang, J. M. Seddon, F. Bresme and R. Chen, *Journal of Chemical Education*, 2018, submitted.

Bibliography to chapter 1

[1] R. J. Scheuplein and I. H. Blank, *Physiological Reviews*, 1971, **51**, 702–47.

[2] J. E. Rim, P. M. Pinsky and W. W. van Osdol, *Journal of Biomechanics*, 2008, **41**, 788–796.

[3] R. R. Warner, M. C. Myers and D. A. Taylor, *The Journal of Investigative Dermatology*, 1988, **90**, 218–224.

[4] J. A. Mcgrath and J. Uitto, *Rook's Textbook of Dermatology Volume 1 8th Edition*, 2010, pp. 1–53.

[5] A. Pappas, *Dermato-Endocrinology 1:2,*, 2009, **1**, 72–76.

[6] C. C. Zouboulis, *Clinics in Dermatology*, 2004, **22**, 360–366.

[7] F. P. Radner and J. Fischer, *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 2014, **1841**, 409–415.

[8] M. Lovászi, A. Szegedi, C. C. Zouboulis and D. Törőcsik, *Dermato-Endocrinology*, 2017, **9**, e1375636.

[9] J. W. Fluhr, M. Mao-Qiang, B. E. Brown, P. W. Wertz, D. Crumrine, J. P. Sundberg, K. R. Feingold and P. M. Elias, *Journal of Investigative Dermatology*, 2003, **120**, 728–737.

[10] J. J. Thiele, S. U. Weber and L. Packer, *Journal of Investigative Dermatology*, 1999, **113**, 1006–1010.

[11] H. M. Sheu, S. C. Chao, T. W. Wong, J. Y. Y. Lee and J. C. Tsai, *British Journal of Dermatology*, 1999, **140**, 385–391.

[12] D. T. Downing and J. S. Strauss, *Journal of Investigative Dermatology*, 1974, **62**, 228–244.

[13] M. E. Lane, *International Journal of Pharmaceutics*, 2013, **447**, 12–21.

[14] P. Kattou, G. Lian, S. Glavin, I. Sorrell and T. Chen, *Pharmaceutical Research*, 2017, **34**, 2036–2048.

[15] R. Mout, D. F. Moyano, S. Rana and V. M. Rotello, *Chemical Society Reviews*, 2012, **41**, 2539–2544.

[16] S. Nafisi and H. I. Maibach, *Cosmetic Science and Technology: Theoretical Principles and Applications*, 2017, ch. 22, pp. 337–361.

[17] S. Gullapalli and M. S. Wong, *Chemical Engineering Progress*, 2011, **107**, 28–32.

[18] Etp-nanomedicine, *Nanomedicine European Funding*, `https://www.etp-nanomedicine.eu/public/about-nanomedicine/european-funding`.

[19] *itn-snal web page*, `https://itn-snal.net`.

[20] J. W. Carter, A. S. Tascini, J. M. Seddon and F. Bresme, *Computational Tools for Chemical Biology*, The Royal Society of Chemistry, 2018, ch. 2, pp. 39–68.

[21] A. M. Kligman, *The Epidermis*, Academic Press, 1964, ch. 20, pp. 387–433.

[22] E. Christophers and A. M. Kligman, *The Journal of Investigative Dermatology*, 1964, **42**, 407–409.

[23] A. Naegel, M. Heisig and G. Wittum, *Permeability Barrier: Methods and Protocols*, Humana Press, Totowa, NJ, 2011, pp. 1–32.

[24] A. Ishida-Yamamoto and S. Igawa, *Cell and Tissue Research*, 2015, **360**, 477–482.

[25] P. M. Elias, R. Gruber, D. Crumrine, G. Menon, M. L. Williams, J. S. Wakefield, W. M. Holleran and Y. Uchida, *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, 2014, **1841**, 314–318.

[26] J. Fluhr, P. Elsner, E. Berardesca and H. I. Maibach, *Bioengineering of the Skin: Water and the Stratum Corneum*, CRC Press edn, 2004, p. 440.

[27] A. Kligman, *Skin Barrier*, Taylor & Francis Group, 2005, ch. 3, pp. 15–24.

[28] L. Norlén, B. Forslind, I. Nicander, B. L. Rozell and S. Ollmar, *Journal of Investigative Dermatology*, 1999, **112**, 72–77.

[29] P. W. Wertz and B. Van Den Bergh, *Chemistry and Physics of Lipids*, 1998, **91**, 85–96.

[30] J. van Smeden, L. Hoppel, R. van der Heijden, T. Hankemeier, R. J. Vreeken and J. A. Bouwstra, *Journal of Lipid Research*, 2011, **52**, 1211–1221.

[31] Y. Masukawa, H. Narita, H. Sato, A. Naoe, N. Kondo, Y. Sugai, T. Oba, R. Homma, J. Ishikawa, Y. Takagi and T. Kitahara, *Journal of Lipid Research*, 2009, **50**, 1708–1719.

[32] J. A. Bouwstra, G. S. Gooris, F. E. Dubbelaar and M. Ponec, *Journal of Lipid Research*, 2001, **42**, 1759–1770.

[33] R. Notman, W. K. den Otter, M. G. Noro, W. J. Briels and J. Anwar, *Biophysical Journal*, 2007, **93**, 2056–2068.

[34] R. Notman, J. Anwar, W. J. Briels, M. G. Noro and W. K. den Otter, *Biophysical Journal*, 2008, **95**, 4763–71.

[35] C. Das, P. D. Olmsted and M. G. Noro, *Soft Matter*, 2009, **5**, 4549–4555.

[36] C. Das, M. G. Noro and P. D. Olmsted, *Biophysical Journal*, 2009, **97**, 1941–1951.

[37] C. Das, M. G. Noro and P. D. Olmsted, *Soft Matter*, 2014, **10**, 7346–7352.

[38] A. Akinshina, C. Das and M. G. Noro, *Phys. Chem. Chem. Phys.*, 2016, **18**, 17446–17460.

[39] R. Notman and J. Anwar, *Advanced Drug Delivery Reviews*, 2013, **65**, 237–250.

[40] H. M. Barriga, E. S. Parsons, N. L. McCarthy, O. Ces, J. M. Seddon, R. V. Law and N. J. Brooks, *Langmuir*, 2015, **31**, 3678–3686.

[41] Y. Uchida and S. Hamanaka, *Skin Barrier*, Taylor & Francis Group, 2005, ch. 6, pp. 43–64.

[42] E. M. Zettersten, R. Ghadially, K. R. Feingold, D. Crumrine and P. M. Elias, *Journal of the American Academy of Dermatology*, 1997, **37**, 403–408.

[43] M. Mao-Qiang, B. E. Brown, S. Wu-Pong, K. R. Feingold and P. M. Elias, *Archives of Dermatology*, 1995, **131**, 809–816.

[44] H. C. Chen, R. Mendelsohn, M. E. Rerek and D. J. Moore, *Biochimica et Biophysica Acta - Biomembranes*, 2001, **1512**, 345–356.

[45] E. H. Mojumdar, G. S. Gooris and J. A. Bouwstra, *Soft Matter*, 2015, **11**, 4326–4336.

[46] L. Norlén, *Journal of Investigative Dermatology*, 2001, **117**, 830–836.

[47] J. A. Bouwstra and G. S. Gooris, *The Open Dermatology Journal*, 2010, **4**, 10–13.

[48] J. Bouwstra, G. Gooris and M. Ponec, *Journal of Biological Physics*, 2002, **28**, 211–223.

[49] M. Lundborg, A. Narangifard, C. L. Wennberg, E. Lindahl, B. Daneholt and L. Norlén, *Journal of Structural Biology*, 2018, **203**, 149–161.

[50] B. O. Forslind, *Acta Dermato-Venereologica*, 1994, **74**, 1–6.

[51] B. Forslind, S. Engström, J. Engblom and L. Norlén, *Journal of Dermatological Science*, 1997, **14**, 115–125.

[52] J. A. Bouwstra, F. E. R. Dubbelaar, G. S. Gooris and M. M. Ponec, *Acta Derm Venereol*, 2000, **208**, 23–30.

[53] I. Iwai, H. Han, L. D. Hollander, S. Svensson, L. G. Öfverstedt, J. Anwar, J. Brewer, M. Bloksgaard, A. Laloeuf, D. Nosek, S. Masich, L. A. Bagatolli, U. Skoglund and L. Norlén, *Journal of Investigative Dermatology*, 2012, **132**, 2215–2225.

[54] A. Hall, J. Repakova and I. Vattulainen, *Journal of Physical Chemistry B*, 2008, **112**, 13772–13782.

[55] J. Telenius, A. Koivuniemi, P. Kulovesi, J. M. Holopainen and I. Vattulainen, *Langmuir*, 2012, **28**, 17092–17100.

[56] N. Nicolaides, *Science*, 1974, **186**, 19–26.

[57] C. C. Zouboulis, M. Picardo, Q. Ju, I. Kurokawa, D. Törőcsik, T. Bíró and M. R. Schneider, *Reviews in Endocrine and Metabolic Disorders*, 2016, **17**, 319–334.

[58] J. L. Burton, *Clinical science*, 1970, **39**, 757–67.

[59] M. R. Motwani, L. D. Rhein and J. L. Zatz, *Journal of cosmetic science*, 2001, **52**, 211–24.

[60] A. Mavon, H. Zahounani, D. Redoules, P. Agache, Y. Gall, P. Humbert, H. Zahouani, D. Redoules, P. Agache, Y. Gall and P. Humbert, *Colloids and Surfaces B: Biointerfaces*, 1997, **8**, 147–155.

[61] A. Mavon, D. Redoules, P. Humbert, P. Agache and Y. Gall, *Colloids and Surfaces B: Biointerfaces*, 1998, **10**, 243–250.

[62] J. Krawczyk, *Skin Research and Technology*, 2015, **21**, 214–223.

[63] P. Capra, G. Musitelli and P. Perugini, *International Journal of Cosmetic Science*, 2017, **39**, 393–401.

[64] C. L. Bratt, P. Wertz, D. Drake, D. V. Dawson and K. A. Brogden, *Lipids and Essential Oils as Antimicrobial Agents*, 2010, pp. 99–121.

[65] D. R. Drake, K. A. Brogden, D. V. Dawson and P. W. Wertz, *Journal of Lipid Research*, 2008, **49**, 4–11.

[66] T. Nakatsuji, M. C. Kao, J. Y. Fang, C. C. Zouboulis, L. Zhang, R. L. Gallo and C. M. Huang, *Journal of Investigative Dermatology*, 2009, **129**, 2480–2488.

[67] J.-C. Tsai, C.-C. Lu, M.-K. Lin, J.-W. Guo and H.-M. Sheu, *Skin pharmacology and*

physiology, 2012, **25**, 124–32.

[68] S. Jung, N. Otberg, G. Thiede, H. Richter, W. Sterry, S. Panzner and J. Lademann, *Journal of Investigative Dermatology*, 2006, **126**, 1728–1732.

[69] M. Morgen, G. W. Lu, D. Du, R. Stehle, F. Lembke, J. Cervantes, S. Ciotti, R. Haskell, D. Smithey, K. Haley and C. Fan, *International Journal of Pharmaceutics*, 2011, **416**, 314–322.

[70] J. Lademann, H. Richter, S. Schanzer, F. Knorr, M. Meinke, W. Sterry and A. Patzelt, *European Journal of Pharmaceutics and Biopharmaceutics*, 2011, **77**, 465–468.

[71] A. Patzelt and J. Lademann, *Expert Opinion on Drug Delivery*, 2013, **10**, 787–797.

[72] C. Das, M. G. Noro and P. D. Olmsted, *Physical Review Letters*, 2013, **111**, 1–5.

[73] R. Thind, D. W. O'Neill, A. Del Regno and R. Notman, *Chem. Commun.*, 2015, **51**, 5406–5409.

[74] M. I. Hoopes, M. G. Noro, M. L. Longo and R. Faller, *Journal of Physical Chemistry B*, 2011, **115**, 3164–3171.

[75] A. Del Regno and R. Notman, *Physical Chemistry Chemical Physics*, 2018, **20**, 2162–2174.

[76] D. Groen, G. S. Gooris, D. J. Barlow, M. J. Lawrence, J. B. Van Mechelen, B. Demé and J. A. Bouwstra, *Biophysical Journal*, 2011, **100**, 1481–1489.

[77] I. Salas-García, F. Fanjul-Vélez, N. Ortega-Quijano, J. Poliak, L. Hudcová, O. Wilfert and J. L. Arce-Diego, Proceedings of 22nd International Conference Radioelektronika 2012, 2012, pp. 1–4.

[78] I. Krasnikov, A. Popov, A. Seteikin and R. Myllylä, *Biomedical Optics Express*, 2011, **2**, 3278.

[79] J. Hu, C. Chen, S. Zhang, X. Zhao, H. Xu, X. Zhao and J. R. Lu, *Biomacromolecules*, 2011, **12**, 3839–3843.

[80] D. Ciumac, R. A. Campbell, H. Xu, L. A. Clifton, A. V. Hughes, J. R. Webster and J. R. Lu, *Colloids and Surfaces B: Biointerfaces*, 2017, **150**, 308–316.

[81] A. S. Tascini, J. Armstrong, E. Chiavazzo, M. Fasano, P. Asinari and F. Bresme, *Phys. Chem. Chem. Phys.*, 2017, **51**, 13725–13730.

[82] A. S. Tascini, M. G. Noro, R. Chen, J. M. Seddon and F. Bresme, *Phys. Chem. Chem.*

Phys., 2018, **20**, 1848–1860.

Chapter 2

Background Theory and Methods

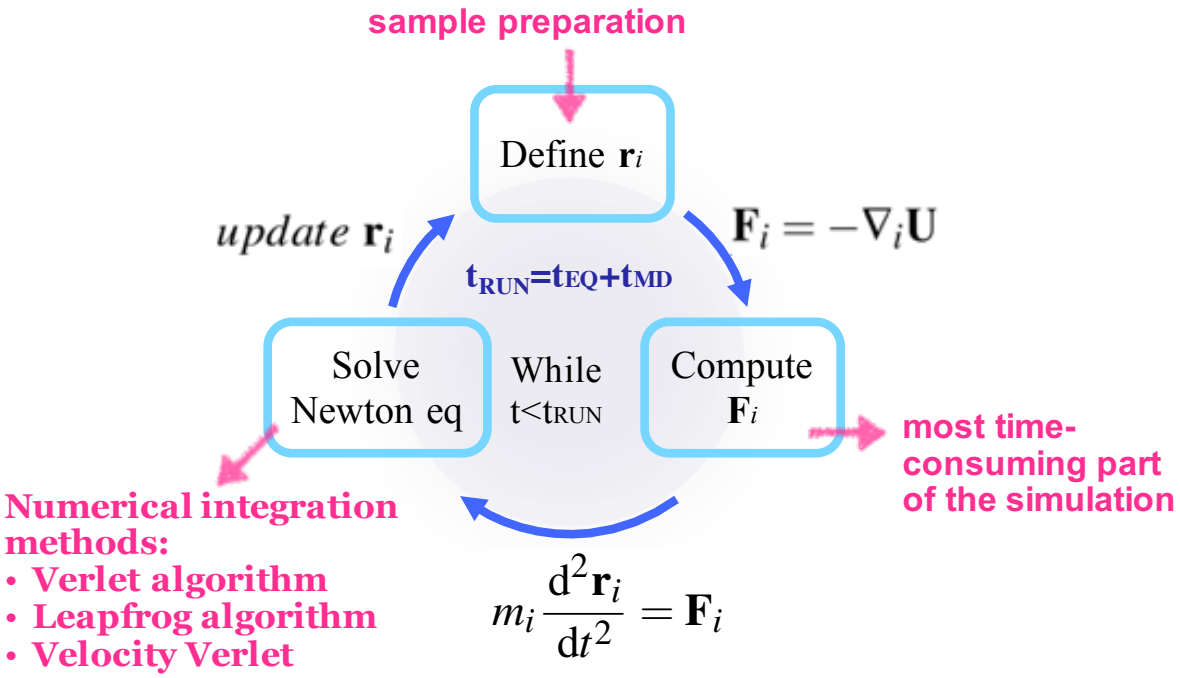


Figure 2.1: Title page figure for chapter 2: molecular dynamics simulation scheme.

The foundations of the molecular dynamics (MD) simulation method were laid in the pioneering works of Alder, Mansigh and Wainwright on computer simulations of hard spheres[1, 2]. Due to the incredible evolution of high-performance computing and theory development in the last fifty years, this method has reached a high level of sophistication that now permits the investigation of complex biological molecules and processes. MD is now a discipline in its own right, that both complements experimental measurement and acts as a unique predictive tool, providing much needed insight at molecular scales, where experiments are not always possible or are difficult. As was discussed in chapter 1, MD has been used previously to investigate the structural and mechanical properties of skin lipids, *e.g.* ceramides[3, 4, 5] and mixtures of ceramides, cholesterol and fatty acids[6, 7, 8, 9, 10, 11], and to investigate their interaction with molecules[12, 13, 14] and nanoparticles[15, 16]. Simulations have also been extensively used to investigate the interaction of lipid bilayers with nanoparticles[17, 18, 19, 20] and peptides[21, 22, 23, 24].

Figure 2.2: Berni Alder stands in his office and looks over the shoulders of Mary Ann Mansigh and Tom Wainwright. They are observing oscillograph traces from the computer generated motion of particles.

The aim of this chapter is to introduce the theory and practicalities associated with classical MD simulations, along with the data analysis tools employed here. Analysis tools and methods which I developed specifically for the work in this thesis are explained in the methods sections of the corresponding results chapters. Material from this chapter has been published in chapter 2 of the RSC book *Computational Tools for Chemical Biology*[25] that I co-authored, and has been here reproduced from ref. 25 with permission from the Royal Society of Chemistry.

2.1 The basics of molecular dynamics

MD is a computational technique in which the behaviour of atoms and molecules is described explicitly using classical mechanics. Applying Newton's equation of motion to a system of N

particles results in N coupled second order differential equations

$$m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = \mathbf{F}_i \left(\mathbf{r}^N \right), \tag{2.1}$$

where $\mathbf{r}_i$, m_i, $\mathbf{F}_i$ are the position, mass and force on particle i, t is the time, and $\mathbf{r}^N$ represents the positions of all N particles.

To study the evolution of the system over time, equation 2.1 must be integrated numerically in discrete time steps (Δt). A number of algorithms have been developed for this purpose, many of them variations of the Verlet algorithm[26], such as the velocity Verlet[27] or leap-frog algorithms[28]. These algorithms rely on the use of a Taylor series, where the position of a particle at time $t + \Delta t$ is expressed in terms of its position, velocity ($\dot{\mathbf{r}}_i(t)$) and acceleration ($\ddot{\mathbf{r}}_i(t)$) at time t, as:

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \dot{\mathbf{r}}_i(t) + \frac{1}{2} \Delta t^2 \ddot{\mathbf{r}}_i(t) + \ldots, \tag{2.2}$$

where $\dot{\mathbf{r}}_i(t) = d\mathbf{r}_i(t)/dt = \mathbf{v}(t)$ and $\ddot{\mathbf{r}}_i(t) = d^2\mathbf{r}_i(t)/dt^2 = \mathbf{F}_i(t)/m_i$.

In order to increase the accuracy of a single Taylor series for $\mathbf{r}_i(t + \Delta t)$, the Taylor series expansions for particle positions forwards ($\mathbf{r}_i(t + \Delta t)$) and backwards ($\mathbf{r}_i(t - \Delta t)$) in time are combined[26]. By neglecting all the terms higher than the second order, the future position of each particle is given by:

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i} \mathbf{F}_i(t). \tag{2.3}$$

A disadvantage of the Verlet algorithm is that the velocity at time t depends on the position at time $(t + \Delta t)$ and $(t - \Delta t)$, and therefore velocities do not update at the same time as positions. This limitation is solved in the velocity Verlet algorithm[27] where the velocity at time $t + \Delta t$ can be derived explicitly from:

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{\Delta t^2}{2m_i} \left[\mathbf{F}_i(t) + \mathbf{F}_i(t + \Delta t) \right]. \tag{2.4}$$

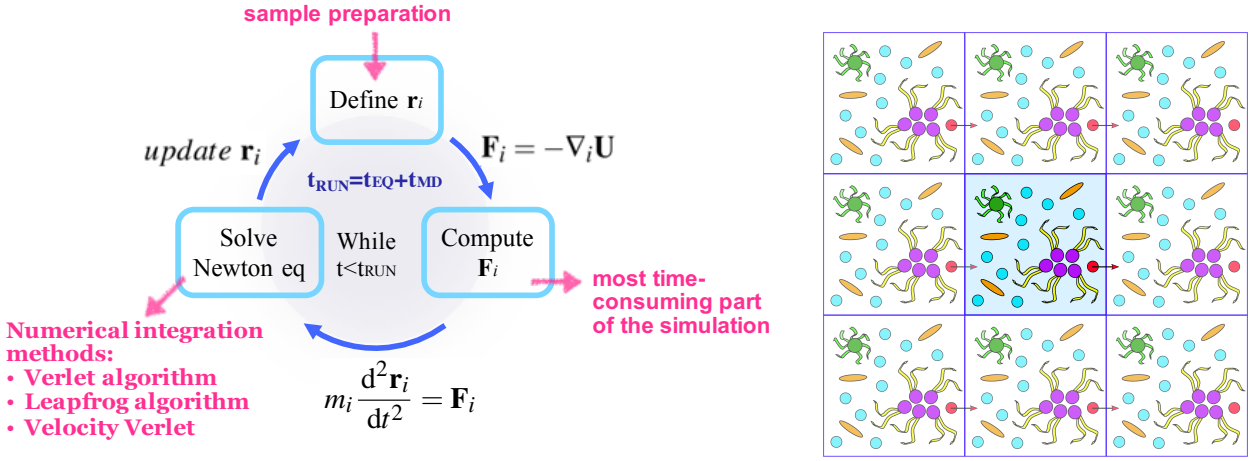
Another widely used approach to integrate the equations of motion is the Leap-frog algorithm[28, 29]. It is the default MD integrator in one of the most commonly used MD software packages, GROMACS[30]. This algorithm uses positions $\mathbf{r}_i$ at time t and velocities $\mathbf{v}_i$ at time $(t - \Delta t/2)$. The forces $\mathbf{F}_i(t)$, determined by the positions at time t, are used to update the velocities, as:

$$\mathbf{v}_i(t + \Delta t/2) = \mathbf{v}_i(t - \Delta t/2) + \frac{\Delta t}{m}\mathbf{F}_i(t) \tag{2.5}$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t + \Delta t/2). \tag{2.6}$$

The resulting trajectory is identical to the Verlet algorithm described above (see equation 2.3). Both the Leap-frog, Verlet and velocity Verlet algorithms satisfy the properties of time-reversibility and the symplectic property, which are crucial for the long-time stability of a numerical solver[31].

It follows from equations 2.3 and 2.4 that to run an MD simulation, the MD software must be provided with the initial positions, $\{\mathbf{r}_1(0), ..., \mathbf{r}_N(0)\}$, and velocities, $\{\dot{\mathbf{r}}_1(0), ..., \dot{\mathbf{r}}_N(0)\}$, for all particles and a means to calculate the forces. In this sense, an MD simulation can be described as a high-resolution *in-silico* experiment (see Figure 2.3a), where the first step involves the preparation of the sample to study, *i.e.* the definition of $\{\mathbf{r}_i(0)\}$ and $\{\dot{\mathbf{r}}_i(0)\}$. To eliminate the edge effects resulting from the finite size of the system ($N \approx 10^3$—10^6 particles) periodic boundary conditions are generally used[32], *i.e.* the simulation box is made periodic so that particles crossing the edge of the box, re-enter the box from the opposite side (see Figure 2.3b). The simulation is then run by repeatedly solving the equations of motion numerically and updating particle positions and velocities. Before obtaining results, during the production run of the simulation, it is important to allow time for the system equilibriate and relax from the initial configuration. This process can be monitored by checking that properties of the system such as the energy no longer change with time. In an MD simulation, as in an experiment, the working time is one of the key parameters to choose (more details about this in section 2.3). Measurements are subject to statistical noise, the effect of which can be reduced by averaging measurements over a longer simulation time. The same concept applies for the system size, the larger the system, the more accurate the measurements are. The analogy between simulations



(a) MD simulation scheme **(b)** pbc

Figure 2.3: a) Scheme capturing the most important steps of an MD simulation: *i.e.* definition of the intial conditions and positions, computing the force depending on the interaction potentials, updating the position by integrating the equation of motion. The entire procedure is repeated for t_{RUN}. b) Two-dimensional schematic representation of periodic boundary conditions (pbc). The particle trajectories in the central simulation box are copied in every direction. A particle exiting the box will enter from the other side.

and experiments allows us to explain the most common mistakes in computer simulations, such as the system not being properly prepared, the system being too small, too short measurements time or the system undergoing an irreversible change during the simulation. Care must also be taken to ensure that the measurements taken from simulations correspond to the properties of interest.

2.1.1 Thermostatting and barostatting

The direct application of Newton's equation of motion results in conservation of the total energy (E) and hence the system is isolated from the surroundings, with a fixed number of particles (N), volume (V) and total energy (E). This ensemble is known as the microcanonical ensemble (NVE). However, many of the problems of interest and most of the experiments to compare with do not occur under constant energy conditions. It is therefore important to use methods which work with ensembles with different sets of thermodynamic control variables, such as the canonical ensemble (NVT), the isothermal-isobaric ensemble (NPT) and the grand canonical ensemble (μVT). Here, algorithms to simulate NVT and NPT ensembles, which are used

throughout this thesis, are described.

Canonical ensemble (*NVT*)

There are several ways to simulate constant temperature conditions. One possibility to control the simulation temperature is to exploit the relationship between temperature and average kinetic energy (E_K), *i.e.*

$$E_K = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i^2 = \frac{N_f}{2} k_B T, \tag{2.7}$$

where N_f is the number of degrees of freedom. The algorithm proposed by Berendsen *et al.*[33] mimics a weak coupling between the system and an external bath with temperature equal to T_0. Any deviation from the bath temperature is corrected according to

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}, \tag{2.8}$$

where τ is the time constant which regulates the rate of exponential decay of the temperature deviation. Despite providing some flexibility in the choice of the coupling strength (by varying τ), this algorithm suppresses the fluctuations of the kinetic energy (especially for short coupling time constants, *i.e.* $\tau \sim 0.01\,\text{ps}$) and does not correctly sample the canonical ensemble, hence caution should be used when deriving thermodynamic properties. The correct kinetic energy distribution can be obtained by including an additional stochastic term[34]:

$$dK = (K_0 - K)\frac{dt}{\tau} + 2\sqrt{\frac{K K_0\, dW}{N_f \sqrt{\tau}}}, \tag{2.9}$$

where K is the kinetic energy, N_f is the number of degrees of freedom and dW is a Wiener process. This approach produces the correct canonical ensemble, absence of oscillations and a first order decay, and is extremely efficient for relaxing a system to the target temperature.

A different approach, known as the Nosé-Hoover thermostat, was proposed by Nosé[35] and later modified by Hoover[36]. According to this method a thermal reservoir (with a fictitious mass Q) and a friction term, proportional to the product of each particle velocity and a friction

parameter, ζ, are added to the equation of motion, which becomes

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} - \frac{p_\zeta}{Q}\frac{d\mathbf{r}_i}{dt}, \tag{2.10}$$

where p_ζ is the momentum of the friction parameter, which has its own equation of motion, $dp_\zeta/dt = (T - T_0)$. In general the strength of this thermostat is regulated by varying the period of the oscillations of the kinetic energy between the system and the reservoir, τ, which is directly related to the thermal reservoir mass, Q, by

$$Q = \frac{\tau^2 T_0}{4\pi^2}. \tag{2.11}$$

Due to the oscillatory character of this relaxation, in this case the actual relaxation time of the coupling is several time-steps longer than the oscillations period and the time constant should be normally 4-5 times larger than the relaxation time used with weak coupling[37].

Isothermal-isobaric ensemble (*NPT*)

Similarly to what was described in the previous subsection, in some cases it may be necessary to hold pressure constant and allow the volume to fluctuate. Following the same logic as for the temperature, an instantaneous definition of pressure can be derived and particle coordinates and box vectors can be scaled to oppose deviations from a prespecified target pressure. Again a range of algorithms exist to perform constant P simulations. The Berendsen algorithm[33] can be applied to rescale the coordinates of the box and relax the pressure towards a reference value P_0, at a rate defined by the relaxation constant τ_P, as:

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_P}. \tag{2.12}$$

The rescaling is made through a matrix μ given by

$$\mu_{ij} = 1 - \frac{\beta_{ij}\Delta t}{3\tau_P}\{P_{0ij} - P_{ij}(t)\}, \tag{2.13}$$

where i and j are the three directions x, y and z, β_{ij} is an element of the matrix describing the isothermal compressibility of the system. Note that these values do not influence the system dynamics much. Despite achieving the correct average pressure, this algorithm does not yield the exact NPT ensemble.

An alternative approach to simulate the true NPT ensemble is the Parrinello-Rahman barostat [38, 39]. Similarly to the Nosé-Hoover thermostat, in this case the box vectors are represented by a matrix which obeys its own equation of motion.

2.2 Forcefields for biomolecular simulations

MD simulations rely on the definition of a forcefield, a mathematical function of the positions of all particles which defines the potential energy of a system configuration, $U\left(\mathbf{r}^N\right)$. Once the forcefield is defined, the instantaneous forces acting on each particle in a given configuration can be computed from the derivative of the potential, as:

$$\mathbf{F}_i = -\nabla_i U\left(\mathbf{r}^N\right) = -\frac{\partial U\left(\mathbf{r}^N\right)}{\partial \mathbf{r}_i}. \tag{2.14}$$

$\mathbf{F}_i$ is then subsequently fed into the MD algorithm to update the particle positions and velocities and advance the simulation by a time step.

In practice, the vast majority of forcefields share the same basic form, defined in equation 2.15, differing only in the values of the tuneable, atom and molecule specific parameters associated with each term.

$$\begin{aligned} U\left(\mathbf{r}^N\right) = &\sum_{ij\in\{\text{Bonds}\}} k_{ij}\left(r_{ij} - r_{ij,0}\right)^2 + \sum_{ijk\in\{\text{Angles}\}} k_{ijk}\left(\theta_{ijk} - \theta_{ijk,0}\right)^2 \\ &+ \sum_{ijkl\in\{\text{Dihedrals}\}} k_{ijkl}\left(1 + \cos\left(n_{ijkl}\phi_{ijkl} - \phi_{ijkl,0}\right)\right) \\ &+ \sum_{ij\in\{\text{LJ}\}} 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right) + \sum_{ij\in\{\text{Coulomb}\}} \frac{Q_i Q_j}{4\pi\epsilon_0 r_{ij}}. \end{aligned} \tag{2.15}$$

Each term in the forcefield function represents a different type of interaction, as described

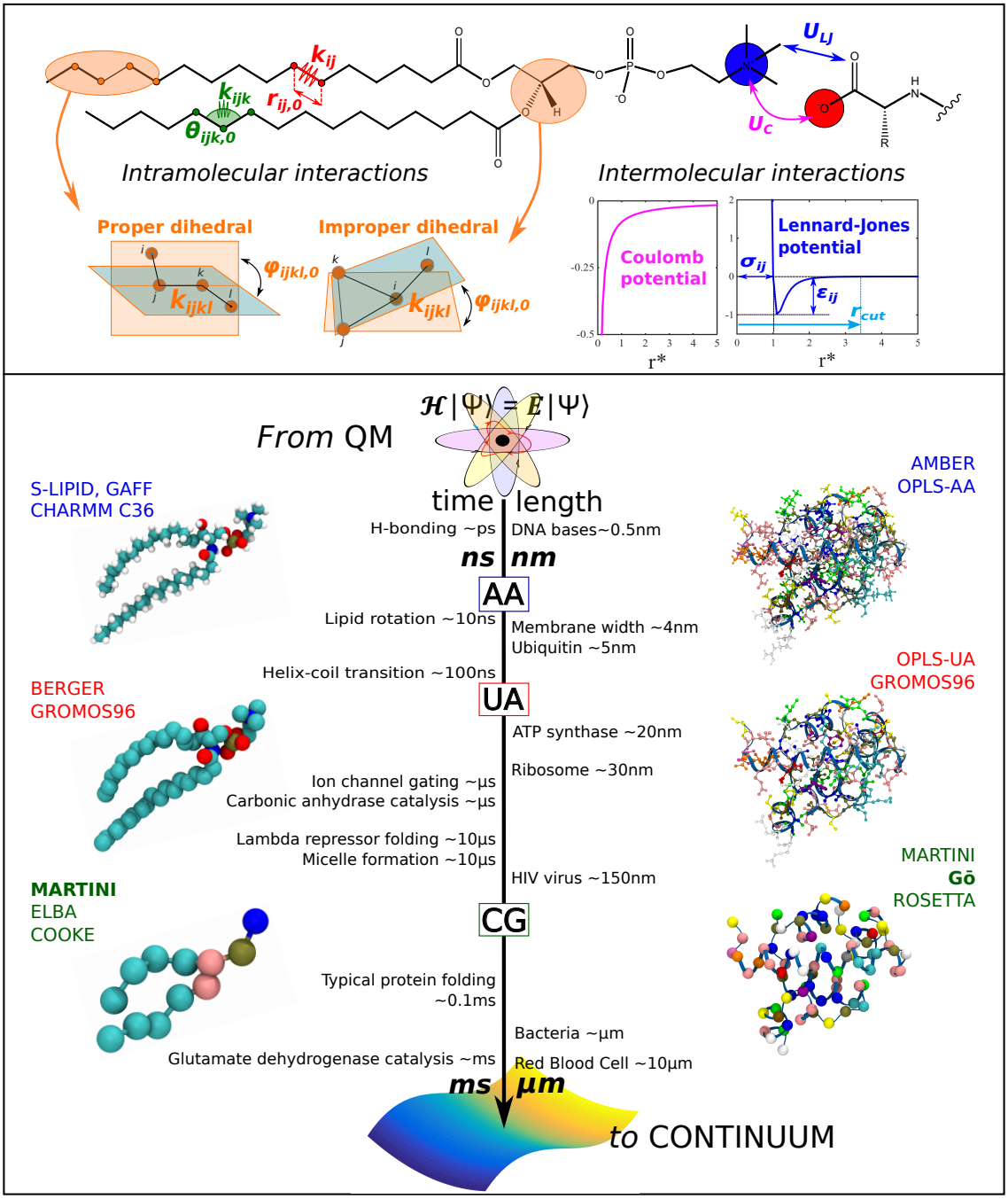


Figure 2.4: Top) Illustration of the different contributions to the forcefield (equation 2.15) applied to a lipid (DPPC) and a general amino acid. k are force constants and $r_{ij,0}$, $\theta_{ijk,0}$, $\phi_{ijkl,0}$ are equilibrium bond lengths, angles and dihedral angles, respectively. Dihedral potentials (periodicity n_{ijkl}) are applied to linear chains of atoms and improper dihedrals are used here for tetrahedral groups. The shape of the Lennard-Jones and Coulomb potentials are also shown. Bottom) Representative structures for a lipid (DPPC (from LipidBook[40])) and protein (λ-repressor (taken from PDB structure 3KZ3[41])) using different resolution models alongside commonly used forcefields at those resolutions and the typical timescales and system sizes accessible in simulations. This figure has been reproduced from ref. 25 with permission from the Royal Society of Chemistry.

below. They can be broadly divided into two groups:

- **Intramolecular (or bonded) interactions**, which account for covalent bonds and molecular flexibility and includes "bonds", "angles" and "dihedrals" contributions.
- **Intermolecular (or non-bonded) interactions**, which often include the Lennard-Jones (LJ) potential, which describes the Pauli repulsion and van der Waals attraction between atoms. Electrostatic interactions between charges, Q, are handled using the Coulomb potential.

An example of how these interactions operate is shown in Figure 2.4-top.

The covalent bond and angle interactions are commonly defined via harmonic potentials with an equilibrium bond length $r_{ij,0}$ and angle $\theta_{ijk,0}$, and associated force constant k_{ij} and k_{ijk}, respectively. However, it is common practise at relevant biological temperatures ($\sim 310\,\mathrm{K}$), where molecular bond vibrations are in the ground state, to replace the harmonic bond term in equation 2.15 by a rigid constraint[42]. The aim of this constraint is to restore the bond length to its equilibrium value each time step, and remove fast bond vibrational motion from the simulation.

The dihedral potential acts on the angle ϕ_{ijkl} between two planes containing atoms ijk and jkl (see Figure 2.4). It can take a number of different forms, but is represented in equation 2.15 as a cosine function (with force constant k_{ijkl}, periodicity n_{ijkl} and equilibrium angle $\phi_{ijkl,0}$). Other common functions are combinations of cosine functions[43] or harmonic potentials, often used for improper dihedrals to give rigidity to rings and tetrahedral sites.

The LJ potential is pairwise additive, *i.e.* it only contains two-body interactions. The effects of interactions involving three or more particles are averaged out and accounted for in the parameters of the effective two-body potential (ϵ_{ij} and σ_{ij} in equation 2.15). ϵ_{ij} and σ_{ij} define the interaction strength and effective particle diameter, respectively. The term proportional to r^{-6} models the theoretical distance dependence of the van der Waals interaction[44], whereas the exponent of the repulsive term has been chosen for computational efficiency (as the square of the r^{-6} term). Rather than parameterising the LJ parameters for every possible pair of atom

types, σ and ϵ are commonly given for single atom types. Interactions between different types are then derived from approximate combination rules.

Calculation of non-bonded forces is the most time consuming part of an MD simulation and the use of periodic boundary conditions means that there is potentially an infinite number of non-bonded atom pairs. For short range interactions such as the LJ potential, the interaction approaches zero sufficiently quickly that it can be neglected beyond a cut off distance, r_{cut} (usually around 2.5–5σ for the LJ potential and it should be less than half of the box length to prevent the same interaction from being counted twice). The tail of the potential can be modified (shifting the potential) to avoid artefacts associated with force discontinuities at r_{cut} and to ensure a stable trajectory. At long distances single LJ interactions are very weak. However, the larger number of particles and the fact that interactions are all attractive at these distances means that the total interaction is still significant and must be accounted for by a dispersion correction term added to the potential energy[45]. Including dispersion corrections to energy and pressure is of significant importance in simulations at constant pressure when properties corresponding to the full potential are required, *e.g.* density or surface tension.

The electrostatic term in equation 2.15, described by the Coulomb potential, is subject to the same problem as the LJ term, but decays much more slowly ($\sim r^{-1}$) which implies that a shifted potential cannot be used reliably. Instead, the electrostatic term is treated using the Ewald summation method[32], in which the potential is divided into a short ranged and long ranged component, the latter computed in reciprocal space. This method is considered the state of the art for computation of electrostatic interactions, however it imposes a periodicity in the system that must be considered carefully in computations of solvation energies[46].

A great variety of forcefields exist for the simulation of biological membranes, making the choice between them often a challenging step. In general, the parametrisation of forcefields can be based on quantum calculations of small fragments of the target molecule or by fitting to experimental properties, both to structural data such as bond lengths from crystal structures, and thermodynamic data, as in the case of organic molecules fitted to reproduce liquid densities and vaporisation enthalpies. When choosing a forcefield, it is vital to understand the philosophy

behind its parameterisation to appreciate the various limitations of the model. Parameters from different forcefields should not be mixed blindly. When setting up a simulation it is also important to choose simulation parameters (*e.g.* potential cutoff) bearing in mind those used in the original parameterisation.

Depending on the degree of resolution, forcefields can be classified as atomistic or coarse grained, as shown in Figure 2.4-bottom. From a higher to lower degree of resolution there are:

- **All atom (AA) forcefields**, in which every atom in a molecule is represented explicitly as a distinct interaction site.
- **United atom (UA) forcefields**, in which non-polar, alkyl hydrogens are merged with the adjacent carbon atom interaction site, reducing the number of constituent atoms in each molecule.
- **Coarse grained (CG) forcefields**, in which molecules can be represented by fewer interaction sites while maintaining some chemical specificity by combining atoms within functional groups or natural molecular subunits, such as amino acids, into a smaller number of beads.

Figure 2.4-bottom includes some of the most commonly used forcefields for lipid molecules (left column) and proteins (right column). Atomistic forcefields include AMBER, CHARMM, GROMOS (UA only) and OPLS, with the first three developed alongside MD software packages of the same name. A large number of different parameter sets have been published, with some of these available within MD software packages or, increasingly, from online databases such as LipidBook[40] or the database of antimicrobial compounds[47]. The advantage of using a UA forcefield is particularly significant in simulations of lipid membranes, due to the high number of non-polar hydrogens in the lipid acyl chains. A widely used UA forcefield, developed to reproduce experimental properties of lipid membranes is the Berger forcefield[48]. This parameter set is based on OPLS[49] and GROMOS[50] forcefields, adjusted to yield the correct volume and heat of vaporisation of pentadecane. It includes the Ryckaert–Bellemans dihedral potential[43] to model the hydrocarbon chains, whose function differs from the one presented in equation 2.15

and is based on expansion in powers of $\cos\phi_{ijkl}$, *i.e.* $\sum_{n=0}^{5} k_{ijkl,n} \left(\cos\left(\phi_{ijkl} - 180^\circ\right)\right)^n$. This forcefield has been used to simulate model skin lipid bilayers in this thesis.

CG forcefields enable a significant increase in the simulation size and time scale by reducing the number of degrees of freedom, which results in a smoother energy landscape. This permits the use of longer time steps (from $\sim 1\,\text{fs}$ in AA and UA simulations to $\sim 10\,\text{fs}$ in CG). Atomistic simulation data can be used to obtain CG parameters, by using systematic coarse-graining approaches, such as Boltzmann inversion[51]. One of the most widely used CG models is the MARTINI model, originally developed for lipids[52], but now extended to proteins, nucleic acids and carbohydrates. This forcefield is based on a 4:1 mapping of heavy (non-hydrogen) atoms to CG beads and was parametrised by computing thermodynamic properties (*i.e.* the free energies of vaporisation, hydration, and partitioning) and comparing them to experimental values. The MARINI parameters set can reproduce bilayer properties, *e.g.* the stress profile, lamellar-micellar transition and free energy of lipid desorption and flip-flopping, with semi-quantitative agreement with experiment[52]. In this thesis, the MARTINI model has been used for the simulations described in chapter 8. Those simulations were performed to explain the microscopic interactions between milk lipids and soap surfactant to the general public, and run alongside an outreach activity aimed at communicating the basics of lipodomics to the general public.

2.3 Extracting information from MD simulations

The output from an MD simulation consists of the positions, velocities and forces of each particle at each time step of the simulation. This immediately provides some structural and dynamic information about the microscopic behaviour of the system. However, in general the properties of interest are macroscopic properties (X), *e.g.* thermodynamic properties. They can often be calculated by taking advantage of the ergodic hypothesis, which states that ensemble averages, $\langle X \rangle$, can be approximated by time averages:

$$\langle X \rangle = \int\int X\left(\mathbf{r}^N, \mathbf{p}^N\right) P\left(\mathbf{r}^N, \mathbf{p}^N\right) d\mathbf{r}^N d\mathbf{p}^N = \lim_{t\to\text{inf}} \frac{1}{t} \int_0^t X\left(\mathbf{r}^N, \mathbf{p}^N; t'\right) dt', \tag{2.16}$$

where $P\left(\mathbf{r}^N, \mathbf{p}^N\right)$ is the probability distribution, defined for the canonical ensemble as:

$$P\left(\mathbf{r}^N, \mathbf{p}^N\right) = \frac{\exp\left(-E\left(\mathbf{r}^N, \mathbf{p}^N\right)/k_B T\right)}{\int \exp\left(-E\left(\mathbf{r}^N, \mathbf{p}^N\right)/k_B T\right) d\mathbf{r}^N d\mathbf{p}^N} \tag{2.17}$$

Equation 2.16 shows that in principle all points in the phase space (*i.e.* all possible values of $\left(\mathbf{r}^N, \mathbf{p}^N\right)$) contribute to the ensemble average, however the major contribution comes from states with a high probability of being visited, *i.e.* higher $P\left(\mathbf{r}^N, \mathbf{p}^N\right)$. This suggests that an ensemble average can be approximated by a time average, providing that the timescale is sufficiently long that a representative sample of high probability states are visited, or, in other words, care must be taken to assess whether the system is really ergodic over the simulation timescale.

For many systems with rugged energy landscapes, simulation timescales may be too short to properly explore the available phase space. For example, proteins might become stuck in a local minimum, without visiting other metastable states. Various advanced simulation techniques exist to improve sampling especially for non-ergodic systems as described below. In addition, MD differs from some other simulation techniques in that it can be used to obtain dynamic properties, often from correlation functions, for example, Green-Kubo relations use correlation functions to calculate transport properties such as diffusion coefficients or thermal conductivities.

2.3.1 Free energy difference between two states

There are some properties which cannot be expressed as ensemble averages and hence the ergodic principle can not be directly applied to derive properties from time averages. One of the most significant examples, is the evaluation of the free energy, A, given by $A = -k_B T \log Q$. In the free energy equation, Q is the canonical partition function, which is proportional to the volume of phase space, *i.e.* the set of all possible system configurations with positions $\mathbf{r}^N$ and

momenta $\mathbf{p}^N$ and hence can not be approximated by a time average. Q can be expressed as:

$$Q = \int \exp\left(-E\left(\mathbf{r}^N, \mathbf{p}^N\right)/k_B T\right) d\mathbf{r}^N d\mathbf{p}^N. \tag{2.18}$$

However, differences in free energy can be calculated as ensemble averages using thermodynamic integration, where the derivative of the free energy is integrated between two different states, λ_1 and λ_0:

$$\Delta A = A\left(\lambda_1\right) - A\left(\lambda_0\right) = \int_{\lambda_0}^{\lambda_1} \frac{\partial A}{\partial \lambda} d\lambda = \int_{\lambda_0}^{\lambda_1} \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_\lambda d\lambda \tag{2.19}$$

The derivative of the free energy can refer to a real, measurable property, *e.g.* the pressure ($(\partial A/\partial V)_T = -P$), as could be calculated experimentally. However, in simulations, free energy derivatives can also represent unmeasurable or non-physical quantities, increasing the number of potential applications of this technique. By introducing a parameter, λ, which switches between two different potential energy functions (*e.g.* setting $U(\lambda) = \lambda U_1 + (1-\lambda)U_0$), "alchemical transitions" can be used to change the chemical identity of atoms or molecules and applied, for example, to measure the strength of binding of different ligands[53]. A good overlap in phase space between the initial U_0 and final U_1 states is required for this technique to be accurate.

2.3.2 Enhanced conformational sampling

Often complex biomolecules are characterised by a rugged energy landscapes and the simulation timescale may be too short to properly explore the available phase space. For example: proteins may become trapped in a local minimum for the entire duration of a simulation, hence producing a non-ergodic trajectory. Advanced simulation techniques have been developed to enhance sampling and avoid the non-ergodic problem. In chapter 7 we use some of these techniques, specifically temperature replica exchange molecular dynamics and mass induced accelerated dynamics to better sample the phase space of an antimicrobial peptide. These approaches are discussed in the following paragraphs.

Replica exchange

Sometimes it is advantageous to allow a drastic change in configuration to immediately jump to a new region of phase space which can then be explored by MD. This can be achieved by introducing a Monte Carlo (MC) step into the simulation. Possible MC moves can include overall molecular rotations and translations, or intermolecular rearrangements, through insertion, deletion or switching of particles. Some biomolecular systems are well suited to MC/MD simulation, for example the structural similarity between lipid molecules permits MC moves to change lipid identity allowing a more extensive probing of phase space[54].

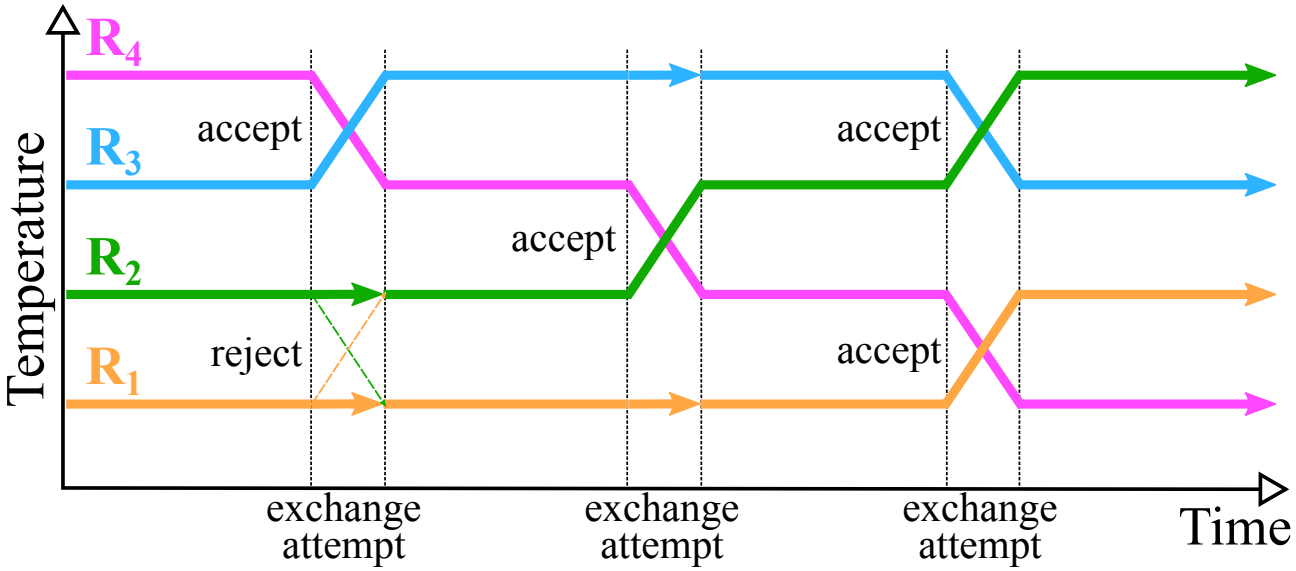


Figure 2.5: Replica exchange molecular dynamics (REMD) methodology illustrated for four replica.

One of the most successful ways of including MC in MD simulations is through replica exchange (REMD). Here, multiple simulations at different temperatures are run in parallel and an MC step is used to frequently exchange configurations between different temperatures (see Figure 2.5). The higher temperatures are used to accelerate the crossing of local energy barriers. The success of this technique depends on several factors, including the chosen temperatures range, the simulation length and the number of replicas (which must be proportional to the number of available processors). A critical parameter is the average exchange probability between each replica, defined as

$$p_{(1\to 2)} = \min\left(1, \exp\left[\left(\frac{1}{k_B T_1} - \frac{1}{k_B T_2}\right)(U_1 - U_2)\right]\right), \tag{2.20}$$

where T_1 and T_2 are the reference temperatures and U_1 and U_2 are the instantaneous potential energies of replica 1 and 2 respectively. $p_{(1\to2)}$ should be kept constant across the replicas and this is usually achieved by choosing an exponential temperature distribution. Patriksson and van der Spoel[55] proposed an algorithm to generate a set of temperatures for use in REMD simulations given a desired exchange probability $p_{(1\to2),des}$. The algorithm is available for online use.

Mass induced accelerated dynamics

For properties that depend only on particle positions, the ensemble average is independent of particle masses. Therefore, unrealistic masses, which narrow the range of motion timescales in a simulation, can be chosen as an alternative route to enhance the efficiency of conformational sampling. The physical motivation behind this approach lies in the reduction of the solvent viscosity due to linear scaling in the solvent mass (as $M_S = \lambda m$, where m is the default solvent mass and λ the scaling factor). The lower viscosity ($\sim\sqrt{\lambda}$) is, indeed, able to speed up the peptide folding dynamics and enhance the conformational sampling while other phenomena take place.

This method was proposed by Lin and Tuckerman[56], via variation of the masses of solvent and protein side-chains to sample the conformational phase space of a 9-residue peptide. A careful choice of solvent and solute masses can lead to a decoupling of their relative motion and a reduction in the solvent viscosity. Protein backbone and side-chain motions can also be decoupled in the same way.

2.3.3 Simulating rare events

Many biological processes are characterised by free energy barriers which are prohibitively large ($\sim$ 10 kJ mol^{-1}), and cannot be sampled within the timescale of a standard MD simulation. Therefore several methods have been developed to promote the sampling of specific events. Most of these methods involve the definition of a reaction coordinate, $\zeta\left(\mathbf{r}^N\right)$, which should

change smoothly and monotonically over the course of the event. The choice of ζ is of critical importance since it needs to fully capture the process of interest, *e.g.* a conformational change. Examples of some possible order parameters used to study protein folding include root mean squared deviation from the native state structure, end to end distance and fraction of native contacts formed.

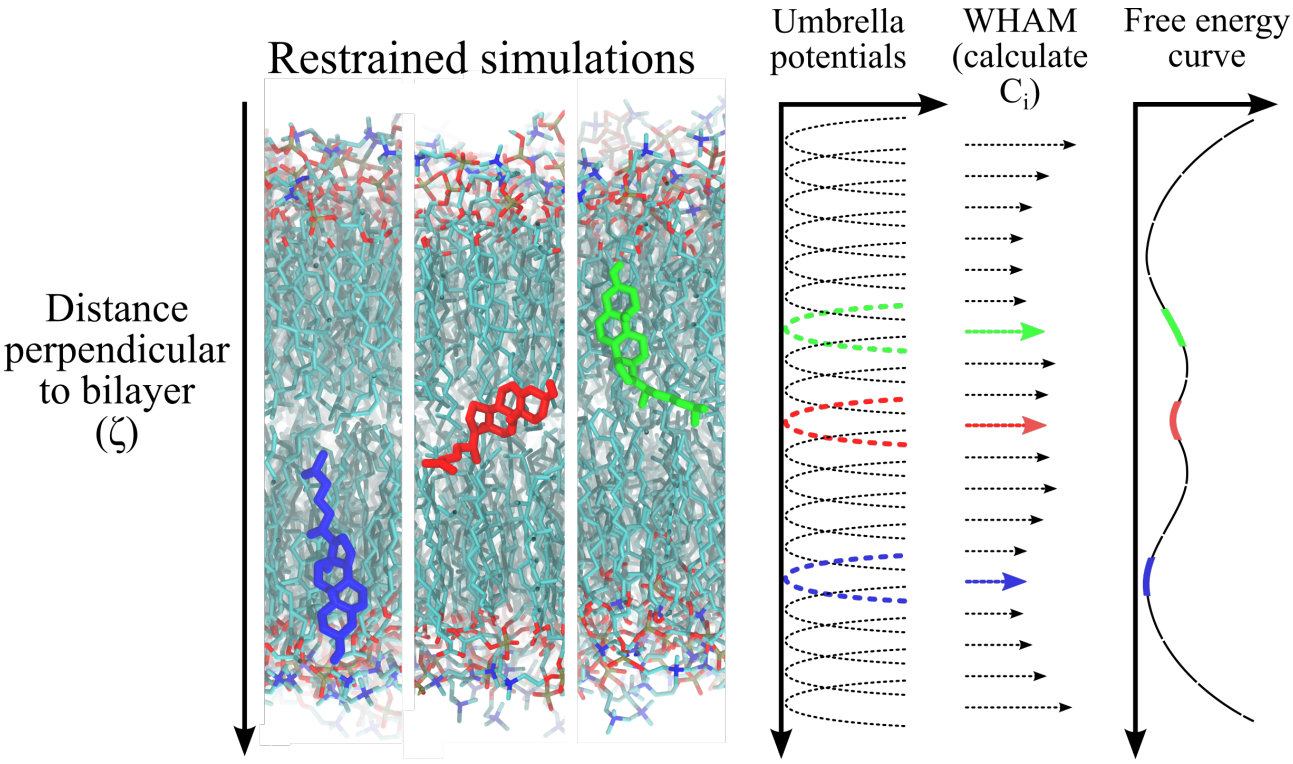


Figure 2.6: Illustration of the umbrella sampling (US) procedure used to determine the free energy barrier of cholesterol flip-flop and a sketch of the WHAM derived free energy curve. Three representative snapshots from restrained simulations using the umbrella potentials are also included.

Adding a restraining potential to the system Hamiltonian is one of the simplest ways to preferentially sample states in a particular region of phase space. In the umbrella sampling (US) method[57], the transition between two states is divided into a number of "windows", each characterised by a different value of the reaction coordinate, $\zeta\left(\mathbf{r}^N\right)$. A separate simulation is run for each window with a, usually harmonic, restraining potential that biases the simulation to sample states within that window, based on the value of $\zeta\left(\mathbf{r}^N\right)$. The effect of the bias potential, $u_i\left(\zeta\right)$, can then be removed to recover the unbiased free energy curve (a_i) for window i, by using

$$a_i\left(\zeta\right) = -k_B T \log\left(P_{bias,i}\left(\zeta\right)\right) - u_i\left(\zeta\right) + C_i, \tag{2.21}$$

where $P_{bias,i}(\zeta)$ represents the biased probability of visiting states along ζ, obtained from histograms computed from simulation i and C_i is a window dependent constant that shifts a_i so that it lines up with the free energy curves from other windows. The weighted histogram analysis method (WHAM)[58] provides an iterative method to obtain self-consistent values for C_i which minimises the error between windows. Figure 2.6 illustrates the application of this technique to studying cholesterol flip-flop.

2.4 MD simulation *vs.* experiment

Thanks to computer hardware improvements, coarse-grained model development and increased efficiency of the available algorithms, the length and timescales reachable with simulations have increased exponentially over the years, allowing a closer connection with experiments. A mutual interlinking between theoretical models, computer simulations and experimental techniques represents, therefore, the best approach to give full insight into the complexity of biological structures.

The first MD simulation in the biophysics field was run in the 1970's[59] and showed the ability of computational methods to shed light on a controversial topic like protein flexibility. Since then, comparison with experimental data has always been used as a starting-point to check the reliability of simulation outcomes and to improve forcefield parametrisation. More recently simulations have also been employed to interpret and give a molecular, microscopic explanation of the macroscopic behaviour observed in experiments. This is especially valuable when applied to contradictory experimental data from different techniques, for example, combining seemingly conflicting NMR and X-ray data to produce a unified model for Cyclophin's catalytic mechanism[60] or to compare solid-state and solution NMR data to determine the structure of the membrane protein gramicidin A[61]. The role of simulation in experimental work is becoming more and more significant: most structure determination by NMR or scattering techniques relies on some form of computational refinement step, *e.g.* simulated annealing. The accuracy of MD complemented structure determination is in most cases higher than with other techniques[62]. The most desirable simulation outcome is to be able to go beyond a simple re-

production of experimental results and provide deeper insight and relevant predictions, which can suggest and inspire new experimental investigations. In this regard, the key point to foster successful experimental/simulation collaborations and develop novel combined approaches to tackle new questions in biophysics is the mutual understanding of the different limitations and practicalities associated with both techniques and awareness of the corresponding assumptions and approximations made[63].

Simulations can be used alongside a wide range of experimental methods. In this section the attention will be focused on the interaction between simulations and experiments in the area of NMR and diffraction techniques, where simulations are an integral part of the analysis and interpretation of the data.

2.4.1 NMR and MD: structure and dynamics

A rich variety of parameters obtained from NMR, for protein and lipid membrane systems, can be expressed as a function of particle positions and momenta and, therefore, easily compared with MD simulations. Such parameters have widely been used for checking the reliability of simulation forcefields, as well as to validate and interpret experimental results, especially when the experimental data does not lead to a unique determination of the samples structure. The "NMRlipid project" is an example of such a collaboration. It provides a platform for discussions of experimental and simulation data and can be accessed through an online repository[64].

Some of the most widely quoted lipid parameters from solid state NMR experiments[65] are bond order parameters which indicate the orientation of particular bonds within lipid molecules and are especially useful for determining the degree of order in the lipid hydrocarbon tails, which show distinctively different values depending on the membrane state, *e.g.* gel or fluid. The C-H bond order parameter is defined as:

$$S_{CD} = \frac{1}{2}\langle\left(3\cos^2\theta - 1\right)\rangle, \tag{2.22}$$

where θ is the instantaneous angle between the C-H bond and the bilayer normal. This definition

is strictly related to its experimental measurement which can be obtained from the quadrupolar splitting, $\Delta\upsilon_Q$ in ^{2}H NMR:

$$|S_{CD}| = \frac{2}{3}\frac{\hbar}{e^2 q Q_D}\Delta\upsilon_Q, \tag{2.23}$$

where $\hbar$ is the Planck's constant, e is the elementary charge, q the electric field gradient and Q_D the deuteron quadrupole moment. Alternatively, it can be also accessed from dipolar splitting for ^{13}C NMR. In AA simulations, the order parameter can be calculated immediately from its definition (*i.e.* equation 2.22). In UA simulations, it is calculated via its relationship related to the C-C bond vectors[66]. Agreement with experiment is in general good for most forcefields, although several differences still persist in the description of the head group region due to difficulties in describing the water-lipid interface.

Regarding proteins, there are also a range of properties that provide a source of comparison between computer simulations and experiments. J-couplings can be converted into dihedral angles using the Karplus relations[67] and compared with average backbone dihedral angles obtained in simulations, to analyse secondary structure. In addition, the Nuclear Overhauser Effect (NOE) can probe the spatial relationship between hydrogens, giving insight on the tertiary structure. Order parameters (indicated with S_2) give dynamic information on protein flexibility based on the angular distributions of bond vector orientations. These properties are often used for forcefield parametrisation[68] and validation[69].

The power of NMR comes from its ability to study structure and dynamics simultaneously. MD simulations are well suited to capturing the conformational flexibility of molecules and they can be used to generate an ensemble of structures which can then be compared with the experimental data. This idea has prompted the development of an integrated NMR-MD approach that has been realised in the past few years. The state of the art of such an approach is discussed in a recent review by Mocci and Laaksonen[70].

Along these lines, Lindorff-Larsen *et al.*[71] have pioneered a new method for structure determination, called Dynamic Ensemble Refinement (DER). In this approach the data from NMR experiments (S_2 and NOEs) are used as restraints to guide the MD simulations. Rather than applying the restraints to every configuration in the trajectory, the restraints are imposed on

the ensemble as a whole. This idea is realised by modifying the restraints during the simulation to ensure that on average, the ensemble will reproduce the NMR data. The advantage of this simulation approach is that it can capture both structural and dynamical information simultaneously since the restraints are applied to the structure, via NOEs, and dynamics, via S_2, in the same simulation.

2.4.2 Structural features from diffraction – solving the phase problem with MD

X-ray diffraction was used to resolve the first protein structures[72] and is still the most commonly used tool for protein structure determination[41]. Lipid membranes are also commonly probed using scattering techniques[73]. However, in biologically relevant environments, membranes are in a fluid state and the lack of order makes structure determination via scattering difficult. In general, scattering techniques cannot provide an unambiguous result due to the so-called "phase problem". Scattering experiments measure the intensity of scattered radiation, $I(q)$, as a function of wavevector q. $I(q)$ is defined as the squared modulus of the form factor, $F(q)$, which is connected to the electron density profile, $\rho_e(r)$, in X-ray scattering (or neutron scattering length density, $\rho_{NSLD}(r)$, in neutron scattering) via the Fourier transform:

$$F(q) = \int_V \rho_X(r)\, e^{iq,r} dV. \tag{2.24}$$

However, obtaining the density profiles from an inverse Fourier transform of the form factor is not possible because experiments only yield the modulus of the form factor, so its amplitude is known, but not its phase. To overcome this problem, scattering data must be fit to a model structure, whose density function depends on fitting parameters which can be derived via comparison with experimental form factors.

Various simple models of lipid membrane structure have been proposed for this purpose[74, 75, 76]. MD can be of great help in this instance, by proposing structures that reduce the number of fitting parameters. An example of this idea is the H2 model, proposed by Klauda *et al.*[77].

Alternatively, it is possible to directly compare form factors derived from simulations with the raw experimental data obtained in reciprocal space[78]. To calculate $F(q)$ of membranes using MD, the number density profile, $N_\alpha(z)$, perpendicular to the membrane has to be calculated. $N_\alpha(z)$ can be easily obtained from a simulation by dividing the simulation box into a number of slices running parallel to the membrane surface (along z) and counting the number of atoms of type α in each box. Then, to compare with X-ray data, $N_\alpha(z)$ is multiplied by the number of electrons in each atom ($n_{e\alpha}$) to obtain $\rho_e(r)$, and finally Fourier transformed to give $F(q)$. To compare with neutron scattering, $N_\alpha(z)$ is multiplied by the neutron scattering length (NSL) to get $\rho_{NSLD}(r)$.

The process of generating X-ray and neutron data from simulations has been automated via the SIMtoEXP software package[79]. This software uses as input the number densities of each atom type, $N_\alpha(z)$. It then calculates $\rho_e(r)$ and $\rho_{NSLD}(r)$ and the form factors, $F_e(q)$ and $F_{NSLD}(q)$. Raw experimental data can also be input alongside simulation data, or obtained from the growing repository of experimental data incorporated in the software.

2.5 Data analysis

Several properties of lipid bilayers and monolayers, nanoparticles and peptides are computed in this thesis. In this section I explain the background theory behind some of the quantitative parameters computed in this thesis. The methods sections in each chapter contain more detailed explanation of the computed quantities, their specific application to the systems under investigation and any new methods which I have developed especially for this work.

2.5.1 Lipid properties

Radial distribution function

In order to extract information regarding the structural properties of the different systems, the radial distribution function, $RDF(r)$, has been used throughout this thesis. The $RDF(r)$

quantifies the probability of finding a group of particles of type B at a distance, r, from a given other group (A), and is computed from:

$$RDF\left(r\right)_{AB} = \frac{\left\langle \rho_B\left(r\right)\right\rangle}{\left\langle \rho_B\right\rangle_{local}} = \frac{1}{\left\langle \rho_B\right\rangle_{local} N_A} \sum_{i\in A}^{N_A} \sum_{j\in B}^{N_B} \frac{\delta\left(r_{ij} - r\right)}{4\pi r^2}, \tag{2.25}$$

where $\left\langle \rho_B\left(r\right)\right\rangle$ is the average density of particles of type B at distance r around A, $\left\langle \rho_B\right\rangle_{local}$ is the average density of particle B inside the maximum radius, N_A and N_B the numbers of particles A and B and r_{ij} the distance between individual A-B pairs. Note that A and B can be different or identical groups of particles and that r stands for a differential spherical slice, rather than an absolute value.

The radial distribution function is very useful in determining the phase behaviour of a substance. The sharper the peaks and troughs in the RDF, the closer to a solid-ordered state. By integrating the RDF up to the first minimum, it is possible to compute the number of neighbouring particles, known as the coordination number. In this thesis the RDF is used to discuss the intermolecular interaction of triglycerides in the fluid phase (chapter 3), and to discuss the phase of lipid monolayers (chapter 3 and chapter 7).

Voronoi tesselation method

The Voronoi tessellation associates a set of points to their closest regions of space (see example in Figure 2.7), and was originally proposed in 1908 by the Ukrainian mathematician George Voronoi[80]. The domain partition obtained through Voronoi tessellation has many applications in problems which involves allocating spaces between groups of objects. For example, it provides a powerful approach to investigate the geometric structure of liquids, the secondary structures of polypeptides[81], and to analyse the lateral organisation of the bilayer structure in biological membranes[82] and around proteins[83].

Definitions - Given a set of n points, $S = \{s_1,\ s_2,\ s_3,\ ...,\ s_n\}$, in $\mathbb{R}^{\mathrm{d}}$ (also known as seeds or generators) the Voronoi tessellation is the partition of $\mathbb{R}^{\mathrm{d}}$ into n polyhedral regions V_i (also

known as Voronoi cells). A Voronoi cell is a convex polyhedron containing the region of space, $\{r \in \mathbb{R}^d\}$, closer to the point s_i than to any other point s_j, so that:

$$V_i = \{r \in \mathbb{R}^d \mid d(r, s_i) \leq d(r, s_j), \forall j \neq i\}, \tag{2.26}$$

where $d(...)$ is the Euclidean distance function. Other distance functions can also be used to obtain variations of the Voronoi diagram. The set of all Voronoi cells and their faces forms a cell complex (see Figure 2.7 for examples in 2D and 3D). There are many parameters which can describe a Voronoi cell, such as coordinates of the generating points and vertices, surface area or volume of the cell, number of facets, number of neighbouring cells. These parameters can be used to obtain a range of information about the properties of the cell itself and its relationship with the surrounding cells in the space.

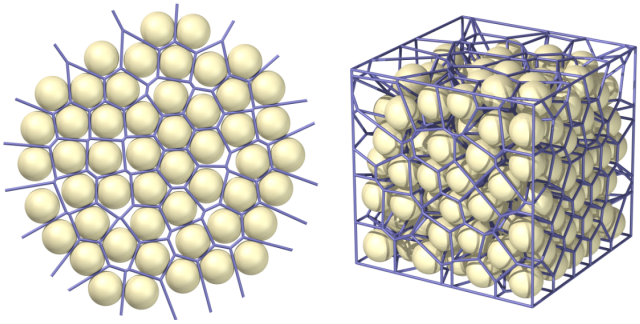


Figure 2.7: Examples of Voronoi tessellation in 2D and 3D.

Qhull *vs.* voro++ There are several open-source software packages to compute the Voronoi tessellation. Within this thesis, QHull[84, 85] is used for obtaining information about the lateral organisation of SC bilayers in 2D and voro++[86] is used to analyse the sebum triglyceride arrangement. The Qhull routine is used in many computational languages such as Matlab and Python. Qhull makes use of an indirect projection method which computes the Voronoi diagram as a single object, specifically via Delaunay triangulation, *i.e.* by determining the circumcenters of the triangulation facets which become the vertices of the Voronoi diagram[84]. On the other hand, voro++ is based on a direct method whereby each cell is computed individually, providing a higher degree of flexibility. The single-cell computation stores the position of edges and vertices of the cell and updates them by repeatedly cutting the cell by planes corresponding to

the bisectors of neighbouring particles[86].

Lateral pressure profile

A microscopic definition of the stress tensor starting from atomic and molecular interactions is challenging. Irving and Kirkwood[87] have led the way in linking different scales through non equilibrium classical statistical mechanics, providing a local definition of the stress tensor, $\sigma(x)$, as a pointwise continuum field, from the position and velocities of individual particles. Their theory was later developed by Noll[88], resulting into the commonly named Irving-Kirkwood-Noll (IKV) theory.

Applying the IKN procedure to derive statistical and pointwise fields from atomistic simulation, requires that every point within the volume of the system is well sampled, which is practically unfeasible. Thus in MD simulations at equilibrium, the IKN pointwise stress can be evaluated from time averages and further averaged spatially, resulting in the so-called Hardy-Murdoch stress tensor[89, 90]. It can be calculated from the summation of a kinetic contribution, $\sigma^K(\mathbf{x})$, due to the flux of momentum associated with the vibrational internal energy of the system, and a potential contribution, $\sigma^V(\mathbf{x})$, that stems from the internal forces between particles, as:

$$\sigma(\mathbf{x}) = \sigma^K(\mathbf{x}) + \sigma^V(\mathbf{x}); \tag{2.27}$$

$$\sigma^K(\mathbf{x}) = -\frac{1}{N_T}\sum_{i=1}^{N_T}\left[\sum_{\alpha} m^\alpha w\left(\mathbf{x};\mathbf{r}_i^\alpha - \mathbf{x}\right)\mathbf{v}_i^\alpha \otimes \mathbf{v}_i^\alpha\right]; \tag{2.28}$$

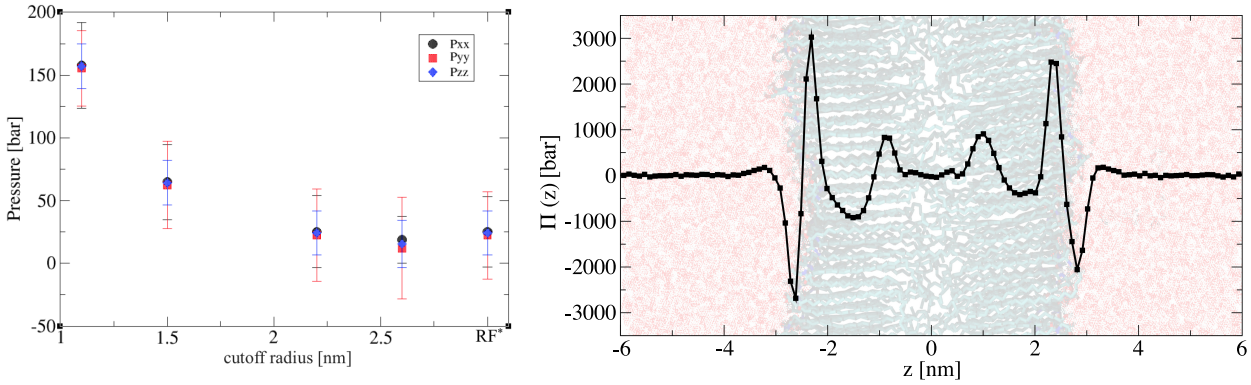
$$\sigma^V(\mathbf{x}) = -\frac{1}{2N_T}\sum_{i=1}^{N_T}\left[\sum_{\alpha,\beta(\neq\alpha)} \mathbf{f}_i^{\alpha\beta} \otimes \mathbf{r}_i^{\alpha\beta} B\left(\mathbf{x};\mathbf{r}_i^\alpha,\mathbf{r}_i^\beta\right)\right]; \tag{2.29}$$

where m^α, $\mathbf{r}_i^\alpha$, $\mathbf{v}_i^\alpha$ and $\mathbf{f}_i^{\alpha\beta}$ are the masses, positions, velocities and pairwise forces at the time-step i; N_T is the total number of time-steps and the symbol $\otimes$ denotes the dyadic product. The spatial average is operated by the weight function $w(\mathbf{x};\mathbf{y})$, supported in a domain Ω_x centred at $\mathbf{x}$, which weights the contribution to the stress from the particle located at $\mathbf{y}$ inside the domain. The bond function $B\left(\mathbf{x};\mathbf{r}_i^\alpha,\mathbf{r}_i^\beta\right)$ results from the integration of $w(\mathbf{x};\mathbf{y})$ over the line segment connecting particles α and β to account for their interaction.

Particular attention has to be given to the decomposition of the forces acting on a particle into pairwise contributions. The decomposition should be central (central force decomposition - CFD) with each pairwise force aligned with the vector connecting the pair of particles in order to guarantee a symmetric stress tensor which satisfies the balance of angular momentum of a simple continuum medium.

An implementation of the local stress calculation in MD simulations at equilibrium via the Hardy-Murdoch procedure was developed by Vanegas *et al.*[91] and is available for the GROMACS simulation package[30]. This implementation systematically treats constraints and applies the central force decomposition in order to construct a symmetric stress tensor. A weakness of this implementation lies in the evaluation of long-range electrostatic interactions, in that the Coulomb contribution includes only electrostatic forces up to a given cutoff radius, because the implementation does not take into account the electrostatic contributions computed in reciprocal space. This approximation might result in an unphysical, positive contribution to the stress tensor components lying on the main diagonal of the stress tensor.

To check the reliability of this implementation a systematic evaluation of the pressure profile through atomistic MD simulations of a box of water, made up of 5187 molecules (SPC water model) was performed, varying the cutoff for the computation of the stress tensor contribution due to Coulomb forces. The analysis was carried out considering four different values of the cutoff radius: 1.1, 1.5, 2.2 and 2.6 nm. It is important to note that this does not change the original trajectory, but the way forces are calculated when analysing each frame. The interactions between water molecules were described with the SPC model. Lennard-Jones forces were calculated with a plain cutoff of 1.0 nm. Long-range electrostatic interactions were computed using the particle-mesh Ewald (PME) method with a real-space cutoff of 1.0 nm and a Fourier grid spacing of 0.15 nm. The temperature was kept constant at 350 K with a Nosé-Hoover thermostat, whilst the pressure was isotropically coupled with a Parrinello-Rahman barostat at 1 bar. The integration time step for the atomistic simulations was 2 fs. The simulation was equilibrated for 10 ns and run for 10 ns, storing positions and velocities every 2 ps. Figure 2.8a shows the dependence of the stress tensor components (σ_{xx}, σ_{yy}, σ_{zz} represented in the graph by the corresponding pressure components $P_{ij} = -\sigma_{ij}$,) with respect



(a) Pressure *vs.* cutoff **(b)** CER2 pressure profile

Figure 2.8: (a) Pressure tensor components as a function of the Coulombic cutoff radius (1.1, 1.5, 2.2 and 2.6 nm). The point denoted as RF^* was computed using the reaction-field method, with a cut-off of 2.2 nm for the electrostatic interactions and a dielectric constant equal to $\epsilon_{RF} = 78.5$. (b) Lateral pressure profile, $\Pi(z) = P_T(z) - P_N(z)$, of a ceramide 2 (CER2) lipid bilayer, overlaid on a simulation snapshot. The peaks and troughs of the pressure profile reflect structural features of the bilayer.

to the cutoff radius. The point denoted as RF, instead, was computed using the reaction-field method, with a cut-off of 2.2 nm for the electrostatic interactions and a dielectric constant equal to $\epsilon_{RF} = 78.5$. Using cutoff radii lower than 2.2 nm results in an artificial stress within the system, which decreases by increasing the cutoff radius. However, small differences are observed beyond a cutoff of 2.2 nm, after which the computational cost increases very rapidly.

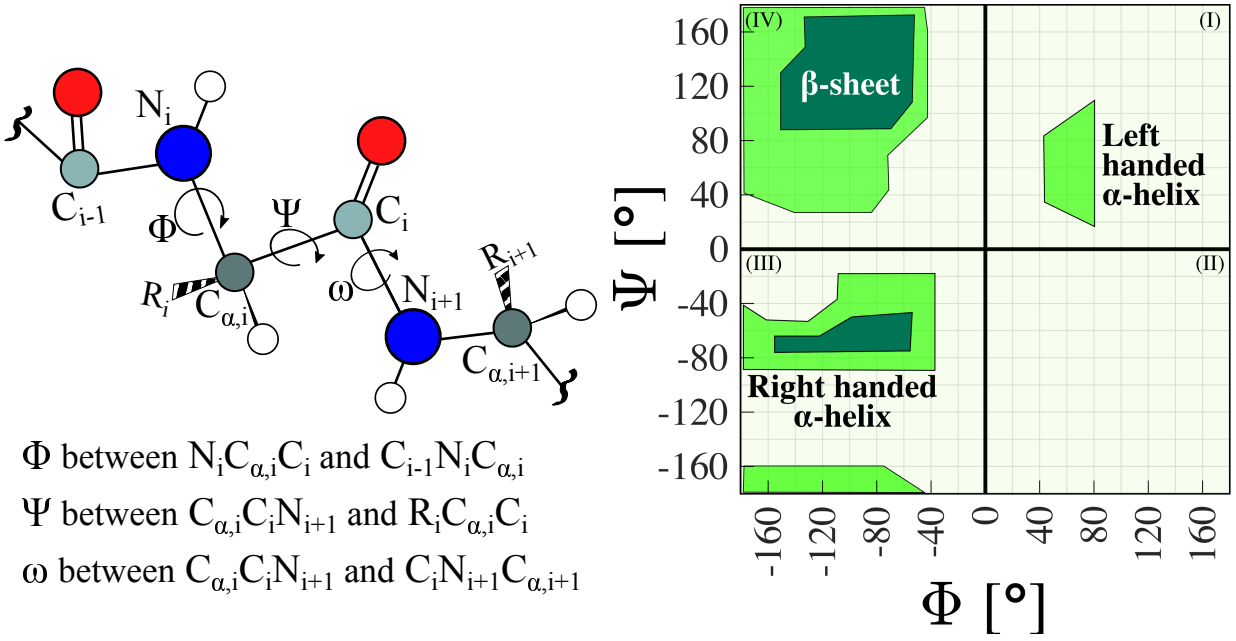
A cutoff of 2.2 nm was then used to obtain the elastic properties of a ceramide 2 lipid bilayer (see Figure 2.8b). The bilayer was simulated in water for 50 ns with the same simulation parameters as in ref. 92. The local pressure was obtained from the contributions of the local stress tangential, $P_T(z)$, and normal, $P_N(z)$, to the membrane plane: $P(z) = P_T(z) - P_N(z)$. The obtained lateral pressure profile (see Figure 2.8b) is in good agreement with the one proposed by Das *et al.*[92].

2.5.2 Peptide properties

Ramachandran plot

The secondary structure of proteins is determined by the backbone dihedral angles, Φ, Ψ and ω (see Figure 2.9a). They are defined as the tortional angles between the normals to the

planes defined by suitable consecutive triplets of atoms as specified in Figure 2.9a. Due to the rigidity of the peptide bond (*e.g.* C_i–N_{i+1}), ω normally measures ∼180° (*i.e.* the partial double-bond character keeps the peptide bond planar). Therefore, Ramachandran *et al.*[93] proposed a method to visualise the protein structure in a 2-dimensional scatter plot of Φ, Ψ, which is known as a Ramachandran plot (or Ramanchadran diagram). An example of the diagram where the value of the angles are related to the resulting peptide secondary structures is represented in Figure 2.9b.



(a) Torsion angles for peptide backbone. **(b)** Ramachandran plot

Figure 2.9: a) Definition of the protein backbone dihedral angles, Φ, Ψ and ω. b) Schematic representation of a Ramachandran plot. Dark green and mid-green areas represent more commonly populated regions of the Ramachandran plot, which correspond to specific elements of secondary structure.

The lightest region in the graph corresponds to conformations with a very low probability to occur, due to steric clashes between atoms. The Φ and Ψ bond angles for each residue in the α-helical structure are very similar and measure $\sim -60°$, conferring regularity to the secondary structure. Similarly, the Φ and Ψ bond angles for every residue in a β strand are similar. Since the residues in a β strand are in an extended form, the Φ and Ψ angles in this conformation are close to $\pm 120°$.

The computation of the Ramachandran plot was used in chapter 7 to discuss the secondary structure of an antimicrobial peptide.

Radius of gyration

The radius of gyration, R_G, gives a measure of the characteristic size of a molecule, and is often used to check whether a protein is stable (relatively constant value of R_G) or if it is unfolding (R_G variation over time).

For a structure composed of N particles, with centre of mass located at $\mathbf{r}_{COM}$, R_G is defined as the root mean square distance of each particle to the centre of mass:

$$R_G = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{r}_i - \mathbf{r}_{COM})^2}, \tag{2.30}$$

where $\mathbf{r}_i$ is the position of particle i.

RMSD

The root-mean-square deviation (*RMSD*) is a measure of the average distance between the atoms of different configurations of the same protein, superimposed on each other (generally computed considering just the backbone C_α atoms) or of small organic molecules. It is generally used to compare a predicted structure to a known reference structure. Also, it can be used as a reaction coordinate to study a protein folding process. In the case of small organic molecules, *RMSD* is often computed to study the configuration of a ligand when binding to a macromolecule.

The *RMSD* is computed as:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2}, \tag{2.31}$$

where δ_i is the distance between atom i and either a reference structure or the mean position of the N equivalent atoms.

Bibliography to chapter 2

[1] B. J. Alder and T. E. Wainwright, *The Journal of Chemical Physics*, 1957, **27**, 1208–1209.

[2] B. J. Alder and T. E. Wainwright, *The Journal of Chemical Physics*, 1959, **31**, 459–466.

[3] R. Notman, W. K. den Otter, M. G. Noro, W. J. Briels and J. Anwar, *Biophysical Journal*, 2007, **93**, 2056–2068.

[4] S. Guo, T. C. Moore, C. R. Iacovella, L. A. Strickland and C. McCabe, *Journal of Chemical Theory and Computation*, 2013, **9**, 5116–5126.

[5] Z. Sovova, K. Berka, M. Otyepka and P. Jurečka, *Journal of Physical Chemistry B*, 2015, **119**, 3988–3998.

[6] R. Notman and J. Anwar, *Advanced Drug Delivery Reviews*, 2013, **65**, 237–250.

[7] C. Das, P. D. Olmsted and M. G. Noro, *Soft Matter*, 2009, **5**, 4549–4555.

[8] C. Das, M. G. Noro and P. D. Olmsted, *Soft Matter*, 2014, **10**, 7346–7352.

[9] C. Das, M. G. Noro and P. D. Olmsted, *Physical Review Letters*, 2013, **111**, 1–5.

[10] A. Del Regno and R. Notman, *Physical Chemistry Chemical Physics*, 2018, **20**, 2162–2174.

[11] T. C. Moore, C. R. Iacovella, R. Hartkamp, A. L. Bunge and C. McCabe, *Journal of Physical Chemistry B*, 2016, **120**, 9944–9958.

[12] R. Notman, J. Anwar, W. J. Briels, M. G. Noro and W. K. den Otter, *Biophysical Journal*, 2008, **95**, 4763–71.

[13] R. Thind, D. W. O'Neill, A. Del Regno and R. Notman, *Chem. Commun.*, 2015, **51**, 5406–5409.

[14] A. Akinshina, C. Das and M. G. Noro, *Phys. Chem. Chem. Phys.*, 2016, **18**, 17446–17460.

[15] R. Gupta and B. Rai, *Journal of Physical Chemistry B*, 2016, **120**, 7133–7142.

[16] R. Gupta and B. Rai, *Scientific Reports*, 2017, **7**, 1–13.

[17] S. Nangia and R. Sureshkumar, *Langmuir*, 2012, **28**, 17666–17671.

[18] R. C. Van Lehn and A. Alexander-Katz, *Soft Matter*, 2011, **7**, 11392–11404.

[19] R. C. Van Lehn, M. Ricci, P. H. Silva, P. Andreozzi, J. Reguera, K. Voïtchovsky, F. Stellacci and A. Alexander-Katz, *Nature Communications*, 2014, **5**, 1–11.

[20] C. L. Ting and Z. G. Wang, *Soft Matter*, 2012, **8**, 12066–12071.

[21] M. B. Ulmschneider, J. P. F. Doux, J. A. Killian, J. C. Smith and J. P. Ulmschneider, *Journal of the American Chemical Society*, 2010, **132**, 3452–3460.

[22] J. P. Ulmschneider, J. C. Smith, S. H. White and M. B. Ulmschneider, *Journal of the*

American Chemical Society, 2011, **133**, 15487–15495.

[23] Y. Wang, C. H. Chen, D. Hu, M. B. Ulmschneider and J. P. Ulmschneider, *Nature Communications*, 2016, **7**, 1–9.

[24] L. Thøgersen, B. Schiøtt, T. Vosegaard, N. C. Nielsen and E. Tajkhorshid, *Biophysical Journal*, 2008, **95**, 4337–4347.

[25] J. W. Carter, A. S. Tascini, J. M. Seddon and F. Bresme, *Computational Tools for Chemical Biology*, The Royal Society of Chemistry, 2018, ch. 2, pp. 39–68.

[26] L. Verlet, *Phys. Rev.*, 1967, **159**, 98.

[27] W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson, *The Journal of Chemical Physics*, 1982, **76**, 637.

[28] R. Hockney, S. Goel and J. Eastwood, *Journal of Computational Physics*, 1974, **14**, 148–158.

[29] W. F. van Gunsteren and H. J. C. Berendsen, *Molecular Simulation*, 1988, **1**, 173–185.

[30] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *Journal of Computational Chemistry*, 2005, **26**, 1701–1718.

[31] M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford University Press, 2010, p. 713.

[32] M. P. Allen and D. J. Tildesley, *Computer Simulation of liquids*, Oxford University Press, 1987, p. 640.

[33] H. J. Berendsen, J. P. Postma, W. F. van Gunsteren, A. Dinola and J. R. Haak, *The Journal of Chemical Physics*, 1984, **81**, 3684–3690.

[34] G. Bussi, D. Donadio and M. Parrinello, *Journal of Chemical Physics*, 2007, **126**, 014101.

[35] S. Nosé, *Molecular Physics*, 1984, **52**, 255–268.

[36] W. G. Hoover, *Physical Review A*, 1985, **31**, 1695–1697.

[37] D. van der Spoel, E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, K. A. Feenstra, R. van Drunen and H. J. C. Berendsen, *GROMACS user manual*, 2010, p. 7778.

[38] M. Parrinello and A. Rahman, *Journal of Applied Physics*, 1981, **52**, 7182–7190.

[39] S. Nosé and M. Klein, *Molecular Physics*, 1983, **50**, 1055–1076.

[40] J. Domański, P. J. Stansfeld, M. S. P. Sansom and O. Beckstein, *Journal of Membrane Biology*, 2010, **236**, 255–258.

[41] RCSB PDC - Protein Data Bank, *PDB Data Distribution by Experimental Method and Molecular Type*, `https://www.rcsb.org/stats/summary`.

[42] B. Hess, H. Bekker, H. J. C. Berendsen and J. G. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.

[43] J. P. Ryckaert and A. Bellemans, *Chemical Physics Letters*, 1975, **30**, 123–125.

[44] J. Israelachvili, *Intermolecular and Surface Forces*, Academic Press, 2011, p. 704.

[45] M. R. Shirts, D. L. Mobley, J. D. Chodera and V. S. Pande, *The journal of physical chemistry. B*, 2007, **111**, 13052–63.

[46] P. H. Hünenberger and J. A. McCammon, *Journal of Chemical Physics*, 1999, **110**, 1856–1872.

[47] G. Malloci, A. V. Vargiu, G. Serra, A. Bosin, P. Ruggerone and M. Ceccarelli, *Molecules*, 2015, **20**, 13997–14021.

[48] O. Berger, O. Edholm and F. Jähnig, *Biophysical Journal*, 1997, **72**, 2002–2013.

[49] W. L. Jorgensen and J. Tirado-Rives, *Journal of the American Chemical Society*, 1988, **110**, 1657–1666.

[50] F. W. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott and I. G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, 1996, pp. 1–1042.

[51] W. G. Noid, J. W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das and H. C. Andersen, *Journal of Chemical Physics*, 2008, **128**, 244114.

[52] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *The Journal of Physical Chemistry B*, 2007, **111**, 7812–7824.

[53] A. De Ruiter, S. Boresch and C. Oostenbrink, *Journal of Computational Chemistry*, 2013, **34**, 1024–1034.

[54] J. De Joannis, P. S. Coppock, F. Yin, M. Mori, A. Zamorano and J. T. Kindt, *Journal of the American Chemical Society*, 2011, **133**, 3625–3634.

[55] A. Patriksson and D. van der Spoel, *Physical Chemistry Chemical Physics*, 2008, **10**, 2073.

[56] I. C. Lin and M. E. Tuckerman, *Journal of Physical Chemistry B*, 2010, **114**, 15935–15940.

[57] G. M. Torrie and J. P. Valleau, *Journal of Computational Physics*, 1977, **23**, 187–199.

[58] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *Journal of Computational Chemistry*, 1992, **13**, 1011–1021.

[59] J. A. McCammon, B. R. Gelin and M. Karplus, *Nature*, 1977, **267**, 585–590.

[60] D. Trzesniak and W. F. van Gunsteren, *Protein Sci.*, 2006, **15**, 2544–51.

[61] T. W. Allen, O. S. Andersen and B. Roux, *J Am Chem Soc*, 2003, **125**, 9868–9877.

[62] B. Mao, R. Tejero, D. Baker and G. T. Montelione, *Journal of the American Chemical Society*, 2014, **136**, 1893–906.

[63] V. Brázdová and D. Bowler, *Atomistic Computer Simulations: A Practical Guide*, 2013.

[64] *The NMRlipids project*, `http://nmrlipids.blogspot.com`.

[65] A. I. I. Tyler, J. A. Clarke, J. M. Seddon and R. V. Law, *Methods Mol Biol*, Springer Science, 2015, vol. 1232, pp. 227–253.

[66] J. J.-P. Douliez, a. Ferrarini and E. E.-J. Dufourc, *The Journal of Chemical Physics*, 1998, **109**, 2513.

[67] M. Karplus, *The Journal of Chemical Physics*, 1959, **30**, 11–15.

[68] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angewandte Chemie International Edition*, 1999, **38**, 236–240.

[69] O. F. Lange, D. Van Der Spoel and B. L. De Groot, *Biophysical Journal*, 2010, **99**, 647–655.

[70] F. Mocci and A. Laaksonen, *Nuclear Magnetic Resonance*, 2015, **44**, 592–616.

[71] K. Lindorff-Larsen, R. B. Best, M. A. DePristo, C. M. Dobson and M. Vendruscolo, *Nature*, 2005, **433**, 128–132.

[72] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff and D. C. Phillips, *Nature*, 1958, **181**, 662–666.

[73] A. I. Tyler, R. V. Law and J. M. Seddon, *Methods in Membrane Lipids: Second Edition*, 2014, pp. 199–225.

[74] M. H. Wilkins, A. E. Blaurock and D. M. Engelman, *Nature New Biology*, 1971, **230**, 72–76.

[75] M. C. Wiener, R. M. Suter and J. F. Nagle, *Biophysical Journal*, 1989, **55**, 315–325.

[76] M. C. Wiener and S. H. White, *Biophysical Journal*, 1991, **59**, 174–185.

[77] J. B. Klauda, N. Kucerka, B. R. Brooks, R. W. Pastor and J. F. Nagle, *Biophysical journal*, 2006, **90**, 2796–2807.

[78] R. W. Benz, F. Castro-Roman, D. J. Tobias and S. H. White, *Biophysical journal*, 2005, **88**, 805–817.

[79] N. Kučerka, J. Katsaras and J. F. Nagle, *The Journal of membrane biology*, 2010, **235**, 43–50.

[80] G. Voronoi, *Journal fur die Reine und Angewandte Mathematik*, 1908, **1908**, 198–287.

[81] F. Dupuis, J. F. Sadoc and J. P. Mornon, *Proteins-Structure Function and Bioinformatics*, 2004, **55**, 519–528.

[82] Y. Zhang, J. W. Carter, A. Lervik, N. J. Brooks, J. M. Seddon and F. Bresme, *Soft Matter*, 2016, **12**, 2108–2117.

[83] A. Lervik, F. Bresme and S. Kjelstrup, *Physical Chemistry Chemical Physics*, 2012, **14**, 3543–3553.

[84] *Qhull*, `http://www.qhull.org`.

[85] C. B. Barber, D. P. Dobkin and H. Huhdanpaa, *ACM Transactions on Mathematical Software*, 1996, **22**, 469–483.

[86] C. H. Rycroft, *Chaos*, 2009, **19**, 041111.

[87] J. H. Irving and J. G. Kirkwood, *The Journal of Chemical Physics*, 1950, **18**, 817.

[88] R. B. Lehoucq and A. Von Lilienfeld-Toal, *Journal of Elasticity*, 2010, **100**, 5–24.

[89] R. J. Hardy, *The Journal of Chemical Physics*, 1982, **76**, 622–628.

[90] A. I. Murdoch, *Quarterly Journal of Mechanics and Applied Mathematics*, 1983, **36**, 163–187.

[91] J. M. Vanegas, A. Torres-Sànchez and M. Arroyo, *Journal of Chemical Theory and Computation*, 2014, **10**, 691–702.

[92] C. Das, M. G. Noro and P. D. Olmsted, *Biophysical Journal*, 2009, **97**, 1941–1951.

[93] G. N. Ramachandran, C. Ramakrishnan and V. Sasisekharan, *Journal of Molecular Biology*, 1963, **7**, 95–99.

Chapter 3

On the top of our skin: the sebum oily layer

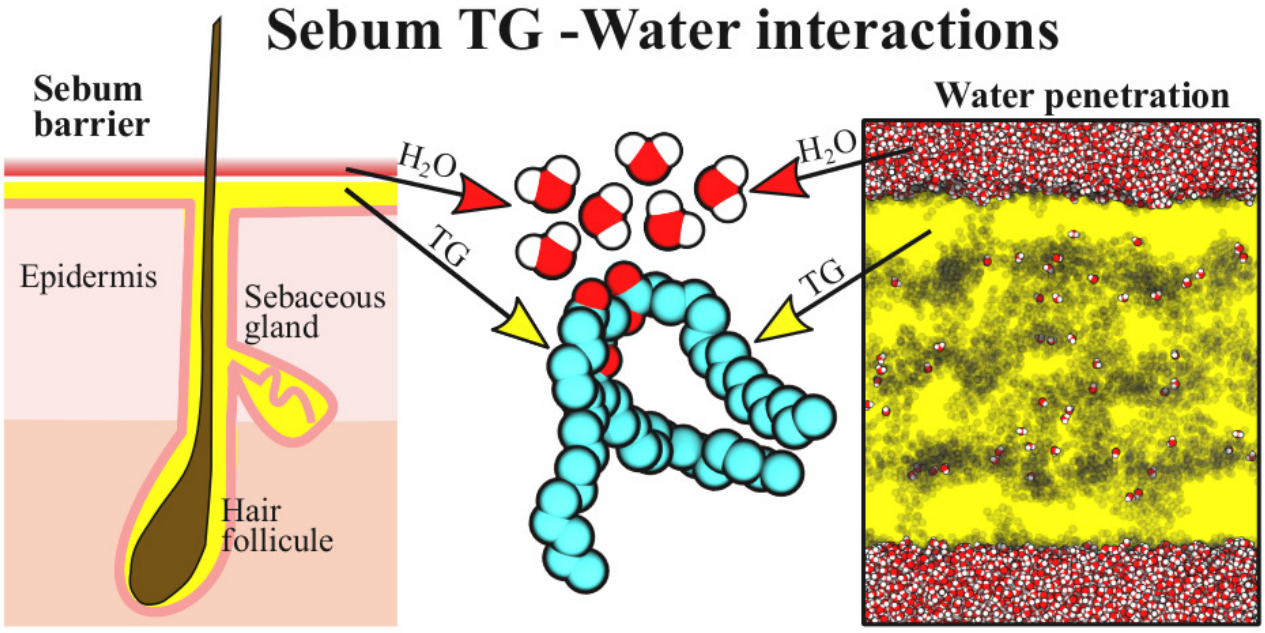


Figure 3.1: Title page figure for chapter 3. This picture has been used as Table of contents for our publication: Tascini *et al.*[1].

3.1 Introduction

The surface of our skin is covered by a thin oily layer called sebum (from the latin meaning *fat, tallow*). Generated in the sebaceous gland, the sebum oil diffuses through hair follicles to cover the surface of our skin. For many years, several studies have been aimed at understanding the connections between the human sebum and the onset of the disease acne[2, 3]. However, in the last decade, thanks to the development of *in vitro* models of the human sebaceous glands[4], the role of sebum oil has been reconsidered and it has been shown to take part in the regulation of the skin hydration and waterproofing properties[5, 6, 7]. In addition, it may confer antimicrobial properties to the skin and assist in the transport of antioxidants, such as vitamin E[8]. As suggested before by Smith and Thiboutot[9], whether the sebaceous gland lipids can be considered *"a friend or a foe"* is still an open question.

The sebum is made up of a mixture of lipids unique to the surface of the human body and, specifically, to this lipid layer. These lipids include triglycerides, wax esters, squalene and free fatty acids. The uniqueness of these lipids can be appreciated by considering the uncommon patterns of unsaturation and the wide range in the acyl chain length[10]. Triacyl glycerols consist of three fatty acid chains esterified to glycerol and are the most abundant component of the human sebum. The fraction of triglycerides in the sebum reported in the literature ranges widely from 20% to 60%. This variability in the reported measurements is due to the capability of some bacteria, such as *Propionibacterium acnes*, to hydrolyse triglyceride molecules to free fatty acids and glycerol[3], which results in a lower fraction of triglyceride (and higher presence of free fatty acids) in individuals afflicted by acne.

Depending on the structure and saturation level of the component fatty acid chains involved in the esterification, triglycerides show different structural and thermodynamic properties. In addition to being an active component of the sebum oil, triglycerides can be found in the human body fat present in the blood stream, where they enable the bidirectional transfer of adipose lipid and blood glucose from the liver via lipoprotein particles[11], and in the tear film, where neutral lipids are fundamental to maintaining the proper wetness and optical quality at the surface of the eye[12, 13]. They also appear in a wide range of products in the food industry

(*e.g.* milk and derivatives[14]), health care (*e.g.* cosmetics) and lubricant films[15]. In all these applications the investigation of the interaction of triglyceride molecules with water is relevant and it is also of critical importance in order to shed light on the role of sebum triglyceride molecules for the water regulation function of our skin surface.

Experimental studies on triglyceride molecules have been performed to investigate their solid and liquid phases. However, the phase behaviour of pure triglycerides from the most stable crystalline form to the liquid phase is complicated by the formation of metastable polymorphic phases, occurring below the melting point. In addition, despite the general agreement on the existence of some ordered structure in the triglyceride liquid phase (which is of great interest in describing the properties of oils, such as the human sebum) the interpretation of experimental results is not unique and different models have been proposed, as recently reviewed by Iwahashi and Kasahara[16]. The presence of small pseudo-lamellar domains (≈ 20 nm in size) in the liquid phase was first described by Larsson[17] and supported later by further experimental studies, such as ^{13}C NMR and Raman Spectroscopy[18]. Cebula *et al.*[19] proposed, instead, the presence of chair-like molecular conformations with oriented chains forming a nematic liquid crystal structure. The existence of Y or propeller shaped triglycerides with disordered chains in the liquid state, which loosely stack on top of each other to form flexible cylindrical rods in a discotic phase, was instead proposed by Corkery *et al.*[20]. In this chapter, this question is addressed by analysing the relative positions of nearest neighbour TG molecules and by employing a Voronoi tessellation method[21] to quantify the number and the shape of TG clusters in the liquid phase.

As it has been discussed in chapter 2, molecular dynamics (MD) simulation is a powerful tool to characterise the structural and mechanical properties of biological systems and to achieve a deeper understanding of the phenomena taking place at the atomistic/molecular level[22, 23]. Simulations can aid the interpretation of experimental results using techniques that require assumptions and approximations, such as X-ray and NMR. Computer simulations can also help to rationalise existing experimental results and visualise processes at the atomistic level, proposing new mechanisms or motivating new questions. MD simulations have been used previously to investigate the properties of pure triglyceride systems in the solid[24, 25] and liquid phases[11, 15, 26]. These studies were of fundamental importance in testing the reliability of forcefield parameters

in reproducing the structural features and dynamics of triglycerides and in helping to comprehend how triglycerides with different chemical structure pack as a function of their intra and inter molecular structures. A deep understanding of the behaviour of triglyceride molecules in contact with air and water is a fundamental preliminary step to address more complex systems. For example, the interactions between triglycerides and phospholipid monolayers have been explored to understand the differences between pure lipid bilayers and more complex systems where lipid molecules with different polarity interact together and with water. These studies are relevant to understand the behaviour of lipid emulsions[27], lipid droplets[28, 29] and tear fluid lipid layers[12].

This chapter presents an MD study of the sebum triglyceride, focusing on the triglyceride bulk and interfacial properties. The chapter is structured as follows. In the Method sections, the different systems investigated in this work (*i.e.* triglycerides in bulk, at the air interface, water interface and a triglyceride monolayer at the air water interface), the simulation parameters and the computations performed in the data analysis are described. The Results section, presents a characterisation of the structure of the bulk TG system, a description of the behaviour of TG at the TG-air or TG-water interface, and a study of the behaviour of TG monolayers. In the section of the TG-water interface the free energy for water penetration within the TG phase is quantified. Finally, the impact of using a different set of charges in the TG glycerol group is examined.

The main results from this chapter have been published in Tascini *et al.*[1]. I wrote the paper draft and realised all the figures. Therefore, most of the contents and figures presented below are also included in ref. 1 and reproduced from ref. 1 with permission from the PCCP Owner Societies.

3.2 Methods

3.2.1 Sebum model: structure and forcefield

The sebum oil is composed of several lipid species, with triglycerides being the most abundant one ($\sim 45\%$, mean weight), followed by wax esters ($\sim 25\%$), squalene ($\sim 12\%$) and fatty acids ($\sim 10\%$)[3]. The fraction of triglycerides and fatty acid concentrations in the sebum reported in literature spans widely, with the former measuring from 20% to 60%, depending on sampling method. The reason behind this, can be found in the ability of some bacteria, such as *P acne*, to hydrolyse triglyceride to free fatty acids and glycerol, which might alter the measured quantity, especially when the analysis is performed on patients affected by *dermatitis*.

Being the first computation model of the sebum oil, the interest was here focused on the sebum most abundant component, the triglycerides. This model was used to study the structural and interfacial properties of the sebum oil. Specifically, Tri-cis-6-hexadecenoin (TG), one of the main components in the sebum formulation[30], was chosen as representative of the human sebum triglycerides. The structure of TG is shown in Figure 3.2. TG is a monounsaturated triglyceride, whose glycerol backbone is connected to three sapienic acid molecules (cis-6-hexadecenoic acid). Sapienic acid is an isomer of palmitoleic acid, with a double bond in position 6. This double bond position is very uncommon in nature[10] and makes this structure unique to humans, hence the name *sapienic* (from *Homo sapiens*).

Bonded and non-bonded interactions were described according to the united atom forcefield by Berger *et al.*[31], previously used in simulations of skin lipids[32, 33, 34]. This parameter set is based on OPLS[35] and GROMOS[36] forcefields and includes the RyckaertBellemans dihedral potential[37] to model the hydrocarbon chains. The unsaturation of the three acyl tails was described using improper dihedrals, according to the description of the double bond in the Berger phosphatidylcholine (POPC) molecule[38].

Following previous simulations of similar triglyceride molecules[11], the glycerol group parameters were fitted to the glycerol group of a generic glycerophospholipid molecule in the Berger database[31, 39]. Atom types and the charge distribution employed in the present paper are

shown in Figure 3.2-bottom.

Water molecules were modelled using the SPC potential for consistency with the Berger lipid parametrisation and the water description employed in other simulations of epidermal lipids[32, 34, 40].

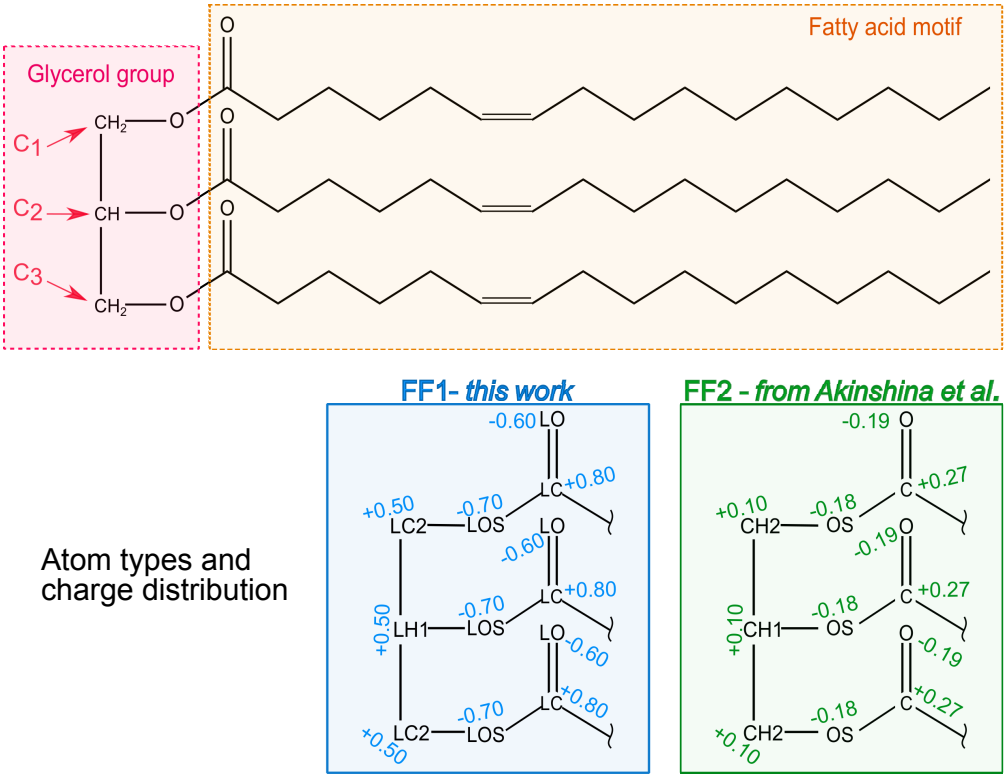


Figure 3.2: Top: Structure of the sebum triglyceride, tri-cis-6-hexadecenoin. The red labels indicate the carbons of the glycerol group: C_1, C_2 and C_3. The glycerol group and fatty acid motifs are grouped into two different rectangles, red and orange respectively, to highlight the partial contributions to the density profiles discussed in the main text. Bottom: Atom types and partial charge distribution of the glycerol group according to the Berger phospholipid forcefield[31, 39] (light blue box) and monoelin parametrisation by Akinshina *et al.*[40] (green box).

3.2.2 Simulation details

To investigate the properties of the TG and its interaction with water, four different input configurations were employed: bulk TG, TG-air and TG-water interfaces and a TG monolayer at the air-water interface. Figure 3.3 shows the four representative snapshots of the system configurations. All simulations in this work were run using the GROMACS/4.5.5 molecular simulation package[41]. The main simulation parameters are summarised in Table 3.1. *NPT* on z denotes the fact that the pressure barostat is applied along the z axis, while the area of the simulation box in the xy plane is kept constant. A description of how the four configurations

Input configuration	N_{TG}	N_{water}	Ensemble	Production run time
Bulk TG	512	-	NPT	300 ns
TG-air	1024	-	NVT	100 ns
TG-water	1024	32 126	NpT on z	200 ns
TG-monolayer	122	9085	NγT	at least 40 ns for each γ

Table 3.1: Main simulation parameters for the different input configurations. N_{TG} and N_{water} are the number of TG and water molecules, respectively. The production run time is the length of the simulation employed to compute the properties.

were constructed and employed to generate the results presented in the Results sections can be found below.

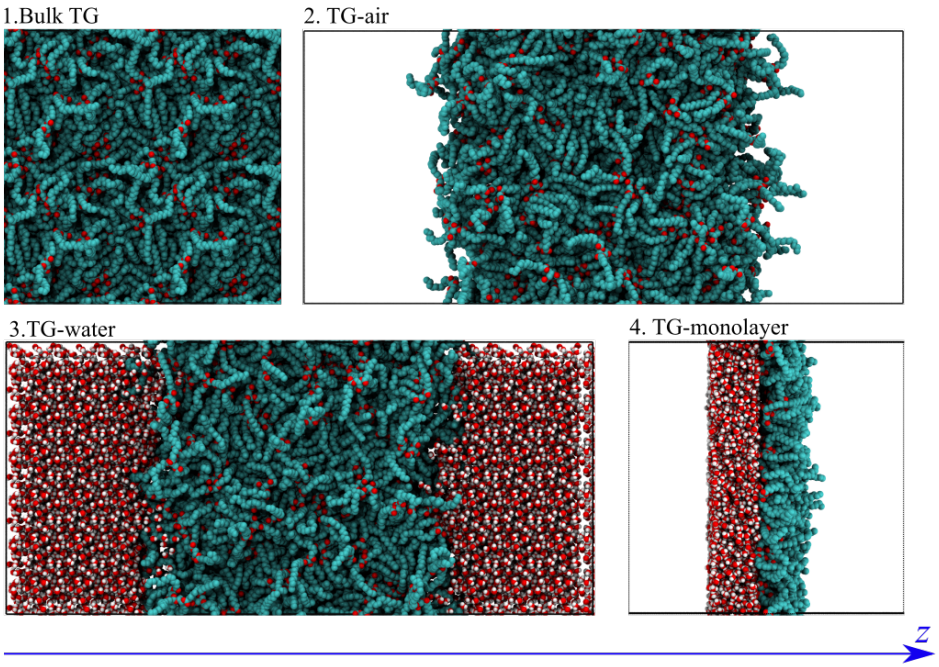


Figure 3.3: 1) Box of 512 triglyceride equilibrated for 80 ns; 2) the box is duplicated along x and y directions and void region is added along the z direction to generate 2 TG-air interfaces; 3) the void space is filled with 32 126 water molecules to generate 2 TG-water interfaces; 4) Monolayer consisted of 122 TG molecules whose glycerol group was partially immersed in a water film with 9085 water molecules. The other surface of the water film was instead left free from TG.

Bulk-TG

Firstly, 128 TG molecules were placed randomly in the simulation box and equilibrated at a temperature, $T = 310$ K, to target the average physiological temperature of our skin. The equilibration was first performed in the NVT ensemble for 200 ps, rescaling the atom velocities with the v-rescale algorithm[42] with a time constant, $\tau_T = 0.1$ ps. At this equilibration stage,

the constant density of the system, based on the box dimensions, was set equal to 910 kg m^{-3}. The system was then equilibrated in the NPT ensemble at $T = 310$ K and pressure, $P = 1$ bar, for 40 ns, using the Nosé-Hoover thermostat and a temperature coupling constant, $\tau_T = 1$ ps, along with the Parrinello-Rahman barostat for the pressure coupling, with a time constant $\tau_P = 2$ ps and compressibility $\kappa = 4.5 \times 10^{-5}$ bar^{-1}. The timestep was set equal to 2 fs. van der Waals interactions were truncated at 1.2 nm and electrostatic interactions were computed in full using the Particle Mesh Ewald algorithm[43]. The TG intra-molecular bonds were constrained using the P-LINCS algorithm and periodic boundary conditions were applied in all directions. After this first equilibration, the system was replicated 4 times in the yz plane resulting in a rectangular box containing a total of 512 lipids. The energy was minimised with the conjugate gradients algorithm[44] and then the system was equilibrated in the NPT ensemble for 80 ns (the resulting equilibrated configuration is shown in Figure 3.3-1). A 120–200 ns production run was then used to compute the bulk properties of the triglyceride.

To confirm the robustness of the results presented below, the bulk-TG simulations were repeated, starting from an independent initial configuration. This configuration was generated by heating up our system to a high temperature ($T_H = 400$ K) followed by a slow cooling of the system down to the chosen working temperature, $T_W = 310$ K, following the simulated annealing technique[45]. The temperature of a bulk-TG configuration (taken at the end of the production-run) was set equal to T_H, leaving the system to equilibrate in the NPT ensemble for 100 ns (with the same simulation settings described above). Then, it was slowly cooled down, following the simulated annealing procedure, by dynamically changing the reference temperature used in the thermostat algorithm, from T_H to 310 K during a 50 ns window. The resulting configuration was equilibrated for an additional 50 ns at $T = T_W$. A production run of 100 ns followed this equilibration procedure.

TG-air and TG-water

The interfacial configurations were generated from the equilibrated TG system, elongating the simulation box in the z direction to create two TG-air interfaces (see Figure 3.3-2). The

equilibrated bulk-TG box of 512 TG molecules was also duplicated along the xy plane to obtain a larger cross sectional interface with a total system size of 1024 TG molecules. In this way, the cross-sectional area of the interface ($A_{xy} \approx 129\,\mathrm{nm}^2$) employed in the simulations was large enough to avoid the periodic error associated with the computation of the surface tension[46]. To compute the surface tension of the TG-air interface, γ_{TG-a}, a 100 ns NVT production run, where the temperature was kept constant using the Nosé Hoover thermostat, was performed. The first 50 ns of simulation were discarded, to allow for the equilibration of the TG-air interface.

For the computation of the TG-water surface tension, γ_{TG-W}, after elongation of the box in the z direction and duplication of the box along xy as above, 32 126 water molecules were added in the void space to create two TG-water interfaces (see Figure 3.3-3). After a short NVT equilibration of 200 ps, the system was simulated in the NPT ensemble, with pressure scaling in the z-direction (*i.e.* perpendicular to the interface) by coupling to a pressure bath at 1 bar allowing the length of the box in this direction to change. According to the procedure used to evaluate the surface tension, the lateral area, A_{xy}, was kept constant at its initial value. The system was equilibrated for 100 ns, followed by a 200 ns production run.

TG monolayer

The triglyceride monolayer was generated using the output configuration of the equilibrated TG-water interface. As discussed below, in the Results section, the interfacial TG molecules tend to re-orient by pointing the glycerol group towards the water interface. Therefore, the monolayer was obtained by selecting all the TG molecules with the central carbon atom of the glycerol group (labelled as C2 in Figure 3.2) close to the first water interface along z (such that their distance from the TG-water interface was ≤ 1 nm - see the density profile in Figure 3.12-right as a reference for the distance). The resulting monolayer consisted of 122 TG molecules whose glycerol group was partially immersed in a water slab with 9085 water molecules, while the other surface of the water slab was instead left free from TG (as shown in see Figure 3.3-4). The correction prescribed for systems with a slab geometry, *i.e.* 2D-Ewald[47], was included in the computation of the electrostatic interactions. The system was equilibrated at constant

temperature for 15 ns. Then, a set of 40 ns production runs was performed with constant surface tensions, γ_t, in the range 40 mN m^{-1}–120 mN m^{-1} (using the $N\gamma T$ ensemble). The system was free to adapt to its equilibrium interfacial area, with the compressibility set to 4.5×10^{-5} bar^{-1}. The Berendsen thermostat and barostat were used in these simulations, following the procedure discussed in ref. 48.

3.2.3 Data analysis

Orientation of the triglyceride chains

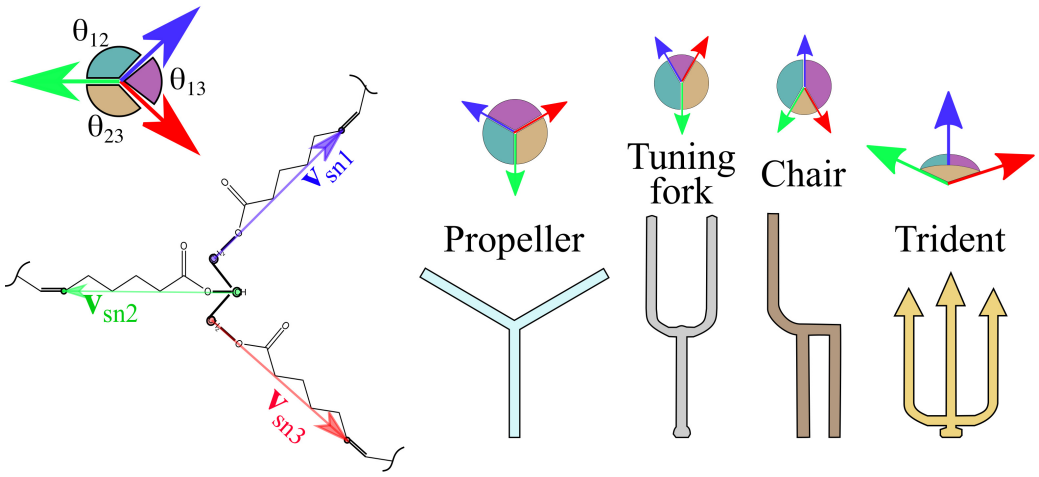


Figure 3.4: (Left) The three vectors describing the three acyl chain orientations, namely $\vec{V}_{sn1}$, $\vec{V}_{sn2}$ and $\vec{V}_{sn3}$, are represented on a sketch of the TG molecule. The inset shows a sketch of the three angles employed for the evaluation of the relative orientations of the chains. (Right) 4 different shapes of TG molecule and representative angle sets. The criteria for classification are described in the text below.

The intra-molecular structure of the sebum triglyceride was investigated by calculating the angles between the triglyceride chains. The kink in the acyl chain, introduced by the presence of the double bond, results in higher degree of disorder after the double bond, towards the end of the chain. Therefore, the direction of the chain was defined as the vector between one of the carbon atoms of the glycerol group (C1, C2 and C3 in Figure 3.2) and the corresponding 6-cis carbon on each chain. The sketch to the left of Figure 3.4 shows the definition of the chain directions given by the vectors $\vec{V}_{sn1}$, $\vec{V}_{sn2}$ and $\vec{V}_{sn3}$ respectively. The angles describing the

orientations were computed from the angles between the corresponding vectors, $\cos\theta_{ij}$, via:

$$\begin{aligned}
\theta_{12} &= \arccos\left(\frac{\vec{V}_{sn1}\cdot\vec{V}_{sn2}}{\|\vec{V}_{sn1}\|\ \|\vec{V}_{sn2}\|}\right),\\
\theta_{23} &= \arccos\left(\frac{\vec{V}_{sn2}\cdot\vec{V}_{sn3}}{\|\vec{V}_{sn2}\|\ \|\vec{V}_{sn3}\|}\right),\\
\theta_{13} &= \arccos\left(\frac{\vec{V}_{sn1}\cdot\vec{V}_{sn3}}{\|\vec{V}_{sn1}\|\ \|\vec{V}_{sn3}\|}\right).
\end{aligned} \tag{3.1}$$

The values of the three angles were used to classify the molecular shape into one of the 4 classes displayed in Figure 3.4-right. From left to right the four classes are:

1. **Propeller (Pr)**: the three angles are uniformly distributed around an average value of 120° (also called Y-shape in previous analysis[20])
2. **Tuning fork (TF)**: two opposite chains are pointing in the same direction and the third one in the opposite direction
3. **Chair (Ch)**: two adjacent chains are pointing in the same direction and the third one in the opposite direction
4. **Trident (Tr)**: the three chains are pointing in the same direction

Four criteria were defined to link the four shapes, described above, to the energetically favourable regions of TG conformations, expressed in terms of the three angle θ_{12}, θ_{23} and θ_{13}. The criteria were defined as follows.

1. **Pr**: $(75 < \theta_{12} < 165)\ \&\ (75 < \theta_{23} < 165)\ \&\ (75 < \theta_{13} < 165)$
2. **TF**: $(95 < \theta_{12} < 180)\ \&\ (95 < \theta_{23} < 180)\,\&\,(0 < \theta_{13} < 75)$
3. **Ch**:
$$\begin{cases} (0 < \theta_{12} < 75)\ \&\ (95 < \theta_{23} < 180)\ \&\,(95 < \theta_{13} < 180) \\ (95 < \theta_{12} < 180)\ \&\ (0 < \theta_{23} < 75)\ \&\,(95 < \theta_{13} < 180) \end{cases}$$
4. **Tr**: otherwise.

The trident region includes all those conformations where at least two of the angle are lower than 95°. It is expected, in fact, that the presence of the double bond in 6^{th} position will result in a wide trident configuration, where one of the three angles could be greater than 95°. In Figure 3.5, the regions assigned to the 4 conformations are shown with different colours: blue for the *propeller*, red for the *tuning fork* and magenta for the *chair*. The region assigned to the trident has been left empty for clarity. Below, in Figure 3.11 and Figure 3.14, the 4 regions are shown overlapped to the 3D-plots of TG conformations for bulk-TG and TG-water interface.

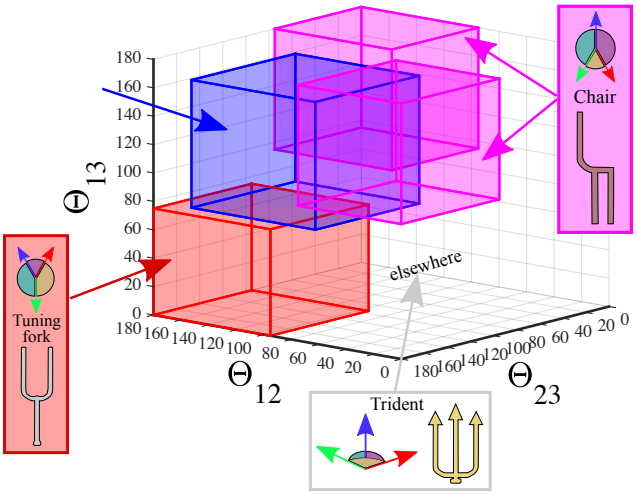


Figure 3.5: Subdivision of the angles phase space into four conformations. The coloured cubes highlight regions assigned to the propeller (blue), the tuning fork (red) and the chair (magenta) conformations. In the remaining region the conformations were assigned to the trident.

Cluster characterization

To quantify the intermolecular organisation of our systems both the radial distribution function[11] and a Voronoi tessellation analysis were used. This Voronoi approach has been previously employed to identify 2D clusters[49] in lipid bilayers; here it is proposed a 3D version of this method.

Given a set of points, s, (often defined in literature as *seeds* or *generators*), a Voroni cell, V_i, can be defined as the convex polyhedron containing the region of space, $\{r \in R^3\}$, closer to the point s_i than to any other point s_j, so that:

$$V_i = \{r \in R^3 \,|\, d(r, s_i) \leq d(r, s_j)\, \forall j \neq i\}. \tag{3.2}$$

Once a system of points is decomposed into Voronoi cells in compliance with the definition expressed in equation 3.2, clusters of cells can be defined according to a suitable criterion. The open source software voro++[21] was used to decompose the TG configuration into 3D Voronoi cells and to list neighbouring particles, defined as particles whose Voronoi cells share a common face. To capture the basic structure of the TG molecule, the head group (*i.e.* the glycerol moiety) was simplified by combining all of the head group atoms together and representing them by a single point located at the position of the central glycerol carbon atom, C2 (see label in Figure 3.2). The tessellation was then performed using this head group point and the carbon atoms of the acyl tails. Clusters were defined by grouping together C2 Voronoi polyhedra which shared a face. The method was used to quantify the size of the TG clusters, the higher the number of adjacent Voronoi polyhedra, the larger the TG cluster.

Interfacial tension and surface pressure

The interfacial tension (γ) of the TG-air and TG-water interfaces was calculated via the microscopic route to the pressure tensor:

$$\gamma = \frac{L_z}{N_{int}} \left[P_{zz} - \frac{1}{2} \left(P_{xx} + P_{yy} \right) \right], \tag{3.3}$$

where N_{int} is the number of interfaces (2 in this case), L_z the box length in the direction normal to the interface, $P_{\alpha\beta}$ the pressure tensor component normal to the direction $\alpha, \beta = x, y$ and z, computed from the virial expression as described in ref. 50.

For the monolayer simulations, the surface pressure of the TG film absorbed at the air-water interface, Π_{TG} was calculated. Π_{TG} is defined by the difference between the surface tension of water at temperature T, $\gamma_W(T)$, and the surface tension of the monolayer at temperature T and area per lipid A, $\gamma_m(A, T)$, as

$$\Pi_{TG}(A, T) = \gamma_W(T) - \gamma_m(A, T). \tag{3.4}$$

At physiological temperature, $T \approx 310\,\mathrm{K}$, the surface tension reported for the SPC water model

is $\gamma_W \approx 52\,\mathrm{mN\,m^{-1}}$ [51]. The monolayer surface tension was obtained from the simulations at constant surface tension γ_t, via $\gamma_m(A,T) = \gamma_t(A,T) - \gamma_W(T)$. The corresponding area per lipid, A_L, is calculated by dividing the cross-sectional area of the simulation box in the xy plane (parallel to the monolayer surface) by the total number of TG molecules, as

$$\langle A_L \rangle = \frac{\langle L_x(t) \times L_y(t) \rangle}{N_{TG}}, \tag{3.5}$$

where $\langle ... \rangle$ represents a time average, $L_x(t) \times L_y(t)$ is the area of the simulation box at time t and N_{TG} the number of TG molecules.

3.3 Results

3.3.1 TG-bulk

Density

Due to the presence of unsaturated chains, the melting temperature of unsaturated triglycerides is much lower ($\approx 0\,^\circ\mathrm{C}$) than the average metabolic temperature of our skin ($\approx 37\,^\circ\mathrm{C}$). The sebum triglyceride, TG, being in contact with the skin, is, therefore, expected to be in the liquid phase. When triglycerides melt from a crystalline polymorphic state to a liquid state, they undergo a small change in volume ($\approx 10\%$), resulting in an highly viscous liquid, with typical viscosities of the order of tens of $\mathrm{mPa\,s}$ [52].

Density values for the liquid phase, found in the literature, range between $875\,\mathrm{kg\,m^{-3}}$ and $930\,\mathrm{kg\,m^{-3}}$ (for Tripalmitolein: $929\,\mathrm{kg\,m^{-3}}$ [53]; Triolein: $907\,\mathrm{kg\,m^{-3}}$ [52] and $915\,\mathrm{kg\,m^{-3}}$ [54]; Tripalmitin $875.2\,\mathrm{kg\,m^{-3}}$ [54]). Figure 3.6 shows the average densities of these triglycerides compared with the simulation results for the sebum TG.

The fluctuations of the density around the well defined average value indicates that the system is well equilibrated and stable in its liquid state. The average density $\approx 930\,\mathrm{kg\,m^{-3}}$ obtained here is very close to the experimental values reported for unsaturated triglycerides with similar

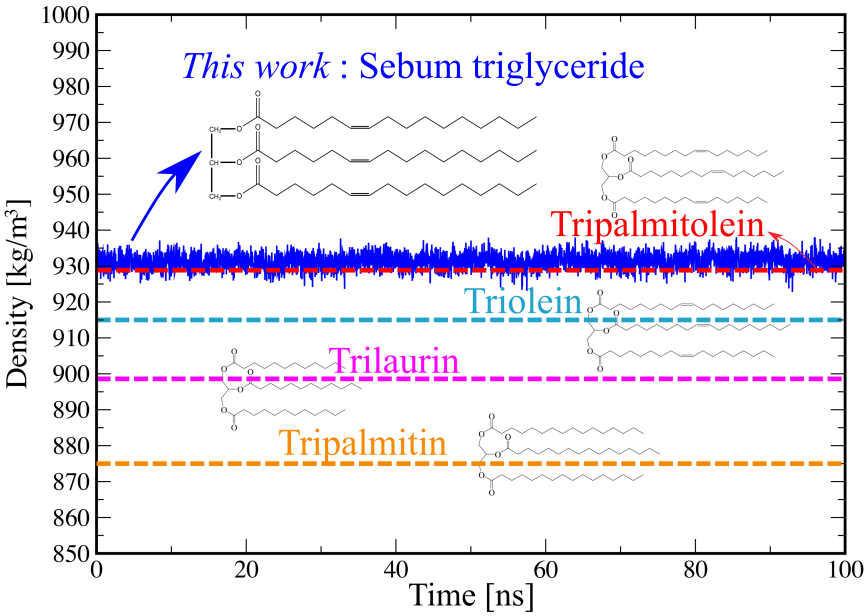


Figure 3.6: Density of TG as a function of time. The dashed straight lines indicate experimental values of other triglyceride molecules. From top to bottom: Tripalmitolein from ref. 53 and Triolein, Trilaurin and Tripalmitin from ref. 54.

structures (*e.g.* Tripalmitolein and Triolein). This result supports the accuracy of the forcefield employed.

Inter molecular arrangement

The radial distribution function (RDF) of the central carbon atom of the glycerol group, C2 (see label in Figure 3.2), was computed to investigate the intermolecular organisation of the triglyceride molecules in the liquid phase. The overall behaviour of C2 presented in Figure 3.7, shows an inhomogeneous structuring of the glycerol groups in the simulation box. The C2-RDF shows two peaks at short distances, a sharp one at $r \approx 0.55\,\text{nm}$ and a second one at $r \approx 0.95\,\text{nm}$, which are followed by a smooth, broad minimum at longer distance, $r \approx 1.6\,\text{nm}$.

The radial distribution function along with the snapshot of the glycerol group distribution, presented in the inset of Figure 3.7, suggest that the TG molecules are not randomly distributed in the simulation box, but instead the TG head-groups are clustered together. An estimate of the number of TG-molecules interacting through the glycerol head group can be obtained from the number of C2 nearest neighbours, by integrating the C2-RDF to its first minimum

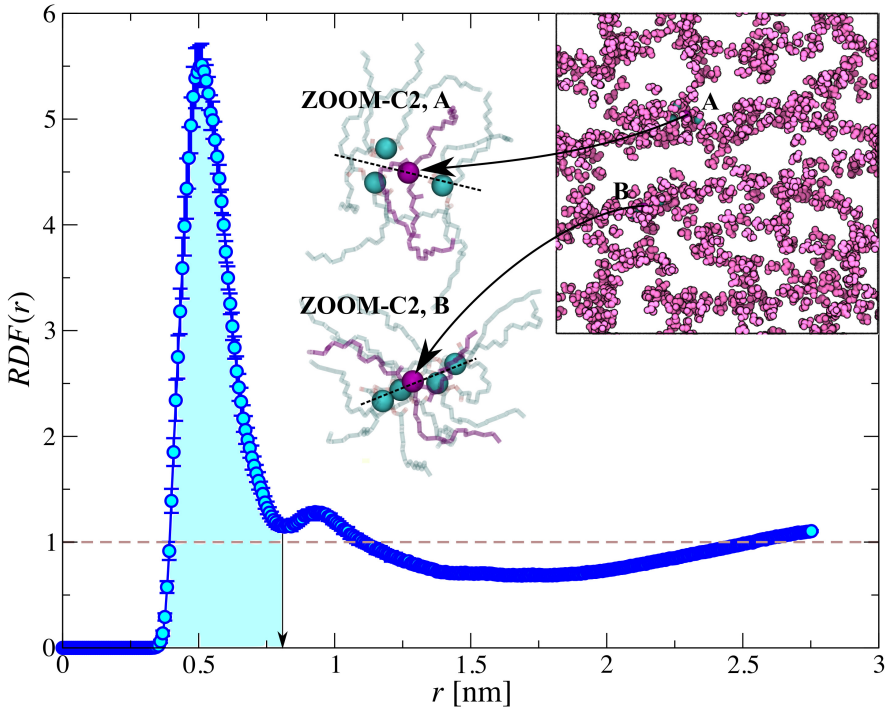
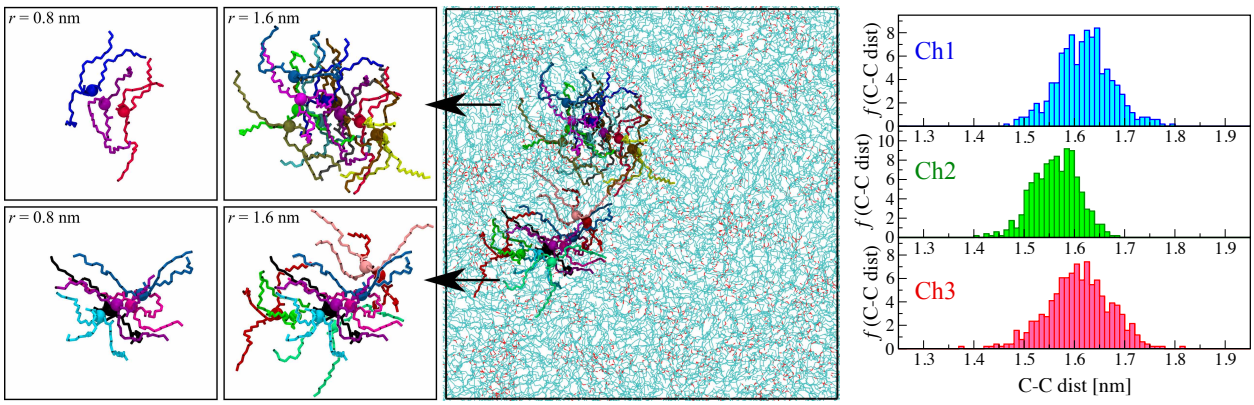


Figure 3.7: Radial distribution function of the central carbon atom of the glycerol group (C2 in Figure 3.2). The coloured area represents the region of integration for the determination of the TG-TG coordination number. The inset at the top-right corner shows the glycerol groups of the TG molecules taken from a representative snapshot of the simulation. The two small insets (ZOOM C2, A and ZOOM C2, B) connected with the snapshot through two black arrows show a two different TG molecules (in purple) with three (ZOOM C2, A) and four (ZOOM C2, B) nearest neighbours, respectively. The nearest neighbours are identified as the TG molecules whose central carbon atom (C2 in Figure 3.2) is placed within a sphere with radius, $r \leq 0.8\,\text{nm}$.

at $r = 0.8\,\text{nm}$ (see light-blue area in Figure 3.7). The coordination number, calculated in this way, is ≈ 3.6. The position of two representative central carbon atoms are highlighted in the inset of Figure 3.7, with three (inset ZOOM C2, A) and four (inset ZOOM C2, B) nearest neighbours respectively. By observing the position of the neighbours, relative to the highlighted C2 (in purple), it is possible to note that the selected C2 and its nearest neighbours are roughly aligned, indicating the existence of elongated clusters.

The second peak of the RDF is related to the confinement of the C2 atoms by steric interaction with other lipid molecules. By integrating up to $r = 1.6\,\text{nm}$ to include this peak, the coordination number is ≈ 11. Inspection of the position of the molecules within a sphere of radius 1.6 nm also indicates that those molecules are preferentially distributed along straight lines (as shown in Figure 3.8a).



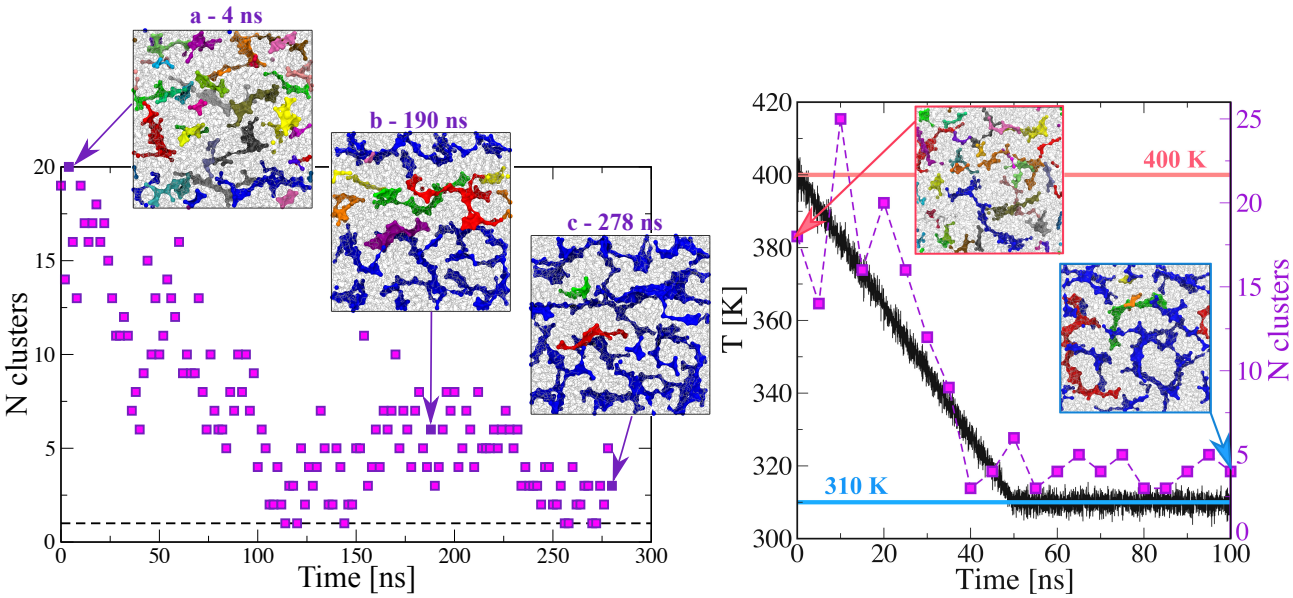
(a) C2 Neighbours **(b)** C-dist histogram

Figure 3.8: (a) Two representative TG molecules (in purple) and their neighbours are highlighted in the TG-bulk phase. The snapshots on the left show the nearest neighbours up to the first ($r \approx 0.8\,\text{nm}$) and second ($r \approx 1.6\,\text{nm}$) minimum of the C2-RDF. (b) Histograms of the distances between C2 (central carbon atom of the glycerol group - labelled in Figure 3.2) and the three terminal carbon atoms of acyl chains Ch1 (top), Ch2 (middle) and Ch3 (bottom).

The position of the broad minimum at $r \approx 1.6\,\text{nm}$ correlates well with the length of the TG acyl chains which can be calculated from the simulations by computing the distances between C2 and the terminal carbon at the end of each chain, as reported in Figure 3.8b where the three histograms show the distribution of these three C-C distances. The histograms are sharp and their maxima correspond to a C-C distance of $\approx 1.6\,\text{nm}$. This suggests that elongated linear-like clusters of glycerol groups are surrounded by the hydrophobic part of the TG molecule.

To further investigate the tendency of the TG molecules to cluster together and to give insight into the number and shape of the clusters an analysis based on the Voronoi polyhedra construction was performed. Figure 3.9a presents the evolution of the number of head group clusters with time, showing a slow decrease in the number of different size clusters. The three simulation snapshots collected at different timesteps (insets *a*, *b* and *c* of Figure 3.9a) illustrate this trend showing the evolution of the head-group organisation from several small elongated clusters that gradually merge over time, resulting in the formation of a single cluster representing an interconnected network involving all the TG head-groups. The speed of the process is linked to the reorganisation of the triglyceride molecules in the bulk.

Figure 3.9b, instead, shows the number of glycerol group clusters (right-hand y axis) and the system temperature (on the left-hand y axis) for the independent initial configuration generated with the simulated annealing protocol. The two insets are two representative simulation



(a) TG Clusters evolution **(b)** SA protocol

Figure 3.9: (a) Evolution of the number of TG head group clusters with time. The three snapshots in the figure represent the simulated system at different times (see darker points in the main plot and labels *a*, *b* and *c*). Atoms belonging to different clusters are highlighted with different colours and joined together with the VMD[55] drawing mode *'Surf'* with probe radius = 2.4 Å[56]. (b) Evolution of the average temperature (left-hand y axis, black curve) and number of clusters (right-hand y axis, shown as magenta squares and dashed line) with time, during the simulated annealing protocol (SA). The two insets illustrate the size and the number of clusters before and after the SA equilibration

snapshots (at the beginning and the end of the simulated annealing equilibration) highlighting the shape of the clusters. This results confirm the evolution of the system from several small clusters to a single percolating network. The number of clusters keeps oscillating between 3 and 1 in the following 100 ns of production run, as previously obtained, with one of the cluster involving almost all the TG head groups (always over 80%).

Intra molecular arrangement

To gain insight into the preferential conformations of the TG molecules the relative orientation of the chains were computed as described in the Method section 3.2.3. The distribution of the angles formed by the relative direction of pairs of chains, θ_{ij}, is represented in the three histograms of Figure 3.10-left. Two of the angles are equivalent by symmetry (θ_{12} and θ_{23}) and display a similar distribution in phase space as expected, with a maximum at about 150°. The θ_{13} distribution, instead, presents a maximum for lower angles ($\approx 60°$). The energetically

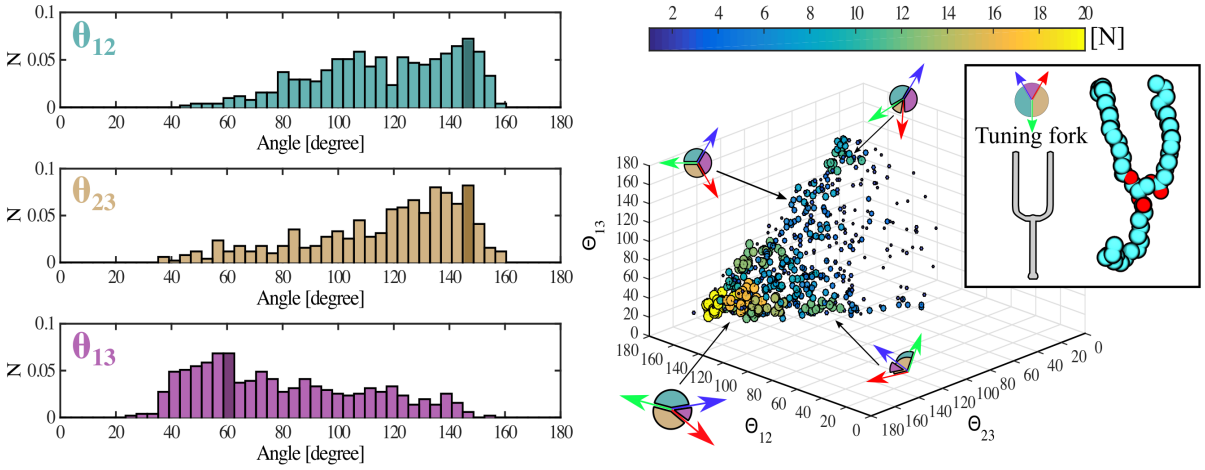


Figure 3.10: Analysis of the intramolecular TG-conformations. Left) The three histograms represent the distributions of the angles, θ_{ij}, formed by pairs of chains. From top to bottom: θ_{12}, θ_{23} and θ_{13}. The darker shaded bars in the histograms highlight the maximum of the three angle distributions. Right) 3D distribution of triglyceride conformation as a function of the three angles described by the mutual orientation of the chains: θ_{12}, θ_{23} and θ_{13}. Each circle corresponds to the average conformation of a TG-molecule in the 3D space. The colour and size of the circles reflects the number density probability of each conformation, using a mesh-grid with $\delta\theta = 20°$. The correspondence between colour and number density is given in the colour bar above the graph. Visual representations of the combination of the angles in the 3D space are shown to highlight the 4 conformational classes (See Figure 3.4). The inset shows the most probable conformation for the TG molecules ($\approx$ 40%). The colouring of the chains here are the same as in Figure 3.4.

favourable regions for TG conformations, expressed in terms of θ_{12}, θ_{23} and θ_{13} are shown in Figure 3.10-right, where each circle represents the average conformation of one TG molecule. The size and the colour of the circles are proportional to the density of their occurrence in 3D-space. The colour bar above the graph gives a measure of the local number density of TG conformations inside a cubic mesh with cubic length, $\delta\theta$, equal to 20° (corresponding to the mesh-grid of Figure 3.10-right). The areas of the diagram corresponding to the four representative TG conformations (described in Figure 3.4) are highlighted with schematic illustrations of the three angles in these conformations. The same plot is shown in Figure 3.11 with the 4 regions linked to the 4 TG shapes superimposed.

The histograms (see Figure 3.10) show a wide distribution of angles in 3D space, corresponding to several different conformations of TG molecules. However, the analysis (summarised in Table 3.2) suggests that the most likely intramolecular arrangement, adopted by $\approx$ 40% of the TG molecules, has the two opposite tails (sn1 and sn3) pointing in one direction and the central one (sn2) pointing in the opposite direction (as shown in the inset of Figure 3.10-right),

resembling the shape of a *tuning fork*.

The group with the smallest percentage of molecules is the open trident conformation (≈ 15%), the only shape of the 4 proposed where the three tails are all roughly pointing at the same direction.

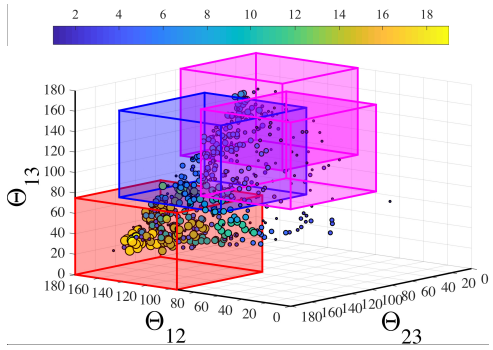


Figure 3.11: 3D distributions of triglyceride conformations for BULK-TG superimposed to the 4 regions describing the 4 TG shapes, as explained in Figure 3.5.

Shape	%
Propeller	27.73
Tuning fork	**39.06**
Chair	17.58
Trident	15.63

Table 3.2: Probability of finding different shapes for molecules in bulk-TG. Results are expressed in % over all the molecules.

The characterisation of the TG ordered structure in the melt phase is far from trivial. The existence of a liquid-state ordering in the triacylglycerols was previously observed in both experimental[17, 19, 20] and computational[11, 15, 26] studies, but the characterisation of the liquid phase is still an open debate, given the existence of different models to describe the structural features of the liquid-ordered structure[16].

From the analysis proposed in this chapter, the TG arrangement in elongated, cylindrical clusters seems to most closely resemble the discotic mesophase structure suggested in ref. 20, in which TG molecules in planar, propeller like conformations stack together. However, our data show a large amount of disorder, with TG conformations which are non-planar and have a wide distribution of angles between chains with a preference for a *tuning fork* conformation. Nonetheless, the head groups do appear to form roughly cylindrical clusters, which quickly merge to form a single network.

3.3.2 TG-air

The TG-air interface computations enable the evaluation of the energy penalty due to the presence of the air interface and to study how the triglyceride molecules preferentially orient

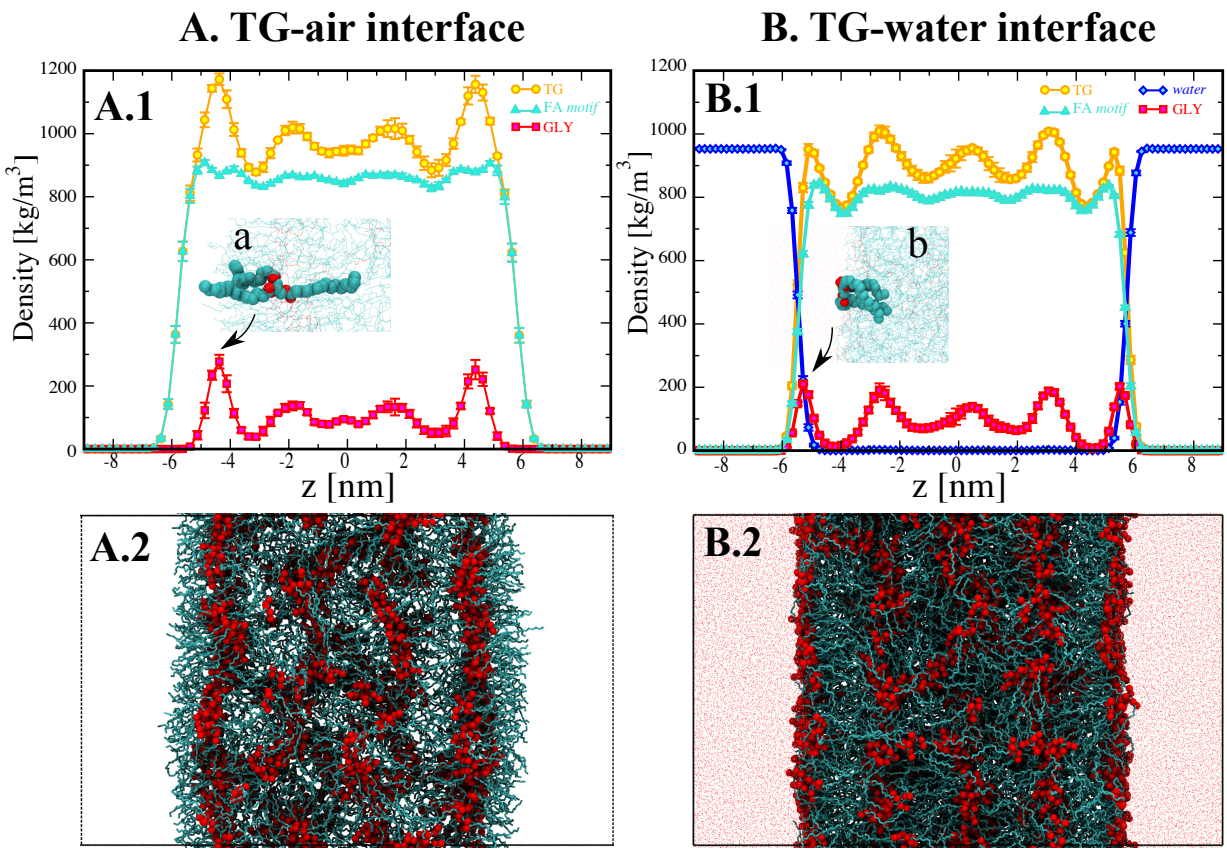


Figure 3.12: . (Left panel) Results for the TG-air interface: the top panel shows the density profiles, highlighting the contribution of different groups of the TG molecule to the total density profile (yellow circles). The glycerol group contribution is displayed with red squares, whilst the free fatty acid motif contribution is represented with light blue triangles. (Right-panel) Results for the TG-water interface: the top figure reports the density profiles, highlighting the contributions of different groups of the TG molecule to the total density profile. The blue diamonds represent the water density profile. The insets show typical conformations of the interfacial TG molecule. The bottom panels represent typical simulation snapshots of the TG-air (left) and TG-water (right) interface, aligned with the corresponding density profiles. Water is represented as red dots.

themselves in order to lower the surface tension of the interface. This analysis can be used as a reference for the TG-water interaction study, and it is of significant interest for the system under investigation, as the sebum oil, placed on the top of our skin, is naturally located at the interface with the external environment.

Specific studies of the surface tension of pure triglycerides are scarce, with a few exceptions in literature (see *e.g.* ref. 57, 58, 59). The surface tension of a single component provides information on the cohesive energy of the fluid.

Table 3.3 lists different experimental values for fatty acid and triglyceride surface tensions reported in the literature. The analysis of previous studies helps to understand the impact of

Surface tension [mN/m], $T = 323\,\mathrm{K}$			
Component	**ref.** 57	**ref.** 58	**ref.** 59
Oleic acid	-	31.9	32.5
Triolein	27.34	-	32
Trilinolein	20.28	-	-
Tricaprylin	24.44	28.81	26.5

Table 3.3: Different surface tension values in $\mathrm{mN\,m^{-1}}$ for a fatty acid and three triglyceride structures at $T = 313\,\mathrm{K}$. Data taken from references 57, 58, 59.

the triglyceride structure on the surface tension, in terms of chain length and unsaturation. Specifically, in the analysis proposed by Benerito *et al.*[57], the length of the chains was found to play a key role in determining the changes in entropy and enthalpy of the surface, showing a minimum as the chain length is increased. Understanding the role of the degree of unsaturation and the position of the double bond is more complex. Even though, the presence of unsaturation was in fact observed to have a small effect on the surface tension, the reported impact of adding a second double bond or the existence of a *cis* or *trans* isomer appears to be relevant.

The surface tension of the sebum TG computed from the TG-air system configuration measures $25 \pm 2\ \mathrm{mN\,m^{-1}}$. This value is slightly lower than the experimental values for triolein (an unsaturated triglyceride with similar chain lengths), but as discussed above, the peculiar position of the double bond characteristic of sapienic acid might play a role here, decreasing the value of the surface tension in a similar fashion to the addition of a second double bond (as for triolein and trilinolein in Table 3.3). Otherwise, the computed surface tension is in good agreement with the reported value for sebum oil surface tension. Butcher and Coonin[60] measured the surface tension of sebum oil with capillary tubes, reporting an average value of $24.9\,\mathrm{mN\,m^{-1}}$ at $T \approx 30\,^{\circ}\mathrm{C}$ and $25.6\,\mathrm{mN\,m^{-1}}$ for higher temperature ($T \approx 40\,^{\circ}\mathrm{C}$). The comparison with experimental data supports the accuracy of the forcefield employed in this analysis.

The TG molecules in contact with air orient their tails toward the air interface, bringing the polar group into the bulk phase. The typical conformation of a TG molecule at the interface is displayed in the inset *a* of Figure 3.12A.1 and in the simulation snapshot below (Figure 3.12A.2). This tendency is reflected in the density profile reported in Figure 3.12A.1: the contribution of the glycerol group (red curve and squares) presents a sharp maximum at $\approx 1.5\,\mathrm{nm}$ from the

interface and an enrichment of the acyl tails contribution (light blue curve and triangles) can be observed immediately adjacent to the two interfaces. The oscillatory behaviour of the density profiles is due to the clustering of the TG molecules in the bulk phase. This lateral organisation of the liquid phase was already noted in section 3.3.1 and it is here further enhanced by the presence of the interface.

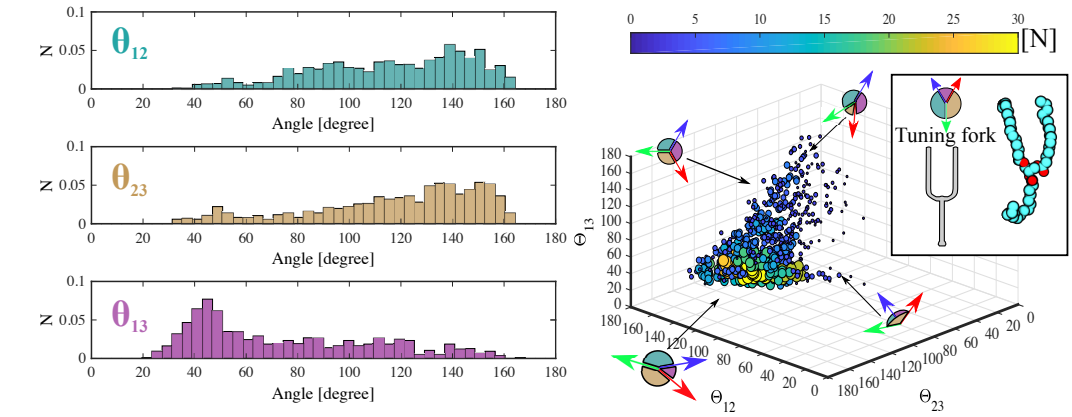
3.3.3 TG-water

Triglycerides are water insoluble molecules. When in contact with water the interfacial triglyceride molecules rearrange themselves to lower the energy penalty due to the interface, with the glycerol groups facing the water phase and the three tails directed towards the bulk TG phase (see inset *b* of Figure 3.12B.1 and simulation snapshot below in Figure 3.12B.2).

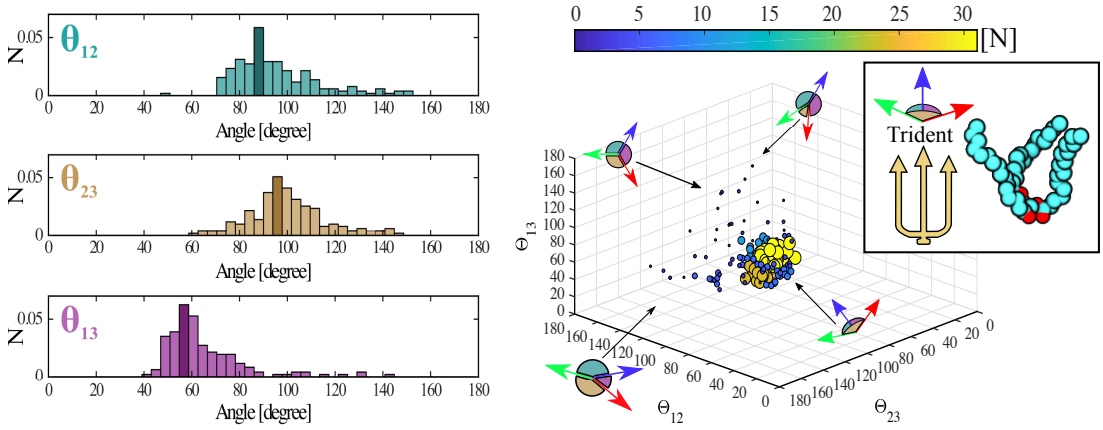
The distribution of the three angles describing the mutual orientations of the acyl chains are reported in the histograms of Figure 3.13-left. The three histograms on the top panel (a) show the distribution of θ_{12}, θ_{23} and θ_{13} for the bulk TG molecules, far from the interface, while the three histograms on the bottom panel (b) are obtained by including in the distribution only the contribution of TG molecules located at the interface. With the same procedure for conformation analysis, described in section 3.3.1, the energetically allowed regions for interfacial TG conformations, expressed in terms of θ_{12}, θ_{23} and θ_{13} are represented in Figure 3.13-right for bulk (top) and interfacial (bottom) TG.

Bulk TG molecules far from the interface show the same behaviours of bulk-TG (see Figure 3.10), presenting wide angle distributions and the occurrence of several conformations within the system. Also in this case, as shown from the inset of Figure 3.13a-right and in Table 3.5, the analysis suggests that the most likely intra-molecular arrangement presents the two opposite tails (sn1 and sn3) pointing towards one direction and the central one (sn2) pointing in the opposite direction, resembling the shape of a *tuning fork*.

For interfacial TG, instead, each angle shows a maximum at $\theta_{ij} \leq 90°$, indicating that most of the molecules at the interface adopt an open trident shape. Unlike in the bulk-TG case, the



(a) Bulk TG



(b) Interfacial TG

Figure 3.13: Analysis of the intra-molecular TG-conformations for bulk TG (a - top) and interfacial TG (b - bottom) at the TG-water interface. (Left) The three histograms represent the distributions of the angles, θ_{ij}. From top to bottom: θ_{12}, θ_{23} and θ_{13}. (Right) 3D distribution of triglyceride conformation as a function of the three angles described by the mutual orientation of the chains: θ_{12}, θ_{23} and θ_{13}. Each circle corresponds to the average conformation of a TG-molecule in the 3D space. The colour and size of the circles reflects the number density probability of each conformation, using a mesh-grid with $d\theta = 20°$. The correspondence between colour and number density is given in the colour bar above the graph. The insets show the most probable conformations for bulk TG molecules, the *tuning fork*, and interfacial TG molecules, the *trident*.

histograms and the 3D-plot on the bottom panel show a sharp distribution, and it is clear that the most probable conformation is the *trident* one ($\approx 70\%$), whose sketch and representative conformation are shown in the inset of Figure. 3.13b-right. In this case all the conformations where one of the chains is pointing in opposite direction to the other two are unfavorable, featuring a very low probability ($\leq 30\%$), as reported in Table 3.4.

The 4 regions of the angles 3D space assigned to the 4 TG shapes are reported below the tables in Figure 3.14 with the conformational distribution superimposed.

Shape	%
Propeller	11.24
Tuning fork	15.98
Chair	3.55
Trident	69.23

Table 3.4: Probability of finding different shapes for interfacial TG in TG-water configurations. Results are expressed in % over all the molecules placed at the interface.

Shape	%
Propeller	23.00
Tuning fork	43.34
Chair	16.23
Trident	17.53

Table 3.5: Probability of finding different shapes for bulk TG in TG-water configurations. Results are expressed in % over all the molecules placed at the interface.

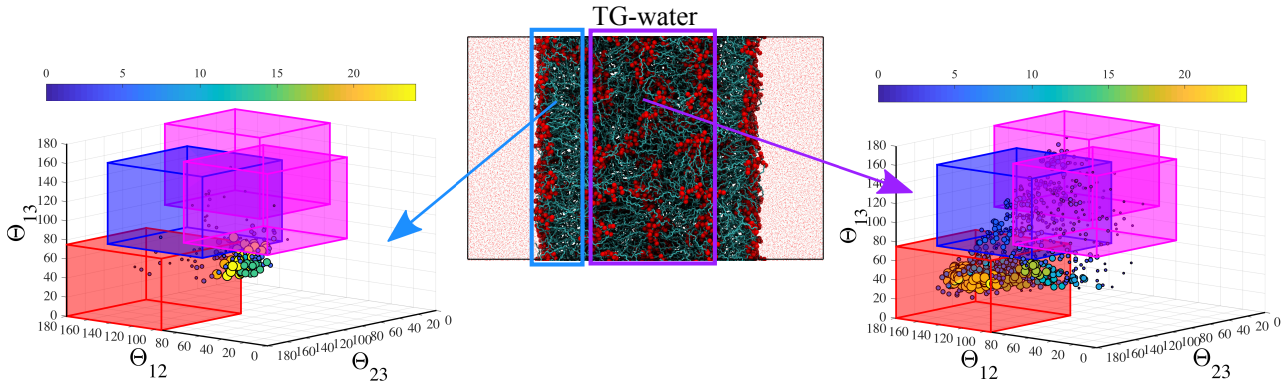


Figure 3.14: 3D distributions of triglyceride conformations for TG-water interface (interfacial on the left and bulk on the right) superimposed to the 4 regions describing the 4 TG shapes, as explained in Figure 3.5.

In the density profiles represented in Figure 3.12B.1, the first peak in the glycerol group distribution (red squares) shows that the glycerol group is adsorbed on the water surface. Further peaks in the distribution are due to the layering of the TG with respect to the interface plane. In section 3.3.1 the existence of glycerol group clustering in bulk TG in the liquid phase was reported. In this case, the ordering induced by the presence of the interface is higher and the TG linear clusters closer to the interface tend to align parallel to the interface (as indicated by the representative simulation snapshot shown in Figure 3.12B.2). The characteristic timescale to form these domains is hundreds of nanoseconds.

By magnifying the contribution of the water oxygen (O-W) to the water density profile 100 times (see black diamonds in Figure 3.15-top), it appears that some water molecules ($\approx$ 50 molecules after 400 ns as represented in the simulation snapshot in Figure 3.15-bottom) are able to permeate the TG phase. This fact can be used to construct the interfacial free energy of water, and to estimate the energy cost of dissolving water in the TG phase. The energy

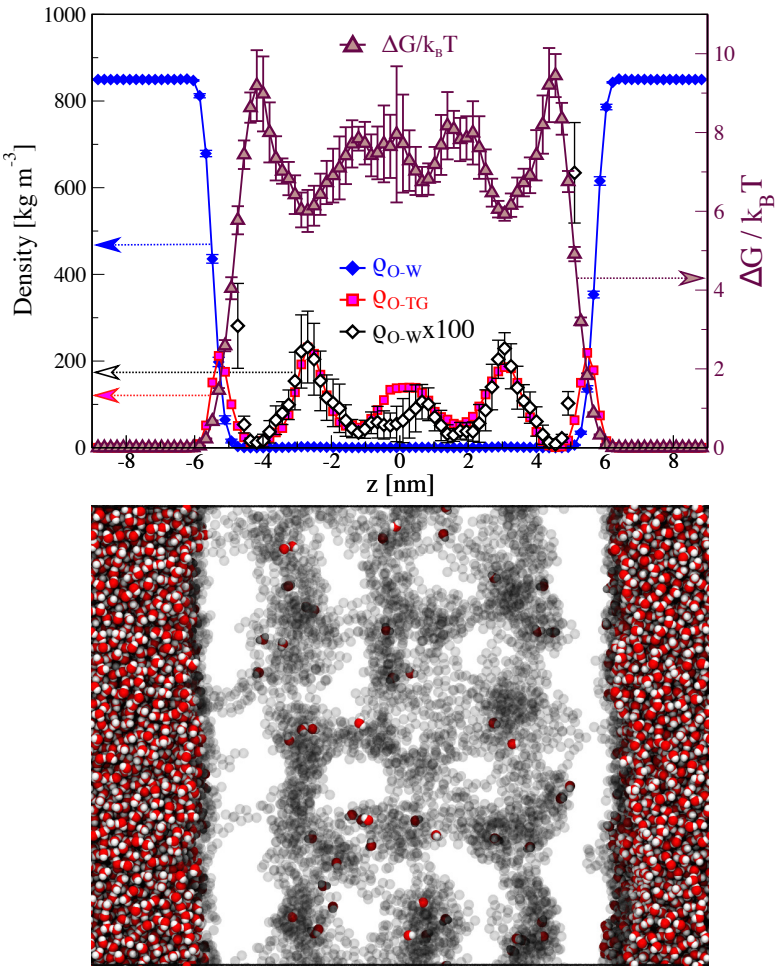


Figure 3.15: Top panel: Density profiles showing the contribution of oxygen atoms in water and TG molecules. Water-oxygens, ρ_{O-W}, are represented with blue diamonds and black diamonds, when magnified by a factor of 100 (ρ_{O-W}x100). TG-oxygens density is represented with red squares. The free energy profile of water insertion in the TG phase is also shown on the right-hand y axis and is represented with brown triangles. Bottom panel: A representative simulation snapshot showing the water molecule distribution in the bulk and inside the TG phase. The carbon atoms in the TG phase have not been rendered for clarity and the TG-oxygens are displayed on the background with grey semi-transparent circles, with darker zones corresponding to a more highly dense oxygen distribution.

barrier to water penetration can be estimated from the water density profile as:

$$\Delta G\left(z\right) = -k_B T \ln\left(\frac{\rho_{O-W}(z)}{\rho_{O-W,BULK}}\right), \tag{3.6}$$

where $\rho_{O-W}\left(z\right)$ is the density of oxygens in water along the direction normal to the interface, z, $\rho_{O-W,BULK}$ is the water density in the bulk phase, k_B the Boltzmann constant and $T = 310\,\mathrm{K}$. ΔG vs z is displayed in Figure 3.15 (right-hand y axis). The maximum energy barrier for permeation of water molecules in the TG phase amounts to $\approx 9.5 k_B T$, which corresponds to an activation energy of $23 \pm 2\,\mathrm{kJ\,mol^{-1}}$ at $T = 310\,\mathrm{K}$. The position of this maximum appears

to coincide with the local maximum in the density of the TG acyl chains. After crossing this activation barrier the water molecules can penetrate the TG phase, preferentially interacting with the oxygens of the TG molecules, O_{TG}. The strong interactions between water molecules and O_{TG} provides a stabilising force which decreases the free energy of water penetration and results in a difference in free energy of $\approx 7k_BT$ between bulk water and water within the TG phase. The hydration of triglyceride head groups was previously observed in triolein simulations and is consistent with the reported water content in olive oil[29]. Here, it is proposed as a possible mechanism of water penetration across the sebum oil, suggesting that triglycerides actively contribute to the water transport across the sebum layer via the oxygen percolating network. This energy cost is comparable with the excess free energy of water in hexadecane ($25\,\mathrm{kJ\,mol^{-1}}$ [61]) and with the peak of the excess free energy to water penetration through a DPPC bilayer ($26 \pm 2\,\mathrm{kJ\,mol^{-1}}$ at $350\,\mathrm{K}$[62] which corresponds to $\approx 9k_BT$). This energy barrier is low as compared with the one of the stratum corneum lipid bilayer ($\approx 40\,\mathrm{kJ\,mol^{-1}}$ at $350\,\mathrm{K}$ or $\approx 15k_BT$[63]), suggesting that the stratum corneum represents the main barrier to water permeation into our body.

The computed interfacial tension for the TG-water interface is $37\,\mathrm{mN\,m^{-1}}$, lower than the reported value for water-alkane interfaces ($\approx 50\,\mathrm{mN\,m^{-1}}$) likely due to the ability of interfacial TG molecules to rearrange themselves with the glycerol group pointing towards the water surface. Simulations of a similar triglyceride at the water interface reported a value of $31 \pm 2\mathrm{mN\,m^{-1}}$ [28]. The differences between these two values might be understood by considering the position of the double bond in the acyl chain, as previously proposed in the TG-air case. The shifting of the double bond closer to the glycerol group might result, in fact, in a more open trident conformation with a higher exposure of the acyl tails to the water surface and consequently a higher energy penalty. This hypothesis remains to be tested.

3.3.4 TG monolayer

To further investigate the interaction of TG and water, a TG monolayer at the air water interface was simulated to compute the surface pressure area isotherm at T=310 K. This

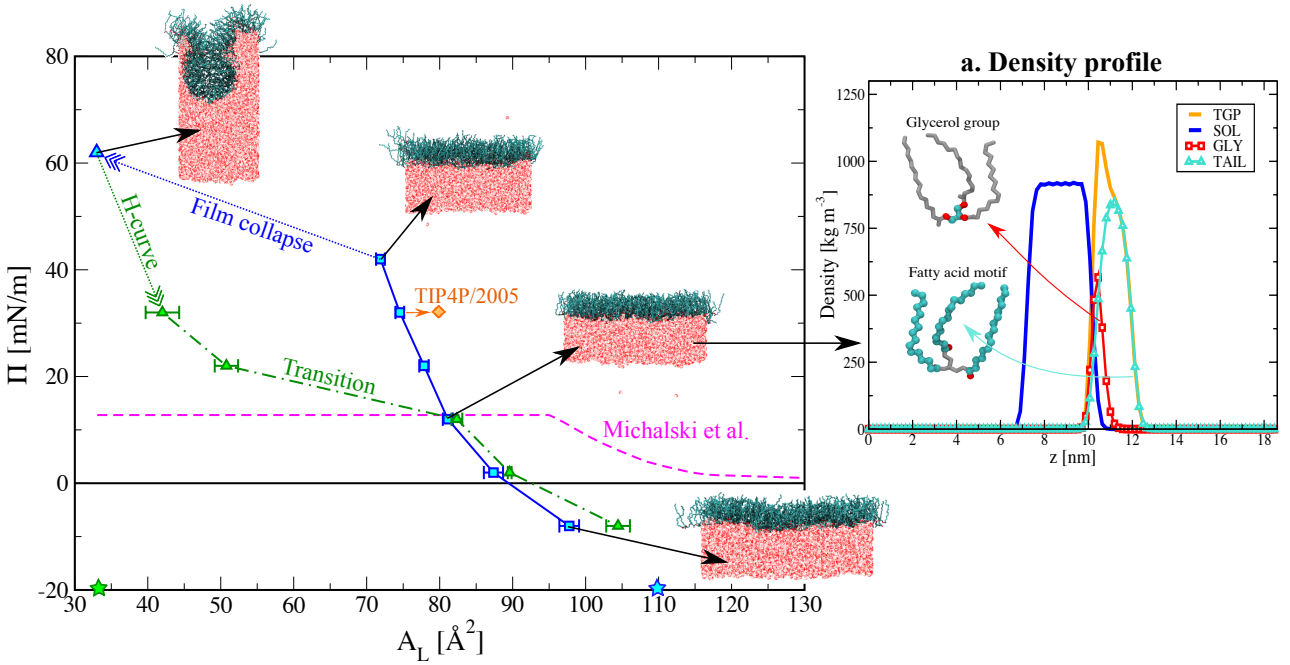


Figure 3.16: The main plot displays the computed surface pressure-area isotherm at $T = 310\,\mathrm{K}$. Different curves and symbols correspond to initial configurations with different areas per lipid, $A_{L,0}$, $110\,\mathrm{Å}^2$ for the blue squares joined by the blue curve and $32\,\mathrm{Å}^2$ for the green triangles joined by a dashed line (see blue and green stars on x-axis, respectively). The experimental isotherm proposed in ref. 64 is shown in the plot with a magenta dashed line. Representative snapshots of the TG monolayer for different surface pressures are shown in the figure to visualise the appearance of the liquid expanded phase until the film collapse at $\Pi \approx 42\,\mathrm{mN\,m^{-1}}$. The inset **a** on the top-right shows a representative density profile for $\Pi = 12\,\mathrm{mN\,m^{-1}}$. The contributions of the glycerol group and fatty acid motif are plotted separately.

analysis enables a direct comparison with experimental results from Langmuir trough studies and is very useful to characterise the fluid states of lipid monolayers under different conditions, including high-pressure/low surface tension regions that are difficult to address experimentally.

The main plot of Figure 3.16 shows the computed surface pressure-area isotherm. The surface pressure, Π, was evaluated from γ_W and γ_m as described in the Methods section (see equation 3.4). The initial monolayer was obtained from the interfacial TG molecules in the equilibrated TG-water configuration (as described in the Methods section) with corresponding area per lipid, A_{L0}, indicated by the blue star on the x-axis of the figure. The corresponding isotherm is the curve represented in blue. As the surface pressure, Π, increases, the TG packing gets tighter, but there is no evidence of a liquid-expanded to liquid-condensed phase transition. The insets in Figure 3.16 display representative snapshots of the equilibrated monolayers at different Π values. From bottom to the top, the first three configurations present an increasing

packing of the acyl chains, in a similar arrangement with all the head-group of the TG molecules facing the water interface in a well defined trident conformation. This is reflected in the density profiles (see inset **a** of Figure 3.16), where the contribution of the glycerol group and the fatty acid motif have been computed separately.

After a critical surface pressure value, $\Pi > 42\,\mathrm{mN\,m^{-1}}$, the TG monolayer collapses, folding into a curved shape (as shown in the simulation snapshot at the top-left corner of Figure 3.16), with a dramatic decrease in the computed area per lipid. Due to the presence of unsaturation in the three fatty acid acyl chains of the TG molecules the formation of the liquid-condensed phase is very unlikely. The presence of the double bonds results in a wider TG conformation, for which reaching the closer molecular packing, typical of the condensed phase, is challenging. This result is confirmed by experimental studies of triolein monolayers at the water interface[64].

To support this hypothesis, the lateral radial distribution functions of the central carbon of the glycerol group (see label C2 in Figure 3.2), RDF_{C2}, was computed for different values of the surface pressure, Π and shown in Figure 3.17. The different lines in figure show RDF_{C2} for different values of Π, in $\mathrm{mN\,m^{-1}}$. Successive RDF functions have been shifted vertically by 0.05 units for greater clarity. The long-range order in the lipid-lipid in-plane radial distribution function within the monolayer has been previously used to detect the transition from a liquid expanded phase to liquid condensed phase, as described in ref. 48 and 65. Even at high values of the surface pressure, in this case, the RDF does not show long range oscillations, which would be indicative of the formation of a liquid condensed phase. This behaviour is reproduced across the different surface pressures investigated here (see Figure 3.17).

To test the reversibility of the pressure-area isotherm curve and the presence of possible hysteresis and metastable states, the computations were repeated starting from the collapsed state, with the area per lipid, A_{L0}, equal to $34\,\mathrm{\AA^2}$ as indicated on the x-axis of the main plot in Figure 3.16 with a green star. The presence of hysteresis is revealed by this new set of simulations with the monolayer reaching a stable equilibrated state only when Π is $\leq 12\,\mathrm{mN\,m^{-1}}$. Above this threshold the collapsed monolayer is not able to rearrange itself, with some of the triglyceride molecules remaining far from the main water interface, forming inverse micelle structures.

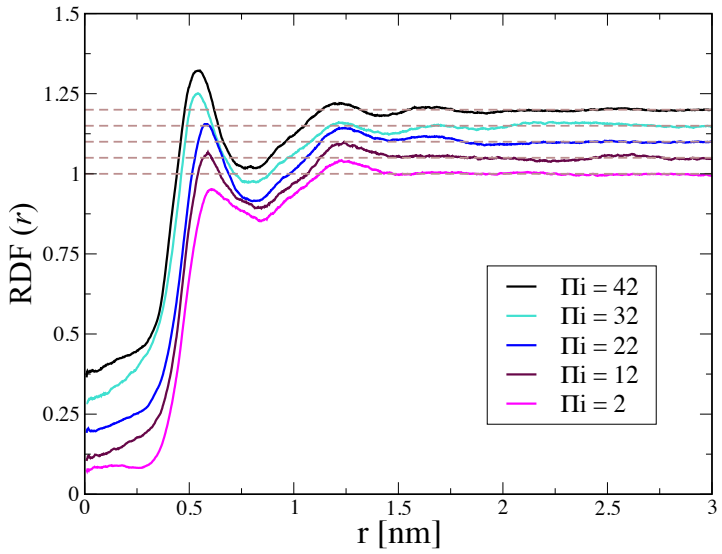


Figure 3.17: Lateral radial distribution functions, *rdf*(r), of the central carbon atom of the glycerol group - C2 in the main paper, for different values of surface pressure expressed in $\mathrm{mN\,m^{-1}}$. The RDF have been shifted upwards by 0.05 units along the y-axis to enable a better comparison of the different curves.

Representative snapshots of the resulting configurations along the hysteresis curve are reported in Figure 3.18, highlighting the threshold of stability for the monolayer. The simulations predict the collapsing surface pressure to be in the range 12–42 $\mathrm{mN\,m^{-1}}$.

The experimental results from Langmuir trough experiments presented in ref. 64 for triolein, a triglyceride with a similar structure to TG were added in Figure 3.16 for a close comparison with experimental results. The reported collapsing pressure for triolein is not far from the value predicted by our computation ($\Pi \approx 13\,\mathrm{mN\,m^{-1}}$) and the unsaturated triglyceride does not experience any liquid expanded/liquid condensed phase transition, in agreement with the simulations. Nonetheless, the comparison between simulations and experiments has to be interpreted carefully, as several sources of discrepancy might exist between experimental and simulated pressure-area isotherms. First of all, not many experimental studies of unsaturated triglycerides are available in the literature. Even where experimental results are available, the sensitivity of the experimental procedure to the set up parameters needs to be taken into account, as previously discussed in ref. 65 in the case of DPPC monolayers where a large range of experimental values can be found in the literature. In addition, from the computational point of view we have to bear in mind the limitations of the SPC water model in reproducing the right value of the water surface tension[51], which is $\approx 20\,\mathrm{mN\,m^{-1}}$ lower than the experimental one

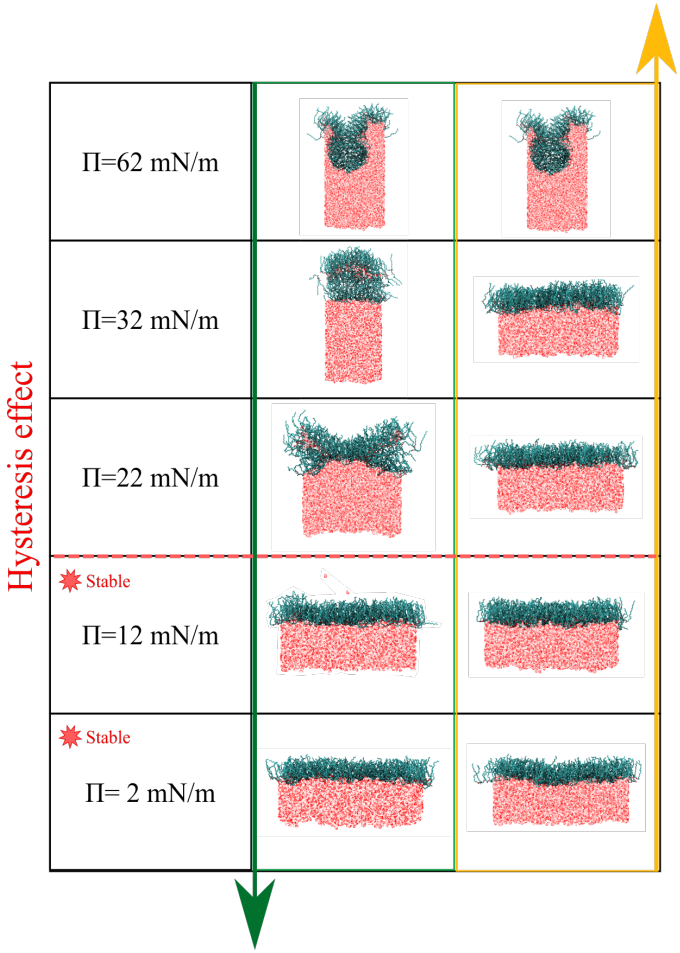


Figure 3.18: Representative monolayer configurations at different surface pressures, Π. The second column (in green) corresponds to the configurations relative to the hysteresis curve, with initial area,$A_L0 = 34\,\text{Å}$. The third column corresponds to the configurations derived by an initial configuration with initial monolayer area $A_L0 = 110\,\text{Å}$. The threshold of stability ($\Pi > 12\,\text{mN}\,\text{m}^{-1}$) is highlighted with a red line.

at 310K. Other water models, such as TIP4P-2005[66] and OPC[67] represent a step forward in attempts to quantitatively reproduce experimental results, with a more accurate representation of the water surface tension. These models have been previously employed in computational studies of lipid monolayers at the air-water interface, see ref 65 and 68 respectively, with encouraging results. To investigate the impact of the incorrect surface tension of SPC water, the analysis was repeated by replacing the SPC water with the TIP4P-2005 model. At equivalent surface pressure, the higher water surface tension of the TIP4P-2005 model shifts the area per lipid of the monolayer towards higher values. This tendency is highlighted in the graph for $\Pi = 32\,\text{mN}\,\text{m}^{-1}$. Despite the limitation of the SPC water model, it remains the best choice in this analysis to guarantee consistency with the water model used for the parametrisation of the

Berger forcefield and as a way to connect to previous simulations of skin lipids performed with the same water model.

Finally, when simulating biological systems the existence of metastable states has to be taken into account. Metastable states have been previously reported[48] for lipid monolayers at large and small areas, where the surface pressure is $\Pi < 0$ and $\Pi > \gamma_w$ respectively. When $\Pi \leq 0\,\mathrm{mN\,m^{-1}}$ the system undergoes a transition to the vapour phase whose activation time might be longer then the simulated 40 ns. The characteristic time of the transition to the vapour phase can be assessed by computing the time required for the area per lipid to diverge. The area per lipid, $A_L(t)$, is defined by equation 3.5 (in Methods section). The curves in Figure 3.19 show

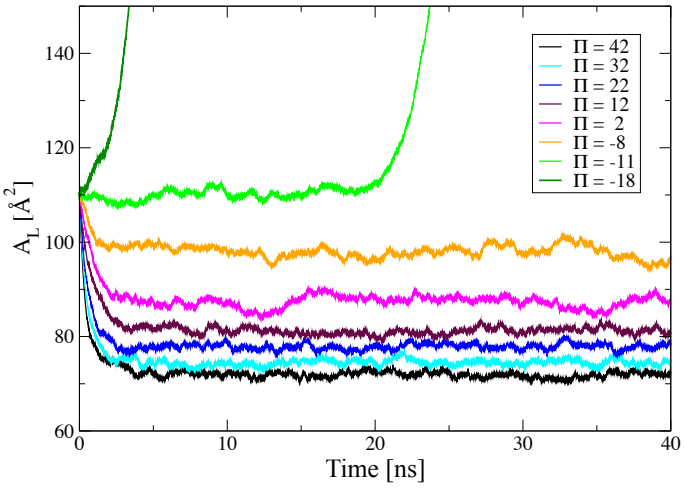


Figure 3.19: Evolution of the area per lipid, A_L, with time. Different lines correspond to different surface pressures in $\mathrm{mN\,m^{-1}}$.

the evolution of A_L for several values of Π. For very negative values of Π, such as $-18\,\mathrm{mN\,m^{-1}}$ - dark green curve in Figure 3.19, the transition to the vapour phase appears after a few nanoseconds. The closer the surface pressure is to $0\,\mathrm{mN\,m^{-1}}$, the longer the time to observe the transition. At $\Pi = -11\,\mathrm{mN\,m^{-1}}$ the transition happens after ≈ 20 ns. At $\Pi = -2\,\mathrm{mN\,m^{-1}}$ a time longer than 40 ns would be required for the transition to take place.

3.4 Impact of the charge distribution of the GLY group

One of the main finding of this chapter is the ability of the glycerol groups to cluster together, enabling water diffusion through the sebum layer. To assess the impact of the triglyceride charge distribution on the TG properties, the set of simulations described above was repeated with a different set of charge for the glycerol group, as shown in Figure 3.2. The choice of using this specific set of charges was motivated by the description of the monoglyceride molecules presented in ref. 40. In this reference, Akinshina *et al.*[40] study how monoglycerides (specifically monoelin) and other free fatty acids, present in cream formulations and in the sebum layer, interact with model SC lipid membranes. It is, therefore, a relevant study for the analysis proposed in this and the following chapters.

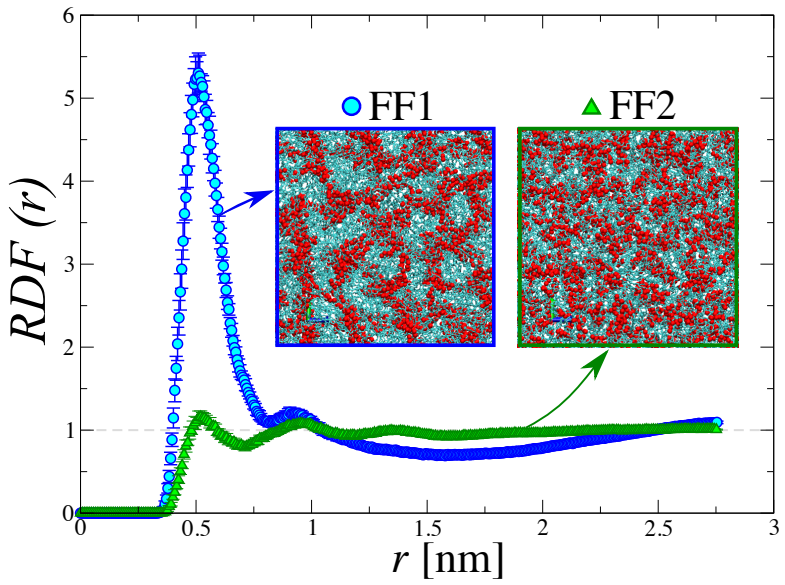


Figure 3.20: Comparison between the C2 RDFs computed using the two different sets of charges, FF1 and FF2 (see Figure3.3 for the charge values). The two insets are snapshots of the system in the two different cases.

Previous works[11] suggested that the impact of the TG glycerol group set of charges is of minor importance in describing the properties of TG molecules. However, the set of charge used in ref. 40 to described the moiety of monoelin is significantly lower than the once reported in the Berger forcefield, which were used in this work and in ref. 11 (see comparison in Figure 3.3-bottom). Therefore, the assessment of their reliability in reproducing the properties of sebum triglycerides appeared to be an important question to address in this chapter. In the following

results relative to the forcefield used in this work are labelled with FF1, and the results relative to the new set of charges introduced in ref. 40 with FF2.

Simulation of TG molecules in bulk show different results in terms of TG density and TG aggregation. The computed density with the FF2 set of charges is higher than the one estimated with forcefield FF1 (950 kg m^{-3} *vs.* 930 kg m^{-3} for FF2 and FF1, respectively) and higher than any experimental values proposed for TG molecules in the liquid phase (see Figure 3.6 for references of experimental values). In addition, in the new set of simulations the TG molecules in bulk lose their ordering and distribute more homogeneously in the simulation box, namely FF2 set of charges is not strong enough to induce clustering of the TG head groups. This result is reflected in the C2-RDF, presented in Figure 3.20. The FF2-RDF (green circles) does not feature a strong peak for short distances and long-range oscillations, but smoothly oscillates toward 1, following a similar trend to pure liquids. The structural differences of the TG in the simulation box can be also observed in the two insets of Figure 3.20.

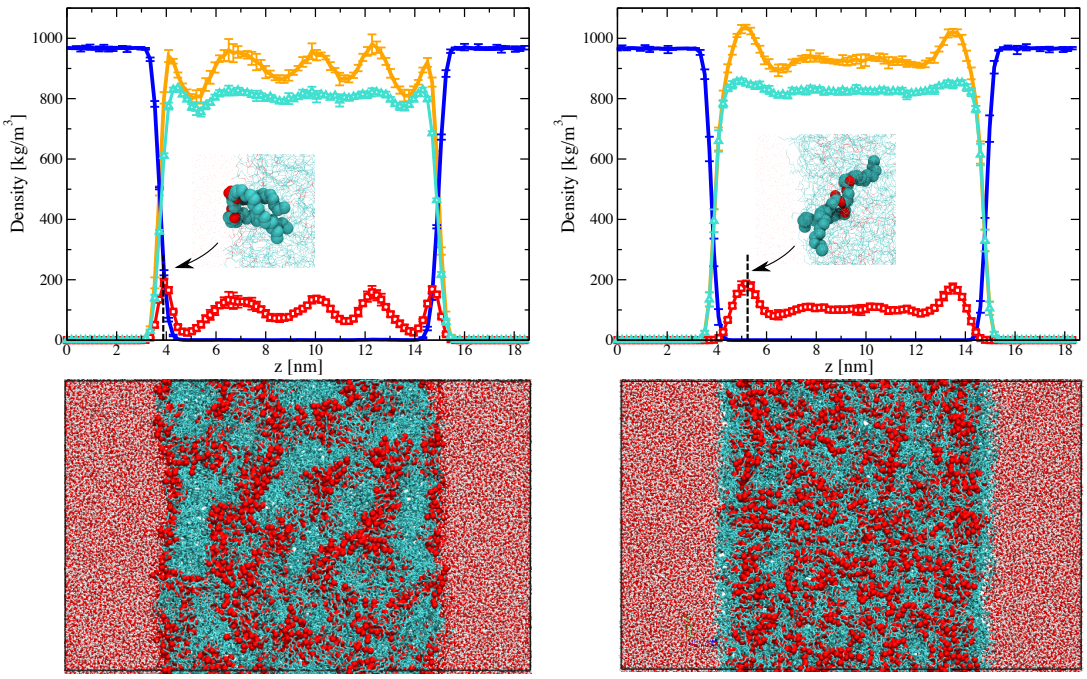


Figure 3.21: Comparison of the interaction between TG and water for FF1 (left) and FF2 (right) set of charges. Density profiles (on the top) and two snapshots of the system (bottom) are highlighted in the figure. The insets on the density profiles show representative interfacial TG molecules.

This difference is especially significant in the interaction of TG with water. As shown in Figure 3.21, the weak TG charges, in the FF2 case, fail to induce the reorientation of TG molecules at the interface. The density profile of the GLY group (red curve in Figure 3.21-top right) shows a peak far from the interface, with the hydrophobic fatty acid tails placed at the TG-water interface. This arrangement disagrees with experimental results on TG-water interface[52].

These simulations show that the set of charges can have an impact on the results and that the low charge values proposed in ref. 40 fails to reproduce the physical properties of TG in bulk and interacting with water, such as the GLY clustering and water penetration. Beyond the results presented in this work these features of TG molecules have been established in several studies and books[11, 17, 19]. Therefore, it would be desirable to repeat the study proposed in ref. 40, with an updated set of charges.

3.5 Conclusion and Future Work

A computational model of sebum oil has been introduced in this chapter. The sebum is represented through its most abundant component, the triglyceride tri-cis-6-hexadecacenoin (TG). The model has been tested by computing the structural and interfacial properties of TG. The united atom Berger forcefield, employed in this analysis, performed well in reproducing the density of the melted triglycerides with similar structures, the surface tension of the sebum oil at the air interface and the collapse pressures of TG-monolayers at the air-water interface. 100s nanosecond simulations show that triglycerides in the liquid phase feature structural ordering, with all TG head-groups clustered together in a single percolating network. The clusters shape is elongated resembling cylindrical inverse micelles. The simulations show that bulk TG molecules in the liquid phase can assume a wide range of different conformations, with a preference for conformations where two of the acyl tails are pointing in opposite direction with respect to the other one and the *tuning fork* being the most stable one ($\approx 40\%$ of the whole population).

To reduce the energy penalty induced by air and water interfaces and to lower the corresponding interfacial free energy, the interfacial TG molecules adopt a preferential orientation. When in contact with vacuum the TG acyl chains, featuring a lower surface tension than polar molecules in contact with air, orient towards the interface. In contact with water, instead, the glycerol group of the interfacial TG molecules preferentially interacts with water and it is absorbed at the water interface, adopting a *trident* conformation (assumed by ∼70% of the interfacial TG).

Addressing the impact the position of the unsaturation of the chains has on triglyceride properties is an interesting question, particularly when considering the uniqueness of the lipids on the skin surface. The lack of a clear understanding of the role that the double bond position has in determining the interfacial properties of TG fluids was pointed out. A possible hypothesis is that a shift of the double bond to a position closer to the polar part of the molecule, will results in a lower surface tension at the air and water interfaces. This effect will be driven likely by the influence of this parameter on the packing and the conformation of the TG molecules at the interface. Further experimental and computational studies are required to confirm this hypothesis, that will be the object of future works.

The possibility of spontaneous permeation of water molecules into the TG phase has been demonstrated here. The number of molecules diffusing into the bulk TG phase in hundreds of nanoseconds was found to be statistically significant and enabled the computation of the free energy barrier to water permeation. The magnitude of the activation energy, $\Delta G \approx 9.5 k_B T$, is relatively low and suggests that water molecules can permeate the TG layer in a short time scale ($\approx$ 100 ns). The estimated difference in free energy between water molecules in bulk water and water molecules in the TG phase is $\approx 7 k_B T$, and reflects the water solubility in the sebum oil. Moreover, the density profiles of water within the TG-phase and TG oxygens feature a similar behaviour, with corresponding maxima and minima (see Figure 3.15), showing that the driving force behind water permeation is the interaction between water molecules and TG-oxygens. This result suggests that triglyceride molecules may play an important role in the water-regulation function of the human sebum on the skin surface, enabling a controlled permeation of water through the skin surface and therefore the hydration of the outermost layers of the skin. The pathway for water permeation is provided by the percolating network

built via the interactions between polar group of TG molecules.

Finally, the TG monolayer at the water-air interface was investigated and the surface-area isotherm curve corresponding to the physiological temperature of our skin, $T = 310\,\mathrm{K}$, computed. The monolayer collapses without undergoing a liquid-expanded to liquid-condensed phase transition which is likely prevented by the presence of the double bond in the TG chains. The absence of this transition agrees with experimental observation reported in ref. 64, obtained with unsaturated triglycerides.

The study presented in this chapter represents a first step in the characterisation of the sebum molecules with a view to addressing the properties of the skin surface and its wettability. The latter will be investigated in the forthcoming chapters.

Bibliography to chapter 3

[1] A. S. Tascini, M. G. Noro, R. Chen, J. M. Seddon and F. Bresme, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1848–1860.

[2] C. C. Zouboulis, *Clinics in Dermatology*, 2004, **22**, 360–366.

[3] A. Pappas, *Dermato-Endocrinology*, 2009, **1**, 72–76.

[4] C. C. Zouboulis, M. Picardo, Q. Ju, I. Kurokawa, D. Törocsik, T. Bíró and M. R. Schneider, *Reviews in Endocrine and Metabolic Disorders*, 2016, **17**, 319–334.

[5] J. Fluhr, P. Elsner, E. Berardesca and H. I. Maibach, *Bioengineering of the Skin - Water and the Stratum Corneum*, CRC Press edn, 2004.

[6] J. C. Tsai, C. C. Lu, M. K. Lin, J. W. Guo and H. M. Sheu, *Skin pharmacology and physiology*, 2012, **25**, 124–32.

[7] F. P. Radner and J. Fischer, *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 2014, **1841**, 409–415.

[8] J. W. Fluhr, M. Mao-Qiang, B. E. Brown, P. W. Wertz, D. Crumrine, J. P. Sundberg, K. R. Feingold and P. M. Elias, *Journal of Investigative Dermatology*, 2003, **120**, 728–737.

[9] K. R. Smith and D. M. Thiboutot, *The Journal of Lipid Research*, 2007, **49**, 271–281.

[10] N. Nicolaides, *Science*, 1974, **186**, 19–26.

[11] A. Hall, J. Repakova and I. Vattulainen, *Journal of Physical Chemistry B*, 2008, **112**, 13772–13782.

[12] J. Telenius, A. Koivuniemi, P. Kulovesi, J. M. Holopainen and I. Vattulainen, *Langmuir*, 2012, **28**, 17092–17100.

[13] P. Kulovesi, J. Telenius, A. Koivuniemi, G. Brezesinski, I. Vattulainen and J. M. Holopainen, *Soft Matter*, 2012, **8**, 5826.

[14] E. DePeters, J. German, S. Taylor, S. Essex and H. Perez-Monti, *Journal of Dairy Science*, 2001, **84**, 929–936.

[15] W. D. Hsu and A. Violi, *Journal of Physical Chemistry B*, 2009, **113**, 887–893.

[16] M. Iwahashi and Y. Kasahara, *Current Opinion in Colloid & Interface Science*, 2011, **16**, 359–366.

[17] K. Larsson, *Fette, Seifen, Anstrichmittel*, 1972, **74**, 136–142.

[18] S. Ueno and A. Minato, *J. Phys. Chem. B*, 1997, **5647**, 6847–6854.

[19] D. J. Cebula, D. J. Mcclements, M. J. W. Povey and P. R. Smith, *Journal of the American Oil Chemists Society*, 1992, **69**, 130–136.

[20] R. W. Corkery, D. Rousseau, P. Smith, D. A. Pink and C. B. Hanna, *Langmuir*, 2007, **23**, 7241–7246.

[21] C. H. Rycroft, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2009, **19**, 041111.

[22] J. W. Carter, A. S. Tascini, J. M. Seddon and F. Bresme, *Computational Tools for Chemical Biology*, The Royal Society of Chemistry, 2018, ch. 2, pp. 39–68.

[23] I. Vattulainen and T. Rog, *Cold Spring Harbor Perspectives in Biology*, 2011, **3**, a004655.

[24] I. Chandrasekhar and W. F. van Gunsteren, *Current science*, 2001, **81**, 1325–1327.

[25] I. Chandrasekhar and W. F. Van Gunsteren, *European Biophysics Journal*, 2002, **31**, 89–101.

[26] A. K. Sum, M. J. Biddy, J. J. de Pablo and M. J. Tupy, *The Journal of Physical Chemistry B*, 2003, **107**, 14443–14451.

[27] G. Henneré, P. Prognon, F. Brion and I. Nicolis, *Chemistry and Physics of Lipids*, 2009, **157**, 86–93.

[28] O. H. S. Ollila, A. Lamberg, M. Lehtivaara, A. Koivuniemi and I. Vattulainen, *Biophysical*

Journal, 2012, **103**, 1236–1244.

[29] A. Bacle, R. Gautier, C. L. Jackson, P. F. Fuchs and S. Vanni, *Biophysical Journal*, 2017, **112**, 1417–1430.

[30] W. K. Stiefel, *Skin treatment with emollient cis-6-hexadecenoic acid or derivatives*, 1977, `http://www.google.co.uk/patents/US4036991`.

[31] O. Berger, O. Edholm and F. Jähnig, *Biophysical Journal*, 1997, **72**, 2002 – 2013.

[32] C. Das, M. G. Noro and P. D. Olmsted, *Biophysical Journal*, 2009, **97**, 1941–1951.

[33] C. Das, M. G. Noro and P. D. Olmsted, *Soft Matter*, 2014, **10**, 7346–7352.

[34] R. Thind, D. W. O'Neill, A. Del Regno and R. Notman, *Chem. Commun.*, 2015, **51**, 5406–5409.

[35] W. L. Jorgensen and J. Tirado-Rives, *Journal of the American Chemical Society*, 1988, **110**, 1657–1666.

[36] W. van Gunsteren, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Biomos ; Zürich, 1996.

[37] J.-P. Ryckaert and A. Bellemans, *Chemical Physics Letters*, 1975, **30**, 123 – 125.

[38] D. Peter Tieleman, H. J. Berendsen and M. S. Sansom, *Biophysical Journal*, 1999, **76**, 1757–1769.

[39] *Berger database, downloadable at*, `http://moose.bio.ucalgary.ca/files/`, Accessed: 2016-11-28.

[40] A. Akinshina, C. Das and M. Noro, *Phys. Chem. Chem. Phys.*, 2016, **18**, 17446–17460.

[41] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *Journal of Computational Chemistry*, 2005, **26**, 1701–1718.

[42] G. Bussi, D. Donadio and M. Parrinello, *Journal of Chemical Physics*, 2007, **126**, 014101.

[43] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *The Journal of Chemical Physics*, 1995, **103**, 8577–8593.

[44] K. Zimmermann, *Journal of Computational Chemistry*, 1991, **12**, 310–319.

[45] K. A. Dowsland and J. M. Thompson, *Handbook of Natural Computing*, 2012, vol. 4-4, pp. 1623–1655.

[46] M. González-Melchor, F. Bresme and J. Alejandre, *The Journal of Chemical Physics*, 2005,

122, 104710.

[47] I.-C. Yeh and M. L. Berkowitz, *The Journal of Chemical Physics*, 1999, **111**, 3155.

[48] D. Mohammad-Aghaie, E. Mace', C. A. Sennoga, J. M. Seddon and F. Bresme, *Journal of Physical Chemistry B*, 2010, **114**, 1325–1335.

[49] Y. Zhang, J. W. Carter, A. Lervik, N. J. Brooks, J. M. Seddon and F. Bresme, *Soft Matter*, 2008, **12**, 2108–2117.

[50] M. González-Melchor, F. Bresme and J. Alejandre, *The Journal of Chemical Physics*, 2005, **122**, 104710.

[51] C. Vega and E. de Miguel, *The Journal of Chemical Physics*, 2007, **126**, 154707.

[52] H. F. DeLuca, D. J. D. J. Hanahan and A. Kuksis, *Handbook of lipid research*, Plenum Press, 1978.

[53] *CAS Chemnet.com The physical and chemical property of 20246-55-3, tripalmitolein (C16:1,-cis-9) grade I is provided by ChemNet.com*, `http://www.chemnet.com/cas/en/20246-55-3/tripalmitolein-(C16:1,-cis-9)-grade-I.html`, Accessed: 2017-03-29.

[54] W. M. Haynes, D. R. Lide and T. J. Bruno, *CRC handbook of chemistry and physics : a ready-reference book of chemical and physical data - 98th edition.*, 2017-2018.

[55] W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33–38.

[56] A. Varshney, F. P. Brooks and W. V. Wright, *IEEE Computer Graphics and Applications*, 1994, **14**, 19–25.

[57] R. R. Benerito, W. S. Singleton and R. O. Feuge, *The Journal of Physical Chemistry*, 1954, **58**, 831–834.

[58] L. D. A. Chumpitaz, L. F. Coutinho and A. J. A. Meirelles, *Journal of the American Oil Chemists' Society*, 1999, **76**, 379–382.

[59] T. Xu, V. Rodriguez-Martinez, S. N. Sahasrabudhe, B. E. Farkas and S. R. Dungan, *Food Biophysics*, 2017, **12**, 88–96.

[60] E. O. Butcher and A. Coonin, *Journal of Investigative Dermatology*, 1949, **12**, 249–254.

[61] P. Schatzberg, *The Journal of Physical Chemistry*, 1963, **67**, 776–779.

[62] S.-J. Marrink and H. J. C. Berendsen, *J. Phys. Chem*, 1994, **98**, 4155–4168.

[63] C. Das, P. D. Olmsted and M. G. Noro, *Soft Matter*, 2009, **5**, 8.

[64] M. C. Michalski, P. Brogueira, A. Gonçalves da Silva and B. Saramago, *Eur. J. Lipid Sci. Technol.*, 2001, **103**, 677–682.

[65] D. Mohammad-Aghaie and F. Bresme, *Molecular Simulation*, 2016, **42**, 391–397.

[66] J. L. Abascal and C. Vega, *The Journal of chemical physics*, 2005, **123**, 234505.

[67] S. Izadi, R. Anandakrishnan and A. V. Onufriev, *The Journal of Physical Chemistry Letters*, 2014, **5**, 3863–3871.

[68] M. Javanainen, A. Lamberg, L. Cwiklik, I. Vattulainen and O. H. Ollila, *Langmuir*, 2018, **34**, 2565–2572.

Chapter 4

How the sebum oil interacts with SC lipids

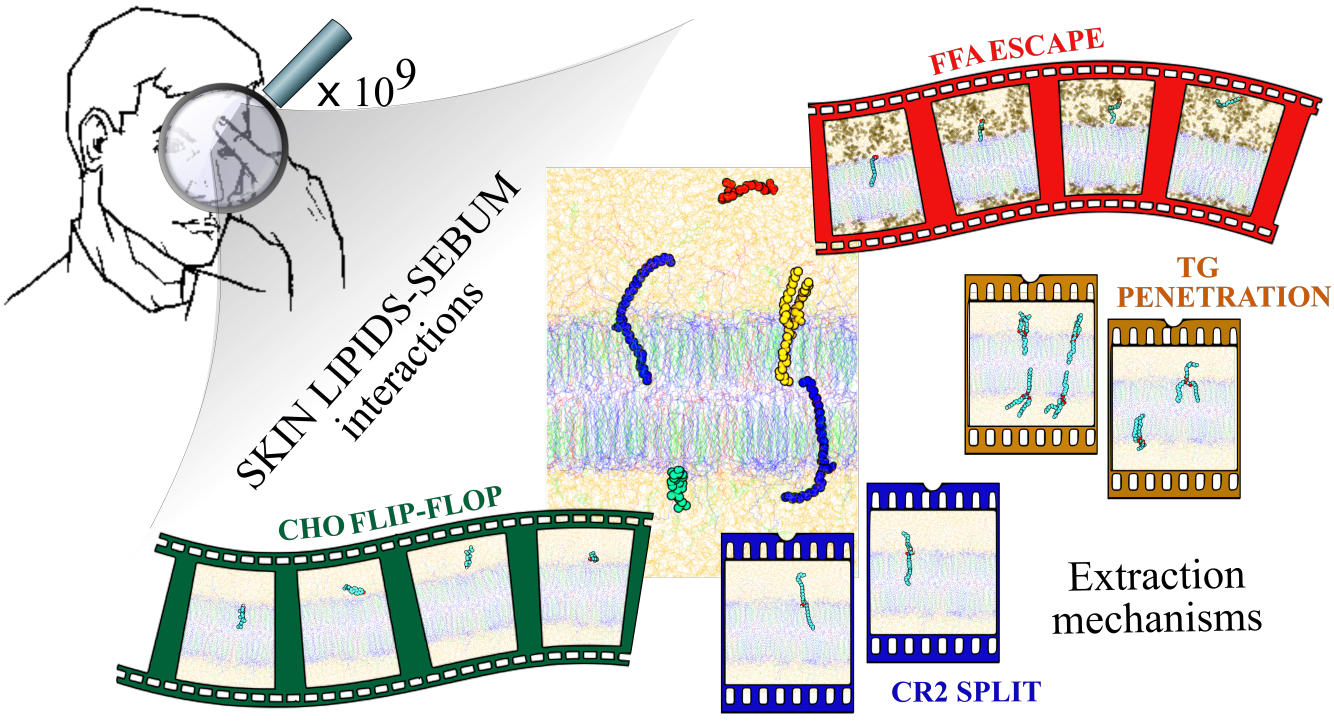


Figure 4.1: Title page figure for chapter 4. This image has been used as table of contents for Tascini, A. S. *et al. Phys. Chem. Chem. Phys.*, 2018, submitted. It was also used for the CCP5 outreach library[1] and contest, for which I won the first prize ex–aequo[2].

4.1 Introduction

The skin surface is the outermost protective barrier between our bodies and the external environment. There are three possible routes for external pathogens, particles and molecules to penetrate the skin barrier: transcellular delivery, intercellular devilery (*i.e.* through the skin lipid mixture), and via the skin appendages, namely the hair follicles (appendageal route), as shown in Figure 1.3. In the former two routes the main barrier for the absorption is represented by the outermost layer of the skin, the SC (SC). The challenges associated to the transcellular and intercellular pathways are the rigidity of the corneocytes envelope[3] and the small permeability of "crystalline" SC lipid mixture[3, 4, 5], respectively. The hair follicles route, instead, offers an opportunity for the permeation of slowly diffusing compounds and very high molecular weight assemblies, such as nanoparticles[6, 7] . The sebum, which fills the gap between the hair follicle and the skin, is the main transport route of the follicular pathway[8]. The high permeation ability of follicles has promoted the development of transdermal drug delivery approaches which target this pathway[6, 9, 10, 11].

The skin surface is coated with a lipid film, composed of a mixture of sebaceous lipids (the sebum oil) and a small fraction of epidermis lipids that can detach from the SC[12, 13]. Depending on the sebaceous lipid concentration on the surface, different anatomical sites can be described as sebum-rich areas, *e.g.* the forehead, and sebum-poor areas, *e.g.* the forearm[14]. The distribution of these areas on the surface of our body corresponds to the location of the hair follicles, through which the sebum oil is released onto the surface. As discussed in the previous chapter, the sebum oil might play a key role in regulating the hydration of the skin and its waterproofing properties[15, 16, 17]. It also features antimicrobial properties[18] and helps in the transport of antioxidants, such as vitamin E[19, 20]. The lack (*hyposeborrhoea*) and the excess (*hyperseborrhoea*) of sebum make the skin more fragile and vulnerable to irritations. Typically, sebum-poor areas are often subjected to discomfort and itching, being more sensitive to all external aggressive agents. Sebum rich areas have instead an higher concentration of antioxidants[20].

For the transcellular, intercellular and appendageal routes, a microscopic understanding of the

properties of the lipid film placed on the skin surface is essential to explain the performance of the skin as a barrier for water and foreign molecule penetration, as well as to design strategies for transdermal drug delivery. Hence, unravelling the interaction between SC lipids and sebum becomes crucial.

As discussed in chapter 2, molecular dynamics (MD) simulations provide a promising tool to characterise the properties of biological systems and to unravel the mechanisms of interaction between biological complexes with atomistic resolution[21]. Simulations have been previously used to describe the structure and phase behaviour of a single component SC lipid model in water[4, 5]. In addition, more complex mixtures of ceramides, cholesterol and free fatty acids have been introduced to understand the contribution of the different lipids to the SC performance and the relationship between the polydispersity of the lipids and the overall packing and fluidity of the resultant membrane[22, 23]. Further simulation studies have explored the effect of compounds such as ethanol[24], DMSO[5] and oils[25, 26] and attempted to characterise the permeation of small molecules through homogeneous skin-lipid membranes[27].

Despite the experimental evidence of interactions between the SC and the sebum oil[13], little attention has been given to the impact that the oil layer might have on the properties of the skin surface. As described in chapter 3, a model of the sebum oil has been recently proposed and its interaction with water has been investigated[17]. Triglycerides, the sebum's most abundant component, are shown to play an important role in the water-regulation function of the human sebum on the skin surface, and might contribute to a controlled permeation of water through the skin surface to ensure a controlled hydration of the outermost layers of the skin. However, how sebum triglycerides interact with SC lipids has remained unexplored and will be addressed here.

This chapter discusses MD simulations of the interaction between sebum triglycerides (TG) and a multicomponent model of the skin, incorporating the three most abundant constituents of the SC: ceramide 2, cholesterol and lignoreic acid. The penetration mechanisms of TG into the SC membrane and three different mechanisms for SC lipid extraction are here uncovered which define the routes followed by SC lipids when they are extracted from the bilayer, and

that they might explain the presence of SC lipids in the sebum films coating the skin surface, previously reported in *in vivo* experiments on human skin[13]. Furthermore, the structural and mechanical properties of a dry bilayer of SC lipids in TG are here evaluated, to investigate the stability of the bilayer in the proximity of the sebum oil.

Part of the results presented in this chapter have been used for a paper draft which has been submitted at *Phys. Chem. Chem. Phys.*: Tascini A. S. *et al.*, Mechanisms of lipid extraction from skin lipid bilayers by sebum triglycerides, *Phys. Chem. Chem. Phys.*, 2018, submitted.

4.2 Methods

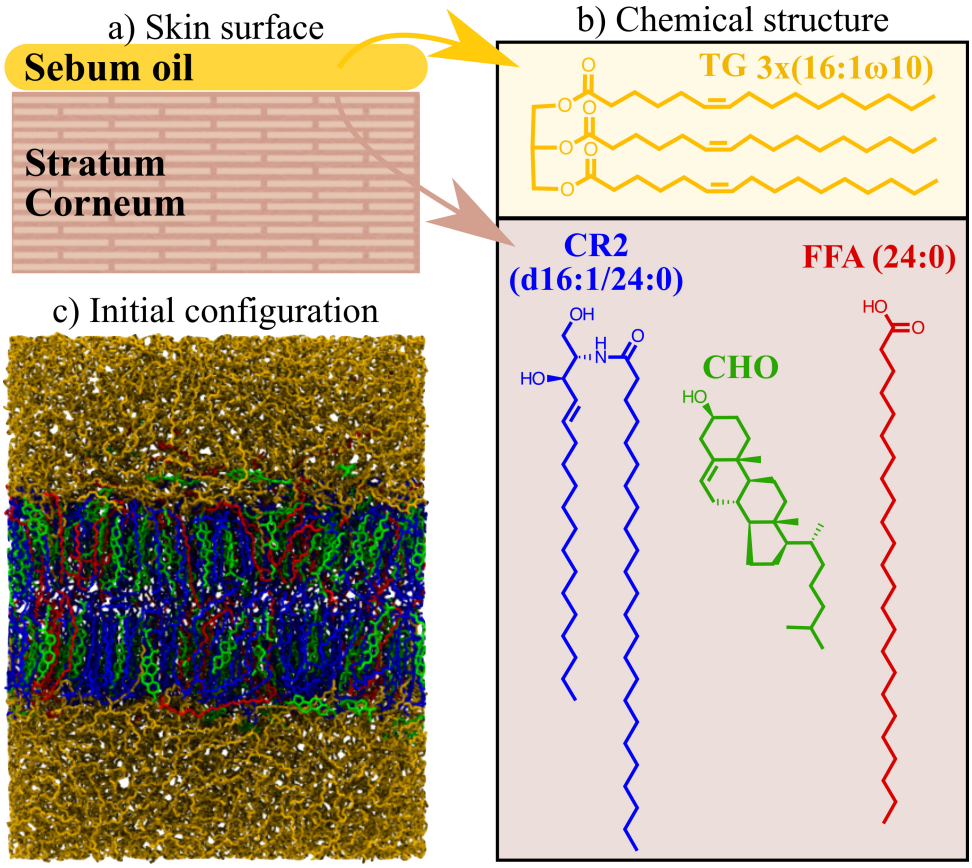


Figure 4.2: (a) Schematic representation of the top of the skin surface, where the sebum oil interacts with lipids from the outermost layer of the skin, the SC. (b) Chemical structure of the lipids investigated in this work. The sebum is modelled with its most abundant component, the triglyceride tri-cis-6-hexadecenoin (TG). The multicomponent model of the SC is made up of ceramide 2 (CR2), cholesterol (CHO) and lignoceric acid (FFA). (c) Representative snapshot of the systems used in this analysis. The colour code is the same as the chemical structures.

4.2.1 Input configurations and simulation details

Following previous studies[4, 28], the SC lipid matrix was modelled as a ternary mixture of its three main components: ceramides 2 (CR2), cholesterol (CHO) and free fatty acids (FFA), here represented by lignoceric acid. The chemical structures of the three components are displayed in Figure 4.2b. A molar ratio of the three components (CR2:CHO:FFA) equal to 2:2:1 was used to mimic the structural and mechanical properties of healthy skin[4].

The SC membrane was constructed using the open source software Packmol[29] to place 288 SC lipid molecules in each of the two leaflets of a bilayer, symmetrically about the bilayer mid-plane. Packmol was used to place the lipid and water molecules inside rectangular regions with a random distributions. To better reproduce the bilayer structure, the top oxygen atoms of the OH group of the hydrophilic head groups and the terminal carbon atom of the longest acyl tail of each lipid molecule were constrained to lie above or under predefined planes (as shown in Figure 4.3). A distance of 1.4 Å between molecules was adopted to avoid overlapping between them. The system was then equilibrated in SPC water, following the procedure described in ref. 4, *i.e.* the system was run for 5 ns with the SC molecules fixed (to hydrate the bilayer properly), and then for 2 ns with only the terminal methyl group on all the lipid tails frozen, so that the rest of the molecules reorient themselves to accommodate the water environment.

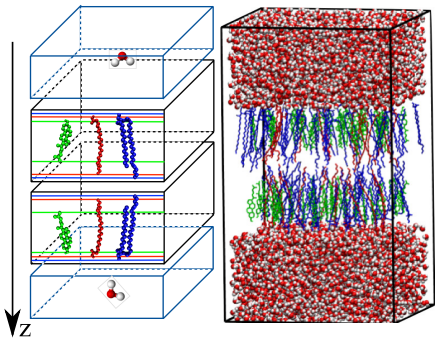


Figure 4.3: The scheme on the left shows the constrains adopted in Packmol[29] to create the bilayer configuration. The output configuration is shown on the right.

The sebum triglyceride, tri-cis-6-hexadecenoin, was used to represent the human sebum, as discussed in chapter 3 and in ref. 17. The chemical structure of a TG molecule is shown in Figure 4.2b. Starting from an equilibrated configuration of the SC bilayer in water, the solvent molecules were replaced with two sebum slabs, made up of 256 TG molecules each. The TG slabs were previously equilibrated for 300 ns following the procedure described in chapter 3 and in ref. 17. A SC-TG configuration, used throughout this work, is shown in Figure 4.2c.

Bonded and non-bonded interactions of all the lipid molecules employed here were described us-

ing the united atom forcefield by Berger[30], previously used in simulations of skin lipids[4, 24, 28]. This parameter set is based on OPLS[31] and GROMOS87[32] forcefields and includes the Ryckaert-Bellemans dihedral potential[33] to model the hydrocarbon chains. The unsaturated region of the acyl tails of the TG molecule was described according to the description of the double bond in the Berger POPC molecule[34] and the *cis* geometry was enforced using a dihedral potential. van der Waals interactions were truncated at 1.2 nm and electrostatic interactions were computed in full using the Particle Mesh Ewald algorithm[35]. The intra-molecular bonds were constrained using the P-LINCS[36] algorithm and periodic boundary conditions were applied in all directions.

After joining the SC bilayer and the TG slabs, a 500 ps equilibration was realised in the NVT ensemble. Temperature was kept constant at 310 K with the velocity rescaling algorithm[37], using a time coupling constant, τ, of 0.1 ps. This was followed by a 400 ns run in the NPT ensemble ($T = 310$ K and $p = 1$ bar), with the Nosé-Hoover thermostat, with $\tau = 1$ ps and the Parinello-Rahaman barostat, applied semi-isotropically with $\tau = 5$ ps and compressibility 4.5×10^{-5} bar^{-1}. This was followed by a 2 µs production run in the NPT ensemble using the same conditions and simulation set-up. The time step of the simulations was 0.002 ps.

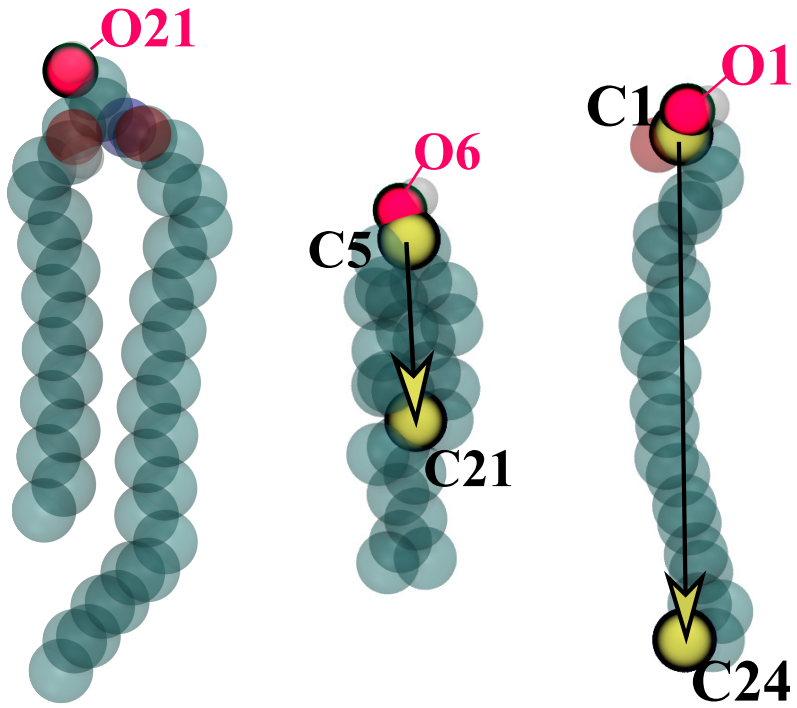
All the simulations were performed with the GROMACS/4.5.5 molecular simulation package[38].

4.2.2 Multicomponent bilayer analysis

Many of the results presented below, *i.e.* bilayer composition, number of extraction events and area per lipid, were obtained through the method explained below, which was implemented in a in-house code.

Definition of a slice within the SC bilayer

In order to identify the boundaries of the lipid bilayer one atom of each lipid type was considered to determine the position of the lipid head groups. Specifically, atoms O21 for CR2, O6 for CHO, O1 for FFA were chosen (see Figure 4.4 for labels). The z positions of the lipid head

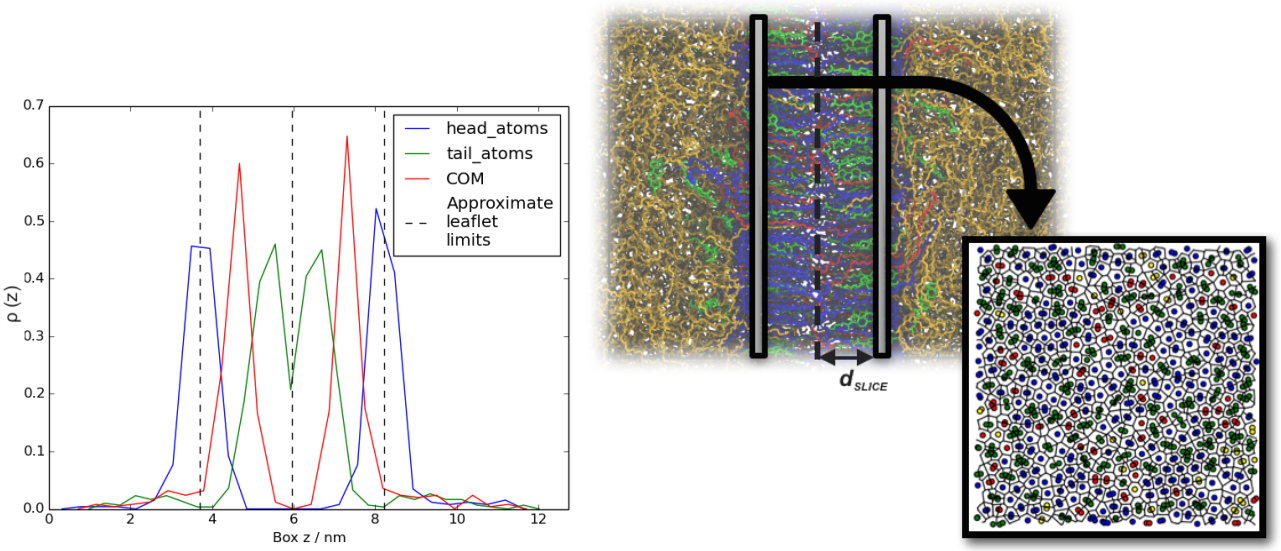


(a) CR2 **(b)** CHO **(c)** FFA

Figure 4.4: (a) CR2 molecule with atom label for O21. (b) CHO molecule with atom labels for O6, C5 and C21. (c) FFA molecule with atom labels for O1, C1 and C24. The lipid molecules are represented with translucent beads with names and coloured beads shown for the labelled atoms.

group atoms were histogrammed (see blue line in Figure 4.5a) and the value of the two maxima in the histograms were used to define the edges of the bilayer (see dashed lines). The bilayer midpoint was defined as halfway between these values. Once the bilayer midpoint was calculated from the histograms, two slices through the bilayer, parallel to the xy plane, were defined at a specified height, d_{SLICE}, above and below the bilayer midpoint. The height, d_{SLICE}, is chosen so that the slice is entirely within one of the leaflets of the bilayer, just below the head group maxima, and does not include the bilayer surface. The slice thickness must be larger than the largest bond length for molecules within the bilayer to ensure that if the slice passes though a molecule, at least one of the molecules atoms are within the slice volume. A graphical example of the definition of the slices is shown in Figure 4.5b. In this analysis d_{SLICE} was set to 1.45 nm and the slice thickness to 0.15 nm.

The atoms in each slice were recorded and used for determining the bilayer properties described below.



(a) Histogram **(b)** Bilayer slice

Figure 4.5: a) Histogram of lipid subgroups, namely head groups (in blue), tails (in green) and centre of mass (red). The leaflets limits and the bilayer mid-plane, estimated by the code, are shown with dashed lines. b) Representative system configuration where grey boxes highlight the position of the two planar slides used to implement the Voronoi tesselation method. The inset shows the corresponding bilayer cross section, with the resulting Voronoi tessellation. Blue lipid and dots correspond to CR2, red to FFA, green to CHO and yellow to TG.

Assessment of the insertion and extraction events

The atoms in each slice were recorded and individual molecules were assigned to one of the two leaflets or to the TG phase by checking whether at least one of their atoms was inside or outside the slices. Specifically, if a SC lipid molecule did not have atoms in either of the two slices, it was assigned as being outside the bilayer and one extraction event was counted. If a TG molecule was found to be in one of the slices it was considered to be anchored to the bilayer. For molecules with more than one chain, *i.e.* TG and CR2, each chain was accounted for separately in order to assign part of the molecule to the bilayer or the TG phase, based on the positions of the atoms in the chain. The resulting configuration of TG or CR2 molecules was then determined by considering the relative assignment of the individual chains.

In TG anchoring the three fatty acid chains were treated separately to check whether TG molecules insert with one, two or three chains and assess the conformation of the inserted molecule. The TG atoms were assigned to the three chains as specified in Figure 4.6.

For the CR2 split event, *i.e.* when a CR2 molecule has the two chains widely spread forming

an angle of $\approx 180°$ with one of the chain in the bilayer and the other one in the TG phase, it was distinguished whether the chain within the slice was the sphingosine chain or the fatty acid chain. The CR2 atoms were assigned to the chains as shown in Figure 4.6. A full CR2 extraction was counted when all the atoms from both chains were found outside the slices.

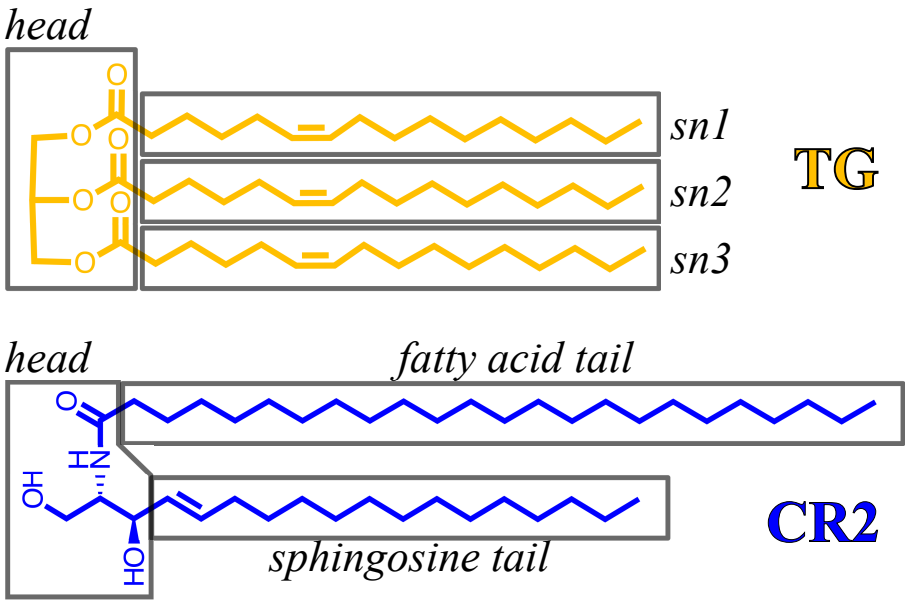


Figure 4.6: Top) The atoms of a TG molecule were separately assigned to one of the three fatty acid chains (*sn1*, *sn2* or *sn3*) or to the glycerol group (*head*). Bottom) Atoms in a CR2 molecule are divided in *head* group atoms, *fatty acid* tail and *sphingosine* tail. The atom assignment is shown with grey rectangular boxes.

Area per lipid

The atom coordinates within the slices were used as generators for the Voronoi tessellation method in two dimensions (see Figure 4.5b). To any atom was assigned its closest region of the 2D space.

Once the Voronoi tessellation was obtained, the area of a planar non-self-intersecting polygon with vertices (x_1, y_1), ..., (x_n, y_n) was calculated by[39]

$$A_P = \frac{1}{2}\left(\begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} + \begin{bmatrix} x_2 & x_3 \\ y_2 & y_3 \end{bmatrix} + \cdots + \begin{bmatrix} x_{n-1} & x_n \\ y_{n-1} & y_n \end{bmatrix} + \begin{bmatrix} x_n & x_1 \\ y_n & y_1 \end{bmatrix}\right). \tag{4.1}$$

The area per lipid of each molecule (i) was obtained by averaging the contributions of the single

polygons, A_P, as

$$A_{L,i} = \langle \frac{\sum_{N_L} A_P}{N_L} \rangle, \tag{4.2}$$

where N_L is the number of molecules of specie L found within the slice and $\langle ... \rangle$ represents a time average.

4.2.3 Free energy pathway of lipid extractions

The free energy for lipid extraction from the bilayer to the sebum phase can be estimated by analysing the density distribution of the SC lipids.

A head group atom for each SC lipid species was selected to track the position of the lipids along z: O21 for CR2, O6 for CHO, O1 for FFA (see Figure 4.4 for the atom names). From the coordinates of the head group atoms along the simulation box, it is possible to determine and estimate the probability of the atoms visiting a particular z coordinate,

$$Z(i) = \frac{\overline{\rho}(z)}{\overline{\rho}(0)},$$

where $\overline{\rho}(z)$ is the average number of lipids at a particular z and $\overline{\rho}(0)$ indicates the reference state, located at the minimum of the distribution. The free energy profile, ΔG, can therefore be obtained from $Z(i)$, by

$$\Delta G = G(z) - G(0) = -k_B T \ln \left(\frac{\overline{\rho}(z)}{\overline{\rho}(0)} \right). \tag{4.3}$$

Another important parameter for the characterisation of the extraction process is the tilt angle (θ_T) of the different lipids. The tilt angle has been found to be critical in explaining intra-leaflet dynamics in lipid membranes, *e.g.* cholesterol flip-flop[40]. The tilt angle of CHO and FFA molecules, along with the position of the head group, were used as reaction coordinates to build the 2D energy landscape, $\Delta G(\rho(z), \theta_T)$, of lipid extraction and to discuss the mechanism of extraction from the bilayer, as reported in the Results section below. The tilt angle is defined as the angle between the lipid molecular axis and the z axis. The molecular axis of CHO was

defined as the vector between atoms C5 and C21, *i.e.* the axis of the rigid ring structure (see Figure 4.4 for the atom names). The FFA molecular axis was defined, instead, from the carbon atoms at the edges of the acyl chain, C1 and C24 (see Figure 4.4 for the atom names).

4.3 Results

4.3.1 Mechanisms of interdigitation of TG and SC

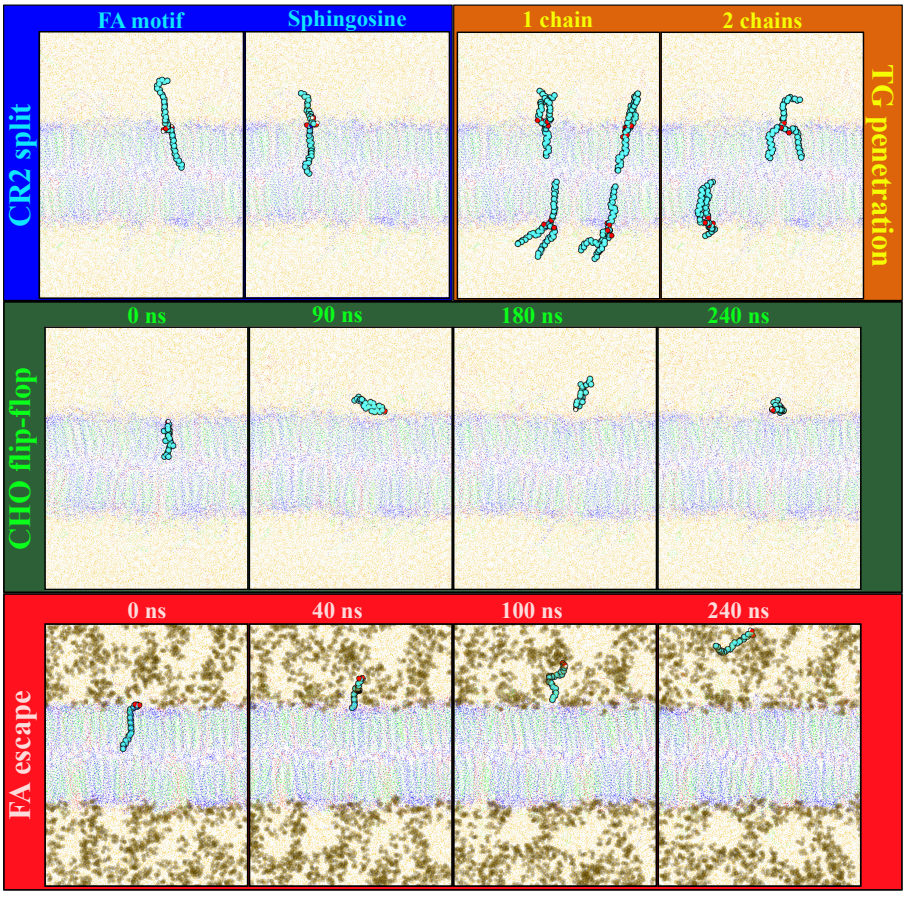


Figure 4.7: Representative snapshots illustrating the main mechanisms of lipid interdigitation. Each lipid contribution is represented in a coloured box. Blue for CR2 splitting, with the free fatty acid (left) or the sphingosine motif (right) in the oily layer; orange for TG penetration with one (left) or two (right) chains in the SC bilayer; green for CHO flip-flop at different time-steps (from left to right: 0, 90, 180, 240 ns); red for FFA escape at different time-steps (from left to right: 0, 40, 100, 240 ns).

When in contact with the SC lipid bilayer some of the interfacial sebum TG start interacting with the SC lipids by inserting their acyl tails inside the bilayer. In a similar fashion some of the SC lipids are extracted from bilayer and dissolve into the oily phase. A small concentration

of lipids detached from the SC mixture is indeed found in the lipid film covering the surface of our skin [13]. The simulations described in this chapter were used to understand the mechanisms of lipid extraction from the SC layer at the molecular level and understand how the outermost skin lipids interact with the sebaceous ones.

To characterise the interactions between TG and SC and the mechanisms of interdigitation, the structural and dynamical features of the lipid molecules involved in this process were examined. Representative simulation snapshots illustrating the different interdigitation mechanisms are summarised in Figure 4.7, where the different contributions are represented in coloured boxes, orange for TG, blue for CR2, green for CHO and red for FFA. Each mechanism is described and further characterised below. The mole fractions of SC lipids (distinguishing the contribution of CHO and FFA) in the sebum phase, χ_{SC}, with respect to the total number of lipids in the sebum phase, $\chi_{SC} = N_{SC}/(N_{TG} + N_{SC})$, are reported in Table 4.1.

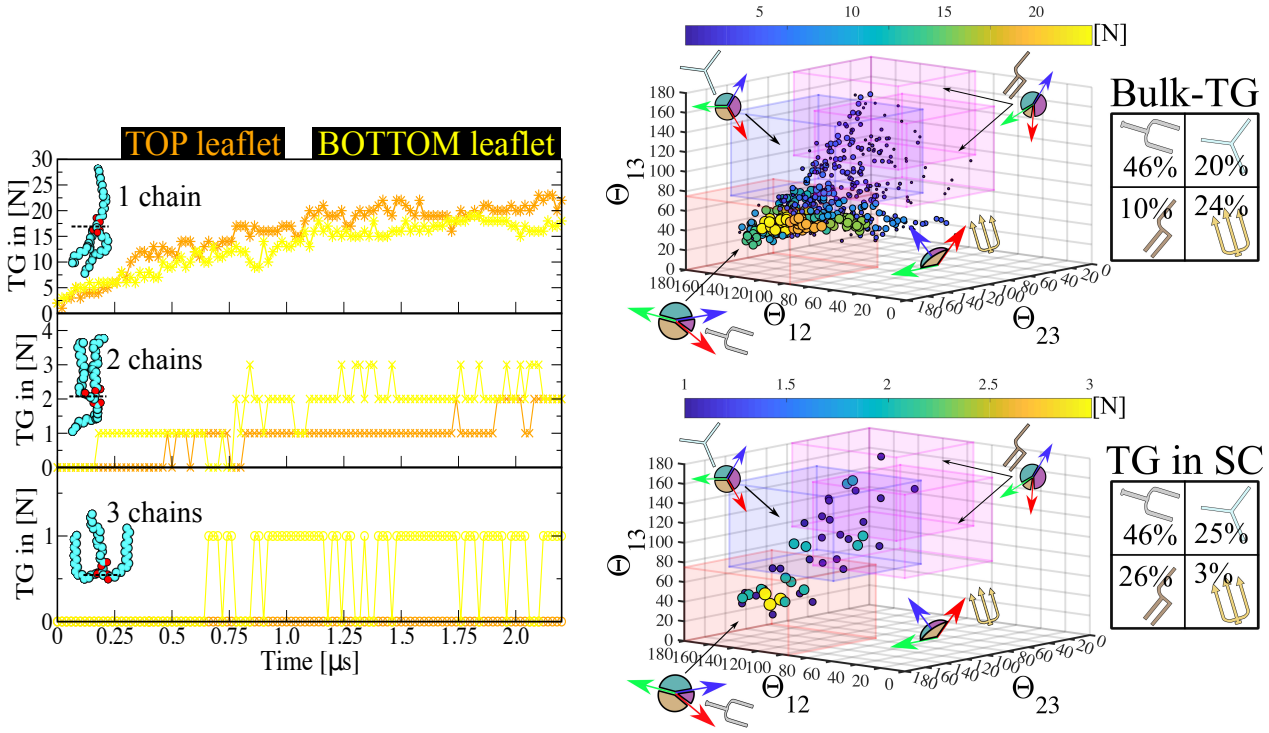
SC	$\chi_{\mathbf{SC}}$
CHO	0.06
FFA	0.04

Table 4.1: Mole fraction of SC lipids (CHO and FFA) extracted and dispersed in the TG phase, χ_{SC}, with respect to the total number of lipids in the oily phase, $\chi_{SC} = N_{SC}/(N_{TG} + N_{SC})$.

TG penetration

Some of the interfacial TG molecules penetrate the bilayer with their acyl chains. The number of insertion events *vs.* time in the top and bottom leaflets is reported in Figure 4.8a. The graph shows that after a steady increase in the number of insertions during the first $\approx 1.1\,\mu s$ of the trajectory, they tend to stabilise towards a constant value (≈ 20 insertions in each leaflet).

The most common type of interdigitation for TG molecules ($\approx 89.7\%$ of the events) is the insertion of one of the three tails, as displayed in the top graph of Figure 4.8a and in the left orange box of Figure 4.7. A few of them are instead penetrating the bilayer with two ($\approx 7.7\%$) or three ($\approx 2.6\%$) chains (see middle and bottom graphs of Figure 4.8a respectively). The insertion probability for more than one chain is much lower due to the higher energy barrier required to align more than one chain and penetrate the bilayer at two nearby points.



(a) Mechanisms of TG penetration **(b)** TG conformations analysis

Figure 4.8: (a) Number of TG molecules anchored to the bilayer *vs.* time in the bottom (yellow curves and symbols) and top (orange curve and symbols) leaflet, with one (top-graph), two (middle-graph) or three (bottom-graph) chains. (b) 3D distribution of triglyceride conformations as a function of the three angles described by the mutual orientation of the chains: θ_{12}, θ_{23} and θ_{13}, for bulk TG (top) and TG in SC bilayer (bottom). The tables next to the 3D-histograms show the % of occurrence of the different shapes.

The preferential conformations of the TG molecules were analysed with the method described in section 3.2.3 and in ref. 17, by averaging TG conformations over the last 0.1 μs of the simulation. Figure 4.8b shows the 3D distributions of TG conformations as a function of the three angles described by the mutual orientation of the chains: θ_{12}, θ_{23} and θ_{13} for bulk (top graph) and interdigitating (bottom graph) TG molecules. As previously discussed in chapter 3 and ref. 17, sebum TG in bulk tends to preferentially assume the shape of a tuning fork, with the two external chains (sn1 and sn3) pointing in the same direction and opposite to the direction of the central (sn2) chain. The presence of the SC bilayer does not affect this result, with $\approx 46\%$ of the molecules assuming the tuning fork conformation in the bulk TG. Interestingly, the number of trident conformations ($\approx 24\%$), instead, increases in comparison to the case of the pure TG phase ($\approx 15\%$ - as reported in chapter 3 and ref. 17). The presence of the SC lipids causes TG nearby the interface to preferentially rearrange their shapes with some of them becoming tridents. This observation is consistent with the structure of the TG molecules at the water-TG

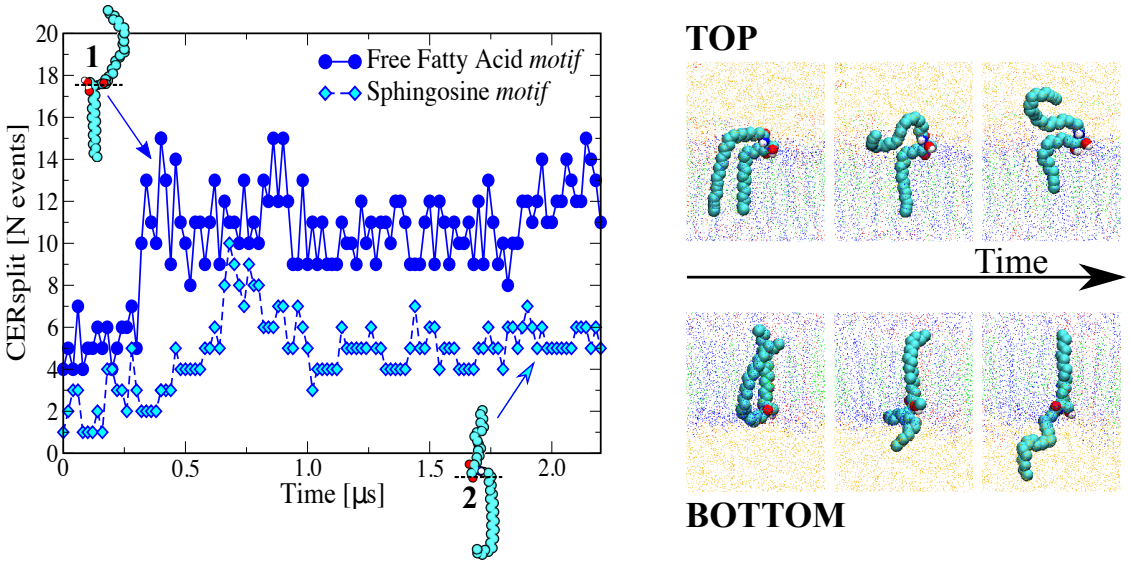
interfaces, where the trident conformation was also observed.

A preference for the tuning fork conformation is noticed in the interdigitating TG molecules, which preferentially insert the middle chain into the bilayer, reflecting the fact that this conformation is the most abundant in bulk. Other conformations, namely the propeller and the chair, are enhanced significantly with respect to the population observed among the non-anchored molecules, while the number of trident configurations is reduced significantly (see inset Figure 4.8b bottom, showing the shape %). The on-off fluctuations in Figure 4.8a bottom suggest that the third chain might be entering the bilayer only superficially, comparing this case to the insertion of only 2 chains.

CR2 splitting

CR2 molecules are stable in the bilayer and after 2 µs just a few molecules (3 out of 224 molecules) left the bilayer and were located at the TG-SC interface. This is due to the strong network of hydrogen-bonds established between CR2-CR2 head groups (as further discussed in section 4.3.2). However, some of the CR2 molecules ($\approx$ 18 out of 224 molecules - $\sim$ 10%) extend one of the acyl chains into the TG oily phase, adopting a conformation characterised by the two chains widely spread forming an angle of $\approx 180°$, as shown in the blue box of Figure 4.7. This event is described as CR2 splitting.

In Figure 4.9a, the number of CR2 splitting events *vs.* time is reported, differentiating between the conformations where the chain leaving the bilayer is the free fatty acid motif (blue circle and inset 1) or the spinghosine motif (light-blue diamonds and inset 2). The graph shows that the free fatty acid motif has a higher probability of leaving the bilayer when a splitting event occurs. This can be understood by considering the asymmetry of the CR2 tails. The longer C24 chain in the free fatty acid tail, in fact, tends to bend in the space between the two leaflets or in the head-group region[5]. When in this second conformation the energy penalty for the splitting event to occur is lower and the C24 chain interdigitates into the oily layer. Figure 4.9b shows the dynamics of this event for two CR2 molecules, from the top (above the time line) and bottom (below the time line) leaflets. In both cases the C24 chain first tilts at the interface



(a) CR2 splitting *vs.* time **(b)** C24 split dynamics

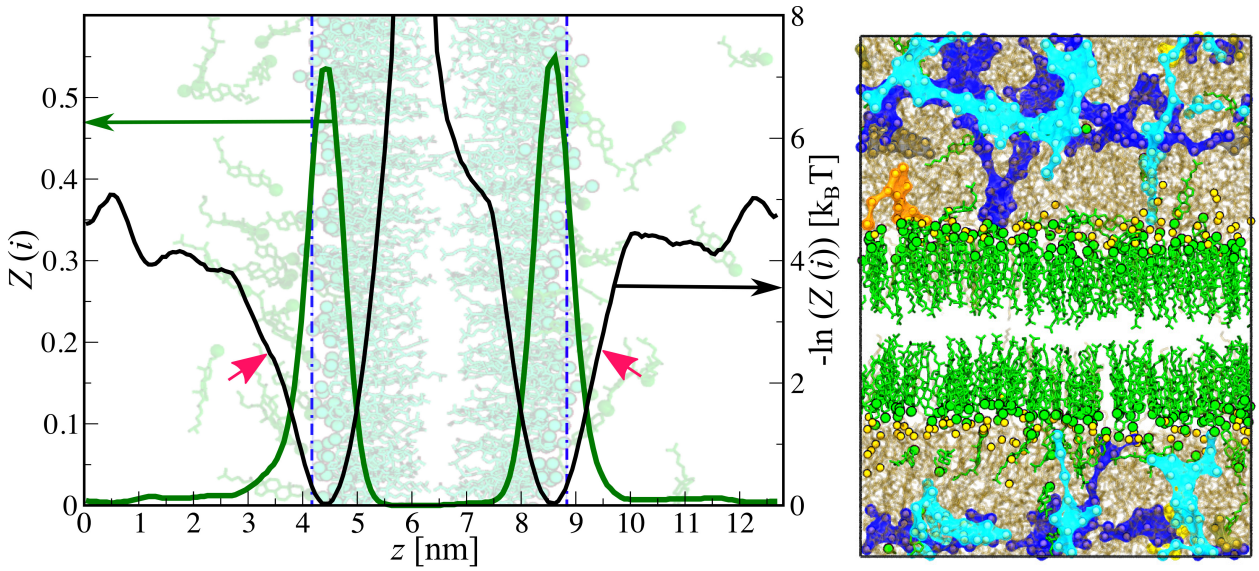
Figure 4.9: (a) Number of CR2 splitting *vs.* time. The blue circles correspond to the number of CR2 split with the C24 tail in the oily phase; the light-blue diamonds correspond to the case where the C18 tail is in the oily phase. Inset 1 and 2 show two representative CR2 splitting event in the top leaflet. (b) Two examples of C24 splitting in the sebum oil are represented through 3 consecutive snapshots for one molecule in the top (above the time line) and bottom (below the time line) leaflets.

and then goes to the TG phase.

The existence of extended configurations for ceramide molecules, such as the one observed here at the TG interface, has been previously noted in studies involving stacks of lipid bilayers[41, 42, 43, 44]. In those cases the bilayers in the middle of the stacks are in an anhydrous environment, as in our simulations. Moore *et al.*[44] reported, using coarse grained models, that in a stack of ceramide bilayers the percentage of ceramide molecules found in extended conformations increases in the inner leaflets where the lipid molecules are not in contact with water.

CHO extraction: a flip-flop event

A significant number of CHO molecules leave the bilayer to go to the TG phase (≈ 14% of the molecules are extracted - see Table 4.1 for the data expressed in terms of mol fraction). Figure 4.10a shows the density profile of the CHO head group (oxygen, O6) obtained from the analysis of a trajectory spanning 1.5 µs. It shows two sharp peaks corresponding to the choles-



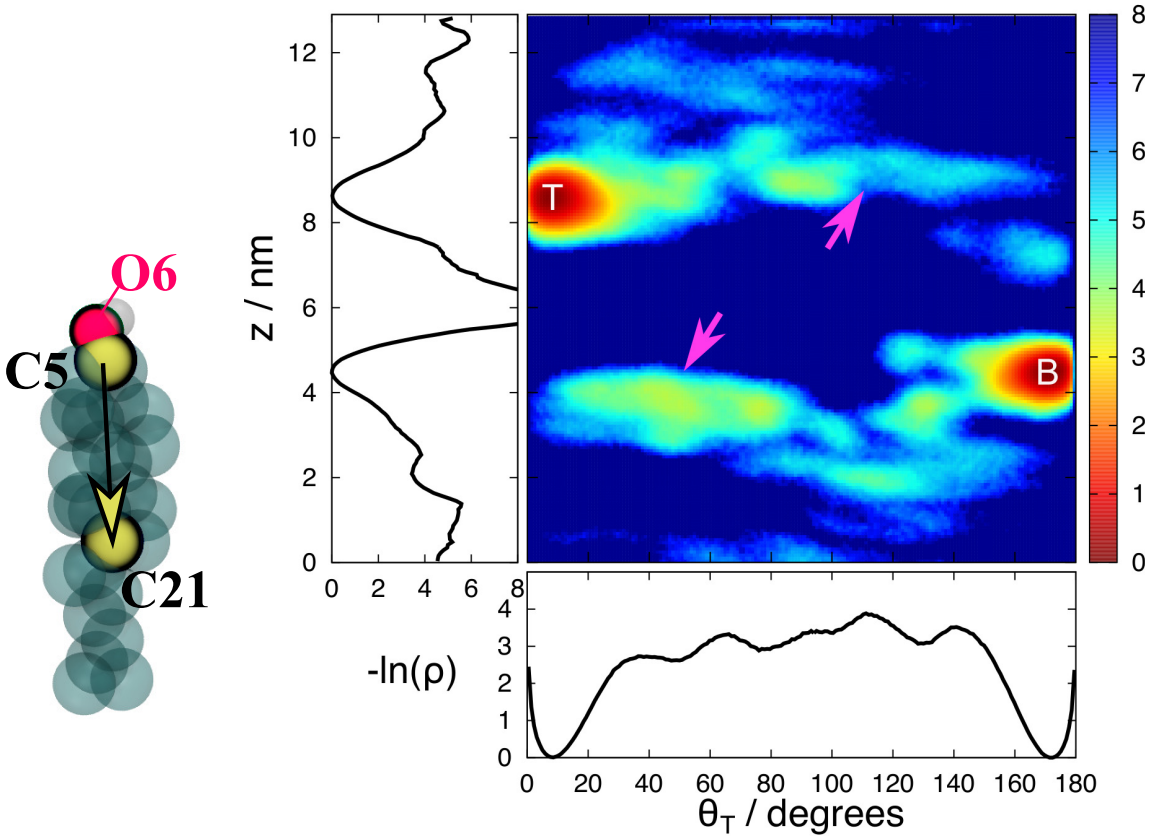
(a) CHO-O6 distribution **(b)** CHO and TG clustering

Figure 4.10: (a) Distribution of the cholesterol head groups (atom O6) along z, $Z\,(i)$, on the left-hand side axis in green and free energy in k_BT units on the right-hand side axis in black (as indicated by the green and black arrows). The arrows indicate the shallower gradient in the CHO distribution close to the interfaces on the TG side. (b) Representative snapshots from the simulation where the TG head groups clusters are highlighted in different colours (blue, light-blue, orange and grey). The CHO molecules inside and outside the bilayer, are also shown in green with most of the extracted species absorbed at the TG-SC interface.

terol remaining in the SC bilayer. The position of these peaks shows that cholesterol molecules preferentially lie below the TG-SC interface (highlighted in the plot with a blue dashed line), as found in previous works[45]. In addition, most of the molecules that are extracted remain in the proximity of the TG-SC interface, resulting in an broad minimum in the free energy profile with a shallower gradient on the TG side, as indicated with red arrows in Figure 4.10a. These molecules were found to preferentially leave the bilayer through a flip-flop event (as shown in the green box of Figure 4.7), with the CHO molecules reverting their orientation before adsorbing at the SC bilayer interface, where they assume a 90° angle with the bilayer normal. Specifically, CHO first lies flat at the SC-sebum interface and from there can either insert into the sebum or back into the bilayer. In Figure 4.10b, a representative snapshot from the simulations is shown where the extracted CHO molecules (in green) are mostly found at the TG-SC interface. Further from the interface, it is possible to measure the free energy barrier to CHO extraction, which measures $\approx 4.5k_BT$.

To further characterise the mechanism of extraction, the 2D free energy landscape for CHO

molecules has been built, using $\Delta G(z, \theta_T) = -k_B T \ln(\rho(z, \theta_T)/\rho_{0,0})$ where $\rho(z, \theta_T)$ is the number of lipid head groups with z coordinate normal to the bilayer surface and tilt angle θ_T with respect to the axis normal to the bilayer plane, and $\rho_{0,0}$ the reference state (*i.e.* the minimum of the 2D profile). The profile is shown in Figure 4.11b. θ_T is defined as the angle between the vector $\overrightarrow{C_5C_{21}}$ and the normal to the bilayer, while z is the position of the head group oxygen, O6, along z (see Figure 4.11a for labels).



(a) CHO label **(b)** CHO 2D free energy

Figure 4.11: (a) CHO molecule shown with translucent beads with labels and coloured beads for the atom O6, C5 and C21. (b) 2D Free energy landscape for CHO extraction as a function of tilt angle and head group z coordinate. The 1D plots are also reported on the x and y axes. The colour bar represents the free energy in k_BT. Letters T and B correspond to the top and bottom leaflets in the bilayer. The magenta arrows indicate CHO flip-flopping to the TG phase, but maintaining contact with the bilayer.

From the plot, it is possible to clearly observe the two free energy minima which correspond to the cholesterol placed in the top and bottom leaflets of the SC bilayer, where θ_T measures $\approx 8°$ and 172°, respectively (labelled with the letter T and B in the plot). Similar values of CHO tilt angle have been previously reported in SC lipid mixtures[45]. Interestingly, a significant number of cholesterol molecules are also present at the interface but with a different range of angles

($\geq 90°$ for the top leaflet and $\leq 90°$ for the bottom one - see magenta arrows in Figure 4.11b), which reflect the tilted cholesterols outside the membrane, but with O6 still in contact with the interface. The two other "clouds" of lower free energy represent the percentage of CHO melted in the TG phase and no longer in contact with the bilayer.

FFA extraction: a lipid escape

FFA molecules also leave the bilayer in similar percentage to CHO ($\approx 15\%$ of the FFA molecules are extracted - see Table 4.1 for data expressed in mol fraction). However, in this case the extraction process appears to differ from the one described above. Most of the FFA are extracted directly without undergoing a flip-flop like event, directly diffusing from the bilayer to the TG phase, as shown in the red box of Figure 4.7.

The configuration assumed by some of the FFA in the bilayer is the first driving factor for the lipid escape. Similarly to CR2, due to the long acyl chain (counting 24 carbon atoms), the FFA aliphatic tail bends near the carboxylic acid group at the bilayer-TG interface (as shown in Figure 4.12). The bending allows a closer interaction between the OH group in FFA and the glycerol groups of TG. It has been previously discussed (in chapter 3 and ref. 17) how TG molecules in the liquid phase cluster via the glycerol group, resulting in a percolating network. At the SC bilayer surface the formation of this network is partially inhibited due to the direct interaction between SC and TG lipids. Figure 4.13b shows that the sites where the TG network meet the bilayer serve as pathway for FFA escape. Once extracted to the sebum phase, FFA OH groups interact with the GLY-clusters and the acyl chains, dragged behind, are placed in the space occupied by the TG aliphatic tails.

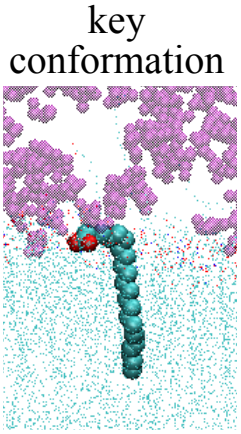
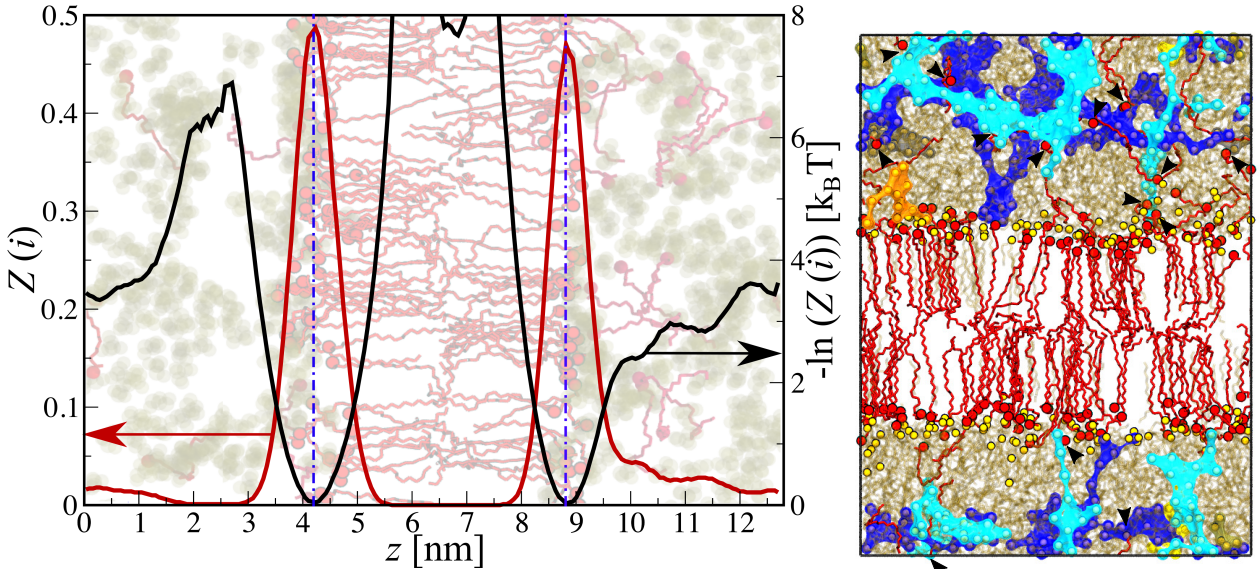


Figure 4.12: FFA with the chain tilted in the head-group region. The purple beads are the GLY of TG molecules.

Figure 4.13a shows the distribution of FFA O1 atoms along z and the corresponding free energy profile. From the density profiles it appears that the FFA extraction can occur via an activation barrier or not. On the left hand side of the bilayer there is an activation barrier of $\sim 7k_BT$ units,

while the free energy difference for the FFA to adsorb in the bilayer is $\sim 3k_BT$ units. In the region of the graph where there is an absorption barrier and consequently a smaller number of events, there is also a gap in the glycerol group network (see Figure 4.13b), which could explain the asymmetry in the free energy distribution. This is probably due to the time required to TG molecules to rearrange themselves and assume the percolating network conformation.



(a) FFA-O1 distribution **(b)** FFA and TG clustering

Figure 4.13: (a) Distribution of the free fatty acid head groups (atom O1) along z, $Z\,(i)$, on the left-hand side axis in red and free energy in units of k_BT on the right-hand side axis in black (as indicated by the red and black arrows). A representative system configuration, at $t = 2.2\,\mu s$, is overloaded to the plot showing the FFA molecules and the glycerol group distribution in the TG phase. (b) Representative snapshots from the simulation where the TG head groups (GLY) clusters are highlighted in different colours (blue, light-blue, orange and grey). The plot shows the position of FFA within the system. Most of the escaped FFA head groups are found to interact with the GLY network, with the tails dispersed in the volume occupied by the TG aliphatic chains. The lipid head groups are represented with a bead and the tails as sticks of different colours: red for FFA and yellow for TG. The other SC lipids are omitted for clarity.

The main difference in the 2D free energy profile for CHO (Figure 4.11b) and FFA (Figure 4.14b) is that FFA molecules outside the bilayer do not show any clear connection between the head group position along the bilayer, z_{O1} and the tilt angle θ_T, calculated between the vector $\overrightarrow{C_1C_{24}}$ and the normal to the bilayer (see Figure 4.14a for labels). In this case, the profile features a homogeneous "cloud", which highlights that once out of the bilayer FFA are free to move in the TG phase, without a preferential direction for the acyl chain. The regions corresponding to FFA within the two leaflets are here also well defined and indicated in Figure 4.14b with the

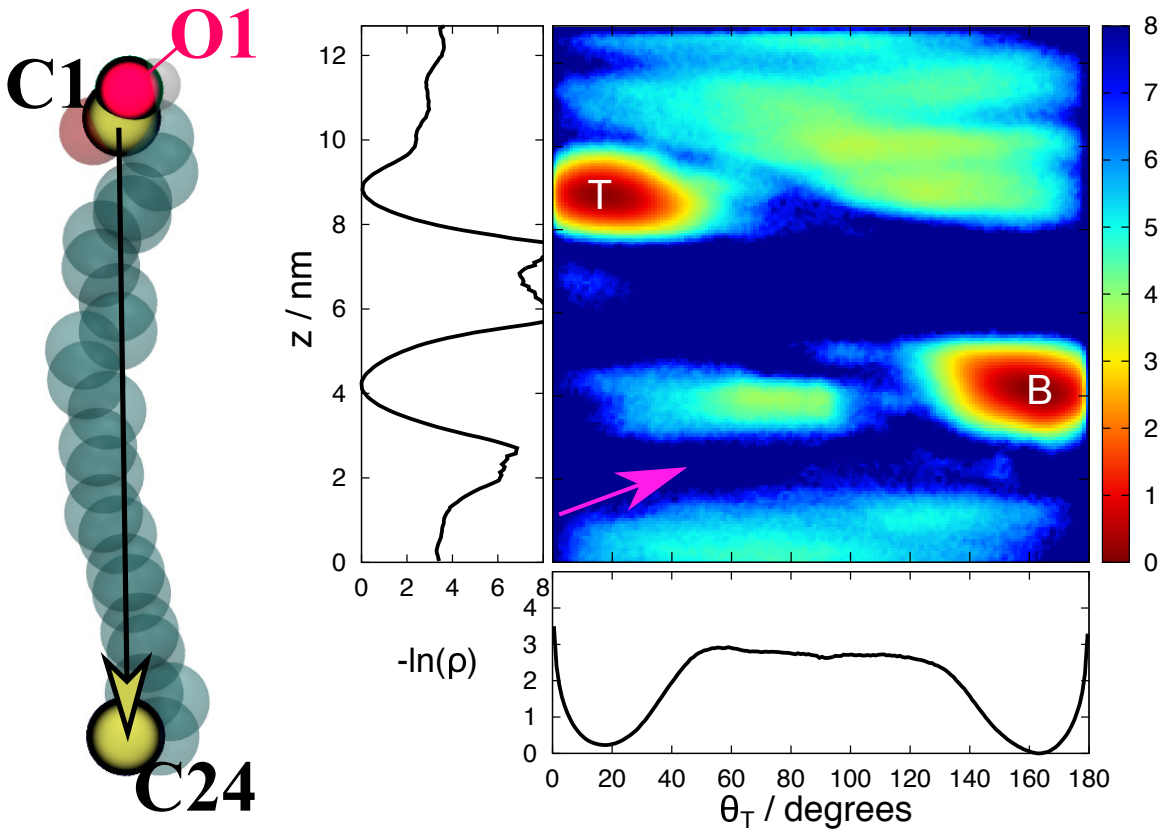


Figure 4.14: (a) FFA molecule shown with translucent beads with labels and coloured beads for the atom O1, C1 and C24. (b) 2D Free energy profile for FFA extraction as a function of the tilt angle and the z_{O1} coordinate. The colour bar represents the free energy in k_BT. The letters T and B indicate the position of the top and bottom leaflets.

letter T and B. In this case, the free energy distribution *vs.* θ_T, presents two wider minima due to the higher flexibility of the long chain fatty acid and the existence of conformations where the FFA bend part of the chain. The minima are located at $\theta_T \approx 15°$ in the top leaflet and $\theta_T \approx 165°$ in the bottom one. In the region below the lower leaflet ($z < 4\,\text{nm}$) one can observe the activation barrier discussed before (see magenta arrow in Figure 4.14b).

4.3.2 Bilayer Stability

The impact of the sebum oil on the SC bilayer was examined by computing a number of properties of the lipid bilayer.

The density profiles of TG (ρ_{TG}) and SC (ρ_{SC}) are shown in Figure 4.15. The data presented in the figure are obtained by averaging the last 200 ns of the trajectory, sampling every 5 ns.

The standard deviations of these averages are plotted as error bars. The density profile of the SC-bilayer equilibrated in water is reproduced in the figure with a dashed line for comparison. The density profile reflects the interdigitation mechanisms defined before, with lower peaks and a more gradual increase in ρ_{SC} at the interface. In the interfacial region, ρ_{TG} smoothly decreases due to the TG reorganisation and the penetration in the bilayer.

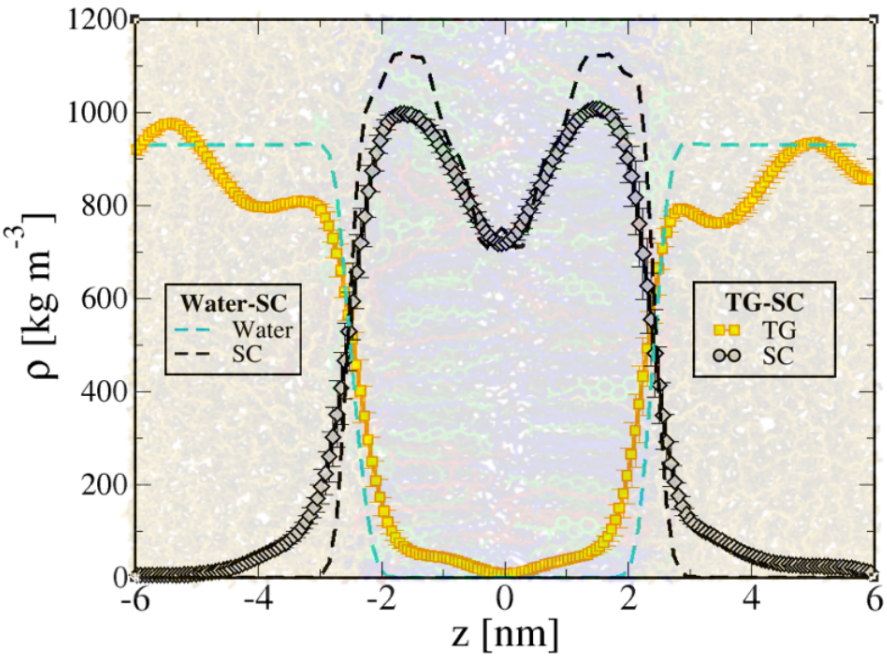


Figure 4.15: Density profile of TG, represented with yellow squares, and SC lipids, represented with dark circles. The density profile of a fully hydrated SC bilayer with the same composition is also added to the figure as a reference with dashed lines (black for SC and light blue for water). On the background a representative snapshot of the system.

A key question in understanding the interaction between sebum TG and the SC bilayer is whether the structure of the main components of the sebum oil is able to destroy the integrity of the SC bilayer. The effect of monoglycerides and fatty acids on the integrity of pure CR2 lipid bilayers in water has been the subject of a recent publication by Akinshina *et al.*[25]. The authors found the degree of unsaturation of the acyl chain to play a crucial role in determining the bilayer integrity, ranging from complete stability against *trans*-unsaturated oils to loss of structure in presence of a single *cis*-double bond oil with an oil content within the bilayer of about 50 - 65%.

In chapter 3, it has been discussed that the composition of the sebum includes lipids with

uncommon patterns of unsaturation and a wide range of acyl chains. Specifically, sapienic acid, the most abundant free fatty acid present in sebum [46] is characterised by a single double bond at the 6th position from the carboxyl end, which might have an impact on the SC lipid bilayer.

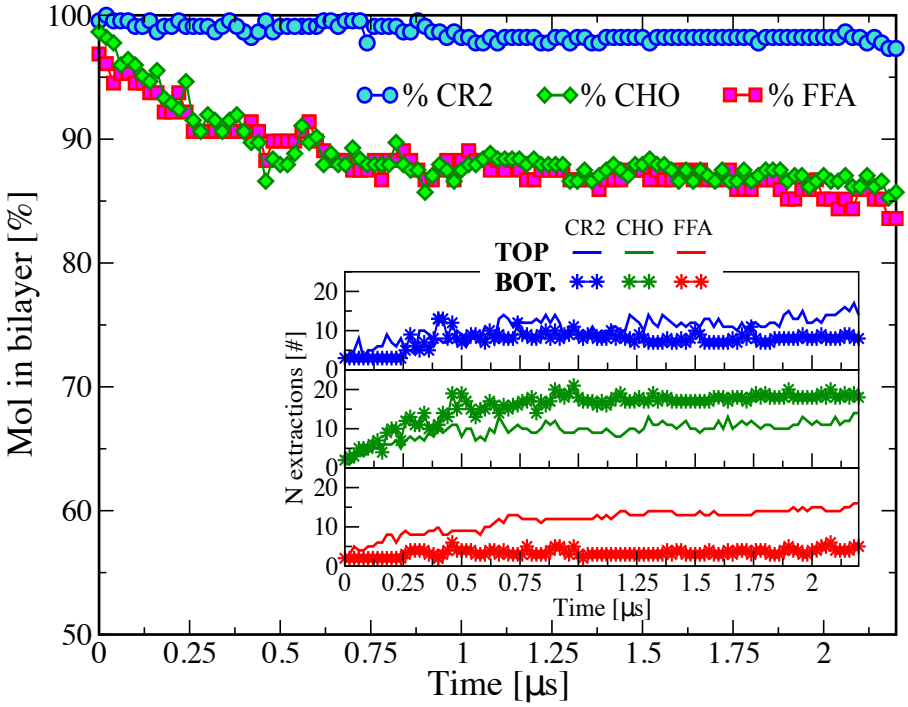


Figure 4.16: Main) Percentage of lipid remaining in the bilayer *vs.* time, blue circles refer to CR2 molecules, green diamonds to CHO and red square to FFA. Inset) Number of extraction event *vs.* time for the three lipid species. From top to bottom: CR2, CHO and FFA. The N of events in the top leaflets are represented with a straight line and in the bottom with asterisks.

To estimate whether the sebum triglycerides would progressively cause a loss of structure in the SC bilayer, the percentage of lipids remaining in the bilayer, $\%N_{SC}$, as a function of time was evaluated. The number of lipids in the bilayer was obtained through the method described in section 4.2.2. $\%N_{SC}$ is reported in Figure 4.16 for each lipid species. As mentioned before, even after 2 µs, the total number of extracted CR2 is just $\approx 2\%$, and $\approx 15\%$ for CHO and FFA, with most of the lipid molecules therefore remaining in the bilayer. In addition, after a first faster reduction for ≈ 1 µs, the lipid percentage levels off towards a constant value (similarly to what was observed for TG insertion, see Figure 4.8a).

In the inset of Figure 4.16, the number of extractions *vs.* time is tracked, distinguishing the number of events in the top (straight line) and bottom (asterisks) leaflet. The number of CR2, CHO and FFA events is slightly asymmetric in the two leaflets, with the number of CR2 splits

and FFA escapes being larger in the top leaflet than the bottom one and *vice-versa* for the CHO extractions. This could be due to some initial asymmetry in the distribution of the lipids in the top and bottom leaflets. The network of TG head groups and the insertion of the TG chains in turn could also locally induce modification in the lipid conformation and drive new extraction events.

The evolution of the area per lipid, A_L, *vs.* time can also be employed as a measure of the bilayer stability. A_L of the different lipids was derived via the Voronoi tessellation method (see section 4.2.2) and shown in Figure 4.17 and Table 4.2.

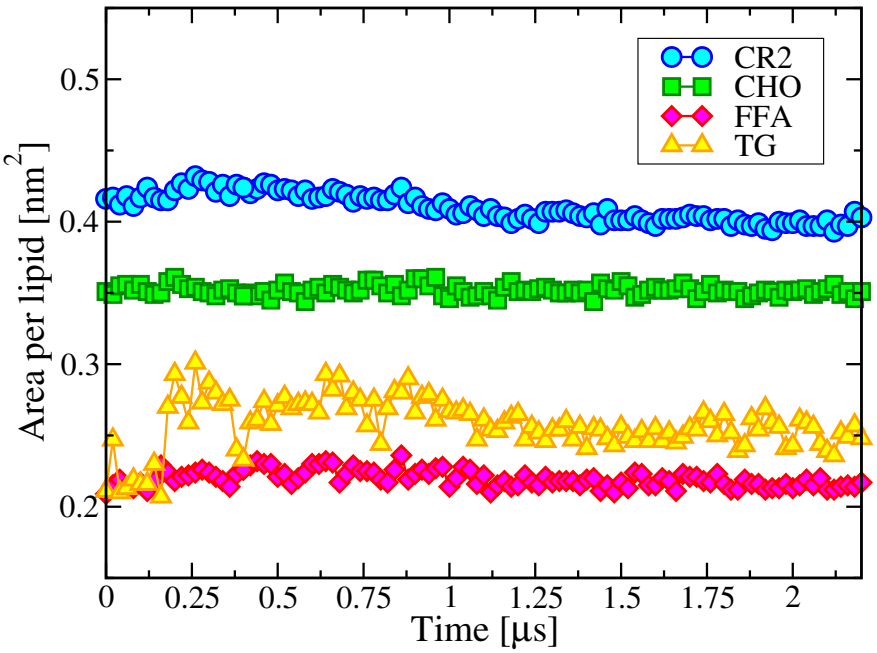


Figure 4.17: Area per lipid *vs.* time for SC and sebaceus lipids, the different lipid contributions are highlighted with different colours and symbols, as specified in the legend. Error bars were not included in the plot for clarity, see Table 4.2 for uncertainty in the measurements.

Lipid	A_L [nm^2]	Err [nm^2]
CR2	0.40356	0.07414
CHO	0.35226	0.06544
FFA	0.21793	0.04405
TG	0.25988	0.10843

Table 4.2: Mean value and standard deviation of A_L measurements. The data are averaged over the simulation length. The error represent the standard deviation of the measurement via Voronoi tessellation method (see section 4.2.2).

All the values of area per lipid, reported in Figure 4.17, stay constant over the period of the simulation and their value is very close to the one measured for SC lipids in water. This indicates that insertion of TG compensates for the extraction events and suggests that CR2, CHO and FFA maintain their favoured area per lipid, without big alterations of the lipid packing caused by the interdigitation mechanisms, supporting the hypothesis of bilayer stability. Remarkably, the value obtained for the area per lipid of CR2 ($\approx 0.4\,\mathrm{nm}^2$) is in agreement with the experimental value obtained by X-ray experiments[47], confirming the accuracy of the proposed forcefield in reproducing experimental data. The small variations of this value in the first 1 µs are caused

	CR2	**CHO**	**FFA**
	$\mathbf{N_{HB(CR2)} = 3.13}$	$\mathbf{N_{HB(CHO)} = 1.45}$	$\mathbf{N_{HB(FFA)} = 1.31}$
CR2	40.6%	51.7%	35.1%
CHO	24.0%	0.0%	19.1%
FFA	8.9%	15.2%	4.6%
TG	26.5%	33.1%	41.2%

Table 4.3: Number of hydrogen bonds, $N_{HB,(i)}$, for SC lipid type: CR2, CHO and FFA. The proposed values are normalised with the total number of lipid of type i in the bilayer. The percentages show how the total number of hydrogen bonds is distributed among the different contributions.

by the CR2 splitting events happening during the equilibration stage.

The TG area per lipid is very similar to the FFA area per lipid. This is due to the fact that TG molecules preferentially penetrate the bilayer with a single chain. The slightly higher value is due to the presence of a double bond in the TG chains and to the existence of a few TG molecule penetrating with 2 or 3 chains.

Hydrogen bonding and the acyl chain order parameter are another two key properties for characterising the structural features of SC bilayers interacting with water, oils or small molecules and address the bilayer stability[4, 5, 25, 26, 27]. CR2 molecules in single component bilayers form a strong network of hydrogen bonds with other lipids and water, which leads to the formation of hexagonal packing[5], making the bilayer rigid and difficult to permeate. The CR2 molecules can in fact be hydrogen bond acceptors and donors at the same time at three different sites. As previously shown by simulations[28, 45] and experiments[48], the addition of CHO and FFA to the bilayer, lipids with just 1 acceptor and 1 donor site each, results in more flexible systems, where long-range hexagonal arrangement is not observed. The effect of TG has been investigated by calculating the hydrogen bonding for the lipid species in the system via the GROMACS tool `g_hbond` with default parameters for the definition of hydrogen bonds, *i.e.* a hydrogen bond is counted for distances between the donor and acceptor atom lower than 0.35 nm and acceptor-donor-hydrogen angles within 30°.

Results are presented in Figure 4.18 and Table 4.3. The graph in Figure 4.18 shows the number of hydrogen bonds (normalised by the total number of lipids in the bilayer, $N_{SC} = 576$) *vs.* time, highlighting the intra-bilayer (SC-SC) and extra-bilayer (SC-TG) contributions. Interestingly, the overall impact of TG on the SC lipid bilayer is a slight increase in the total number of

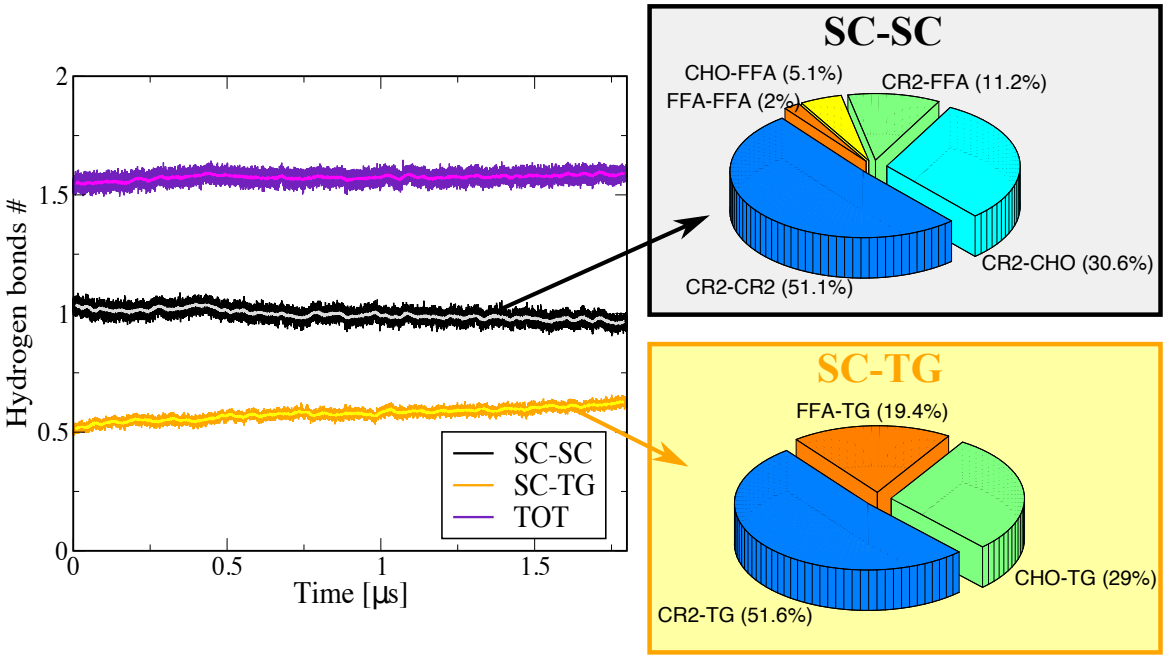


Figure 4.18: Left) Number of hydrogen bonding per lipid *vs.* time. The intra-bilayer (SC-SC) and extra-bilayer (SC-TG) contributions are represented separately. Right) Lipid contribution (expressed as %) to SC-SC h-bond (top square) and SC-TG h-bond (bottom square).

hydrogen bonds (purple curve). The rate of reduction in the number of hydrogen bonds within the SC lipids is in fact lower then the rate of growth of TG-SC hydrogen bonds. This might be related to the insertion of TG molecules into the bilayer, with a consequent reorganisation at the interface allowing closer contact and better alignment between donor groups of SC lipids and TG acceptors. The pie charts on the right of Figure 4.18 show the contributions of the different lipid species to the total hydrogen-bonding, averaged over the last 10 ns of the simulation. Within the SC the main contribution to hydrogen bonding is the CR2 molecules bonding between themselves and other SC lipids ($\approx$ 93% of the total hydrogen bonding involves a CR2 molecule). Looking at the SC-TG hydrogen bonding, one can observe that a significant amount of FFA and CHO molecules are hydrogen bonded with TG molecules, suggesting that in addition to the interfacial hydrogen bonds, some of these interactions involve the extracted molecules, dispersed in the TG phase. Table 4.3 shows the total number of hydrogen bonds per molecules, $N_{HB(i)}$, normalised by the number of lipids of the observed species (in bracket, where i is CR2, CHO and FFA respectively). Interestingly, FFA molecules tend to preferentially form hydrogen bonds with TG acceptors, confirming the interaction between the OH group of FFA and TG glycerol group as a pathway for the escape of FFA.

The acyl chain order parameter, S_z, can be calculated by

$$S_z = \frac{1}{2}\langle 3\cos^2\theta_z - 1\rangle, \tag{4.4}$$

where θ_z is the angle between the normal to the bilayer (here the z axis) and the molecular axis at atom C_n, defined as the vector from atom C_{n-1} to atom C_{n+1}. S_z is equal to 1 when the chain is aligned to z and -0.5 when it is perpendicular. Small S_z indicates isotropic distributions of the tilt angle. Therefore, S_z offers a measure of the ordering of the acyl chain, along with a good opportunity to compare experimental results with NMR experiments (see chapter 2). The order parameters presented below were evaluated with the GROMACS tool `g_order`. Figure 4.19 displays the order parameters for the CR2 sphingosine motif (on the left) and the C24 chains of CR2 and FFA molecules (on the right) as a function of the C_n atom index. The different curves correspond to the order parameter for molecules inside and outside the bilayer (as specified in the legend with IN and OUT). In addition, the order parameter of a SC bilayer equilibrated in water with the same lipid composition was added to the graph with a dashed line, to allow an easy comparison and point out the effect of the sebum on the SC lipid bilayer. The order parameter of the molecules remaining in the bilayer follows the same trend as the bilayer in water, featuring smaller values in the proximity of the head group and the leaflets mid plane regions where the acyl chains are tilted, and a high level of ordering within the leaflet where $S_z \approx 0.8$. Due to the penetration of TG molecules the order of the chains is slightly lower than the case of a bilayer fully hydrated in water. This may be also due to the extraction of FFA from the bilayer, FFA in the SC lipid mixture tends to increase the order of the bilayer straightening the tails of CR2 lipids. The TG tails penetrating the bilayer, instead, present a *cis*-double bond and are expected to decrease the order. This difference is negligible for the C24 chain of CR2 molecules near to the middle of the bilayer (atom index > 14) where the bent C24 tails are in the fluid state. FFA chains present lower values for atom index ≈ 6 due to the existence of conformations where the first part of the FFA chain is tilted at the interface (see Figure 4.12).

Looking at the order parameter of the chains outside the bilayer it can be clearly seen how

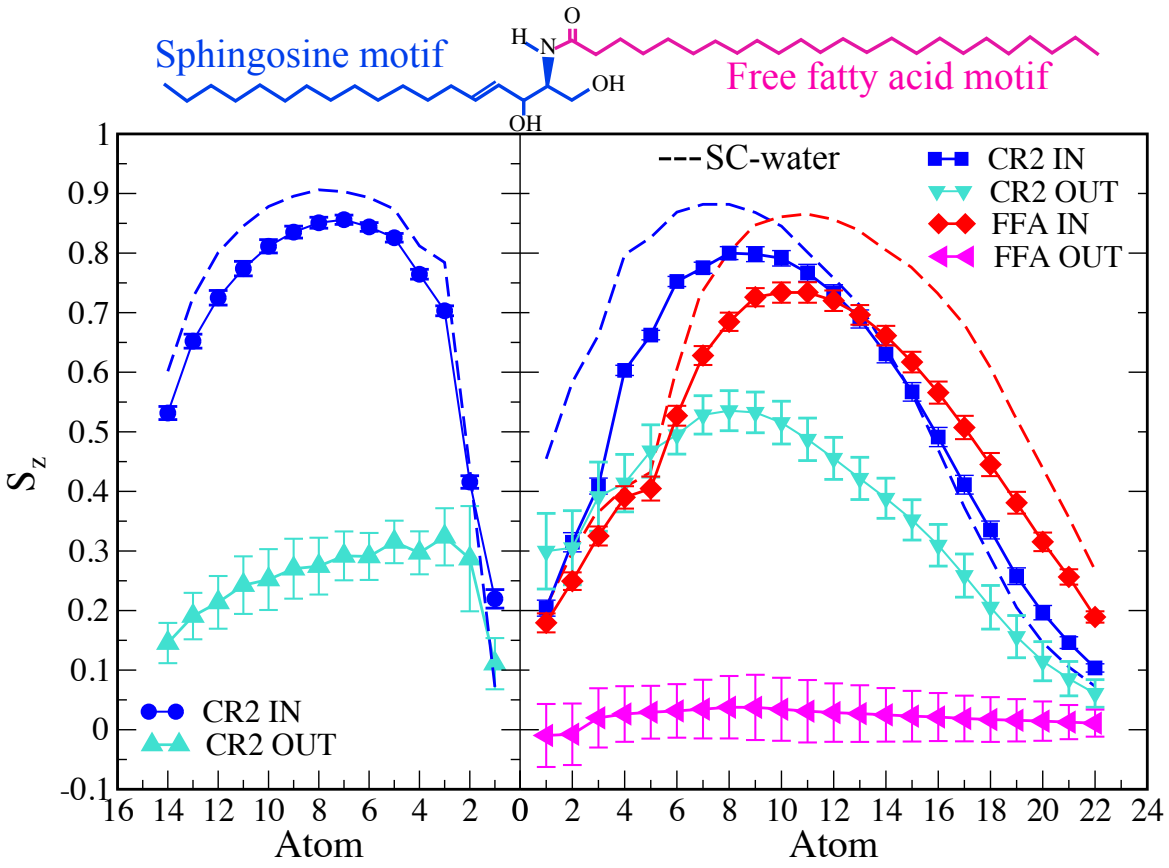


Figure 4.19: Order parameter, S_z, of the lipid tails *vs.* atom index. S_z for the C16 tails of CR2 molecules are reported on the left and S_z for the C24 tails of CR2 and FFA molecules on the right. Different curves correspond to the contribution of molecule inside (IN) and outside (OUT) the bilayer. The dashed curves represent the tail order parameters of a SC lipid bilayer in water.

these tails are much less ordered then the chains within the bilayer. As discussed before, CR2 molecules interdigitate with TG by splitting their chains with one of them immersed in the TG phase. The extracted chains feature S_z values which are typical of lipids in the liquid phase. Extracted C24-CR2 chains (CR2 OUT - right graph of Figure 4.19) feature the same tilt angles as FFA at the head group region (atom index < 6), showing a similar conformation to the FFA molecules, but with the rest of the chain in the TG phase rather than inside the bilayer. On the other hand, extracted FFA are randomly dispersed in TG, losing any contact with the bilayer and showing $S_z \approx 0$, typical of an isotropic distribution.

After 1 µs at the extraction plateau the composition of the bilayer is very similar to the initial one (CR2:CHO:FFA=2:2:1 molar ratio) with an enrichment in TG and a small reduction in the concentration of FFA and CHO. The fractional concentration of TG chains within the bilayer is equivalent to ≈ 0.1 TG chains per SC lipids. Therefore, despite the presence of a *cis*-double

bond, the small concentration of TG chains able to penetrate the bilayer is not enough to cause a loss of integrity in the SC bilayer, in agreement with the prediction of Akinshina *et al.*[25].

The lateral pressure profile, Π, for the TG-SC system was computed with the procedure described in section 2.5.1 and is shown in Figure 4.20. Data over the last 100 ns of simulation run were used, by averaging over 20 000 frames. The cut off radius for the Coulombic interaction is set up to 2.2 nm.

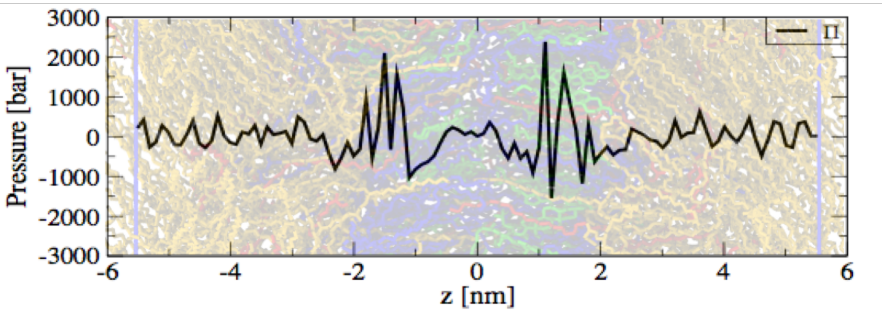


Figure 4.20: Lateral pressure profile for the TG-SC system. The profile is superimposed to a TG-SC system configuration.

The profile presents some small oscillation around 0 in the TG phase due to the network of glycerol groups in the TG phase. At the interface the complex interactions between TG and SC head groups results in three peaks and wells. Π features a local maximum at the centre of the bilayer due to the repulsive interaction between SC tails caused by the leaflets interdigitation. By integrating the lateral pressure profile along z, the average surface tension, γ, was computed and measured $\sim 0\,\mathrm{mN\,m^{-1}}$. The Gaussian modulus, k_G, can be also obtained by direct integration of the second momentum of the lateral pressure profile, as

$$\int_0^{L_z} \Pi\,(z)\, z^2 dz = k_G, \tag{4.5}$$

where L_z is the length of the simulation box along z. It was obtained a value of $k_G = 83/k_BT$, which is comparable with the values reported by Das and Olmsted[49] for SC lipid mixtures in water. A large positive value of k_G supports the stability of the SC lamellar phase.

4.4 Conclusions

The interactions between sebum TG and the main components of the SC lipid matrix, CR2, CHO and FFA, were investigated here, *via* µs MD simulations, to understand how the sebaceus lipid mixture affects the properties of the skin surface.

TG penetrates the bilayer with some of its acyl chains and leads to the extraction of lipids from the SC bilayers. The inserted TG molecules preferentially feature a tuning fork shape with the middle chain inside the bilayer and the other two outside. The presence of a *cis*-double bond in the TG chains and the interactions between glycerol groups and SC lipids head groups make the bilayer less ordered and drive the extraction of SC lipids. The mechanisms of extraction of SC lipids were here characterised, highlighting the tendency for CR2 molecules to split their chains across the TG-SC interface, for CHO to flip-flop in and out the bilayer and for FFA to escape from the SC bilayer to the oil phase. The free energy barrier for CHO and FFA extraction were computed by constructing a 2D free energy landscape to explain these processes in terms of the position of the lipids' head groups and molecules' tilt angle as reaction coordinates.

Overall, despite initial SC lipids extraction and TG chains insertion with a constant rate for $\approx 1\,\mu s$, the number of events levels off to a constant value. In addition, the bilayer is stable over time, with most of the lipids remaining in the SC bilayer. This result has been confirmed by computing a number of properties of the SC bilayer, such as area per lipid, hydrogen bonding, tail order parameter and lateral pressure profile. The findings of this chapter explain the small concentration of skin lipids present in the lipid film covering the surface of our skin and describe the active role of sebum triglycerides on the outermost layer of the skin.

4.5 CPP5 Outreach

The contents and main results of this chapter were summarised in one page and submitted as an entry for the CCP5 outreach library for schools, both for pupils and teachers, and the general public. More information are available online, at the CCP5 website[1, 2].

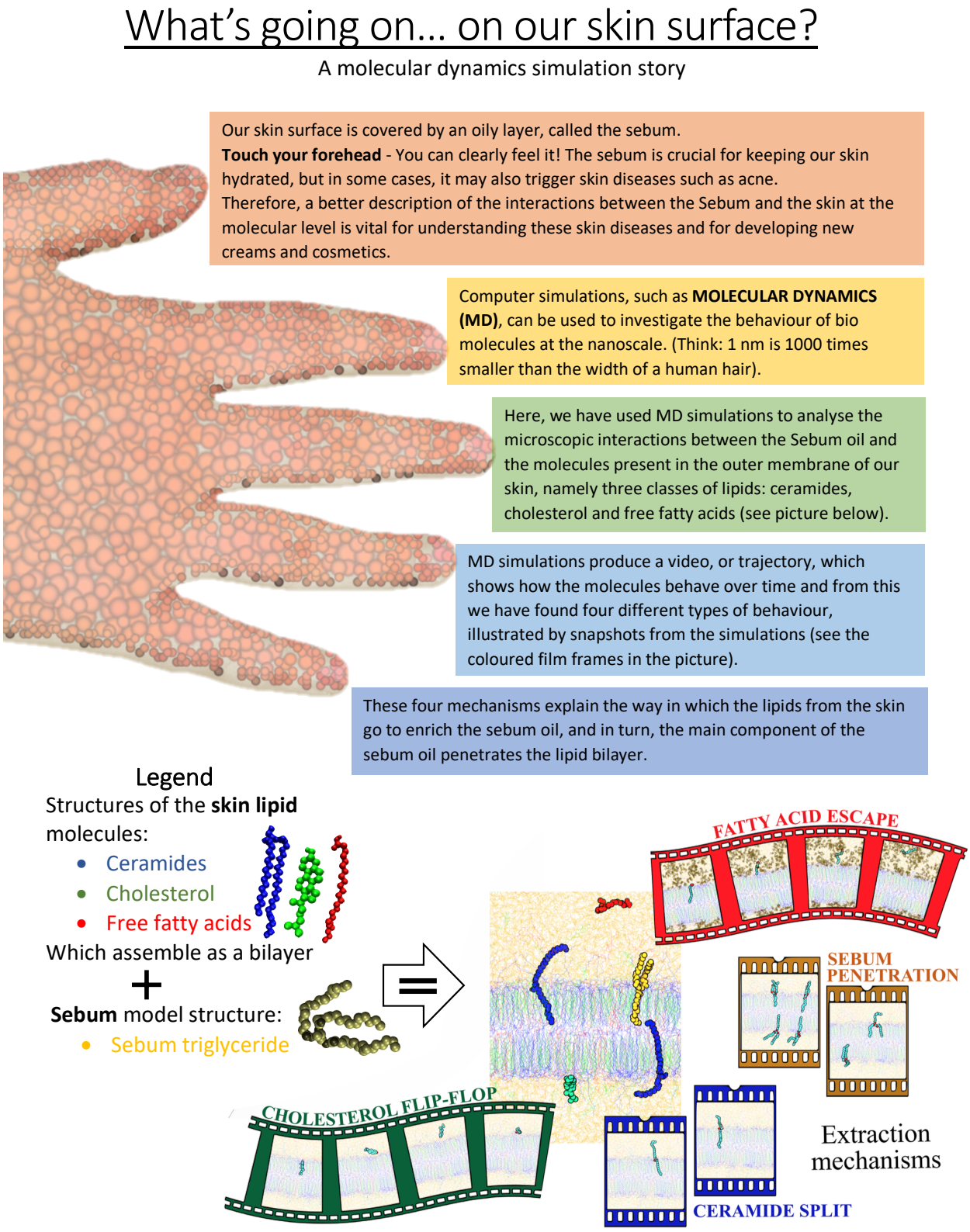


Figure 4.21: Text and image entry for CCP5 Outreach library[1] and contest[2], with which I won the first prize ex–aequo.

Bibliography to chapter 4

[1] CCP5, *CCP5 Outreach library*, `https://www.ccp5.ac.uk/outreach`.

[2] CCP5, *CCP5 Outreach prize*, `https://www.ccp5.ac.uk/ccp5prize`.

[3] L. Norlén, *International journal of cosmetic science*, 2006, **28**, 397–425.

[4] C. Das, P. D. Olmsted and M. G. Noro, *Soft Matter*, 2009, **5**, 4549–4555.

[5] R. Notman, W. K. den Otter, M. G. Noro, W. J. Briels and J. Anwar, *Biophysical Journal*, 2007, **93**, 2056–2068.

[6] J. Lademann, H. Richter, S. Schanzer, F. Knorr, M. Meinke, W. Sterry and A. Patzelt, *European Journal of Pharmaceutics and Biopharmaceutics*, 2011, **77**, 465–468.

[7] M. Schneider, F. Stracke, S. Hansen and U. F. Schaefer, *Dermato-Endocrinology*, 2009, **1**, 197–206.

[8] P. Kattou, G. Lian, S. Glavin, I. Sorrell and T. Chen, *Pharmaceutical Research*, 2017, **34**, 2036–2048.

[9] A. Patzelt and J. Lademann, *Expert Opinion on Drug Delivery*, 2013, **10**, 787–797.

[10] M. Morgen, G. W. Lu, D. Du, R. Stehle, F. Lembke, J. Cervantes, S. Ciotti, R. Haskell, D. Smithey, K. Haley and C. Fan, *International Journal of Pharmaceutics*, 2011, **416**, 314–322.

[11] S. Jung, N. Otberg, G. Thiede, H. Richter, W. Sterry, S. Panzner and J. Lademann, *Journal of Investigative Dermatology*, 2006, **126**, 1728–1732.

[12] R. S. Greene, D. T. Downing, P. E. Pochi and J. S. Strauss, *Journal of Investigative Dermatology*, 1970, **54**, 240–247.

[13] H. M. Sheu, S. C. Chao, T. W. Wong, J. Y. Y. Lee and J. C. Tsai, *British Journal of Dermatology*, 1999, **140**, 385–391.

[14] D. T. Downing and J. S. Strauss, *Journal of Investigative Dermatology*, 1974, **62**, 228–244.

[15] K. R. Smith and D. M. Thiboutot, *The Journal of Lipid Research*, 2007, **49**, 271–281.

[16] C. C. Zouboulis, M. Picardo, Q. Ju, I. Kurokawa, D. Törőcsik, T. Bíró and M. R. Schneider, *Reviews in Endocrine and Metabolic Disorders*, 2016, **17**, 319–334.

[17] A. S. Tascini, M. G. Noro, R. Chen, J. M. Seddon and F. Bresme, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1848–1860.

[18] M. Lovászi, A. Szegedi, C. C. Zouboulis and D. Törőcsik, *Dermato-Endocrinology*, 2017, **9**, e1375636.

[19] J. W. Fluhr, M. Mao-Qiang, B. E. Brown, P. W. Wertz, D. Crumrine, J. P. Sundberg, K. R. Feingold and P. M. Elias, *Journal of Investigative Dermatology*, 2003, **120**, 728–737.

[20] J. J. Thiele, S. U. Weber and L. Packer, *Journal of Investigative Dermatology*, 1999, **113**, 1006–1010.

[21] J. W. Carter, A. S. Tascini, J. M. Seddon and F. Bresme, *Computational Tools for Chemical Biology*, The Royal Society of Chemistry, 2018, ch. 2, pp. 39–68.

[22] C. Das, M. G. Noro and P. D. Olmsted, *Biophysical Journal*, 2009, **97**, 1941–1951.

[23] C. Das, M. G. Noro and P. D. Olmsted, *Physical Review Letters*, 2013, **111**, 1–5.

[24] R. Thind, D. W. O'Neill, A. Del Regno and R. Notman, *Chem. Commun.*, 2015, **51**, 5406–5409.

[25] A. Akinshina, C. Das and M. G. Noro, *Phys. Chem. Chem. Phys.*, 2016, **18**, 17446–17460.

[26] M. I. Hoopes, M. G. Noro, M. L. Longo and R. Faller, *Journal of Physical Chemistry B*, 2011, **115**, 3164–3171.

[27] R. Notman and J. Anwar, *Advanced Drug Delivery Reviews*, 2013, **65**, 237–250.

[28] C. Das, M. G. Noro and P. D. Olmsted, *Soft Matter*, 2014, **10**, 7346–7352.

[29] L. Martinez, R. Andrade, E. G. Birgin and J. M. Martínez, *Journal of Computational Chemistry*, 2009, **30**, 2157–2164.

[30] O. Berger, O. Edholm and F. Jähnig, *Biophysical Journal*, 1997, **72**, 2002–2013.

[31] W. L. Jorgensen and J. Tirado-Rives, *Journal of the American Chemical Society*, 1988, **110**, 1657–1666.

[32] F. W. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott and I. G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, 1996, pp. 1–1042.

[33] J. P. Ryckaert and A. Bellemans, *Chemical Physics Letters*, 1975, **30**, 123–125.

[34] D. Peter Tieleman, H. J. Berendsen and M. S. Sansom, *Biophysical Journal*, 1999, **76**, 1757–1769.

[35] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *The*

Journal of Chemical Physics, 1995, **103**, 8577–8593.

[36] B. Hess, H. Bekker, H. J. C. Berendsen and J. G. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.

[37] G. Bussi, D. Donadio and M. Parrinello, *Journal of Chemical Physics*, 2007, **126**, 014101.

[38] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *Journal of Computational Chemistry*, 2005, **26**, 1701–1718.

[39] D. Zwillinger, *CRC Standard Mathematical Tables and Formulae*, 2003, p. 367.

[40] S. Jo, H. Rui, J. B. Lim, J. B. Klauda and W. Im, *Journal of Physical Chemistry B*, 2010, **114**, 13342–13348.

[41] L. Norlén, *Journal of Investigative Dermatology*, 2001, **117**, 830–836.

[42] D. Kessner, A. Ruettinger, M. A. Kiselev, S. Wartewig and R. H. Neubert, *Skin Pharmacology and Physiology*, 2008, **21**, 58–74.

[43] M. A. Kiselev, *Crystallography Reports*, 2007, **52**, 525–528.

[44] T. C. Moore, C. R. Iacovella, R. Hartkamp, A. L. Bunge and C. McCabe, *Journal of Physical Chemistry B*, 2016, **120**, 9944–9958.

[45] A. Del Regno and R. Notman, *Physical Chemistry Chemical Physics*, 2018, **20**, 2162–2174.

[46] A. Pappas, *Dermato-Endocrinology 1:2,*, 2009, **1**, 72–76.

[47] B. Dahlén and I. Pascher, *Chemistry and Physics of Lipids*, 1979, **24**, 119–133.

[48] J. A. Bouwstra and G. S. Gooris, *The Open Dermatology Journal*, 2010, **4**, 10–13.

[49] C. Das and P. D. Olmsted, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2016, **374**, year.

Chapter 5

How wettable is the skin surface?

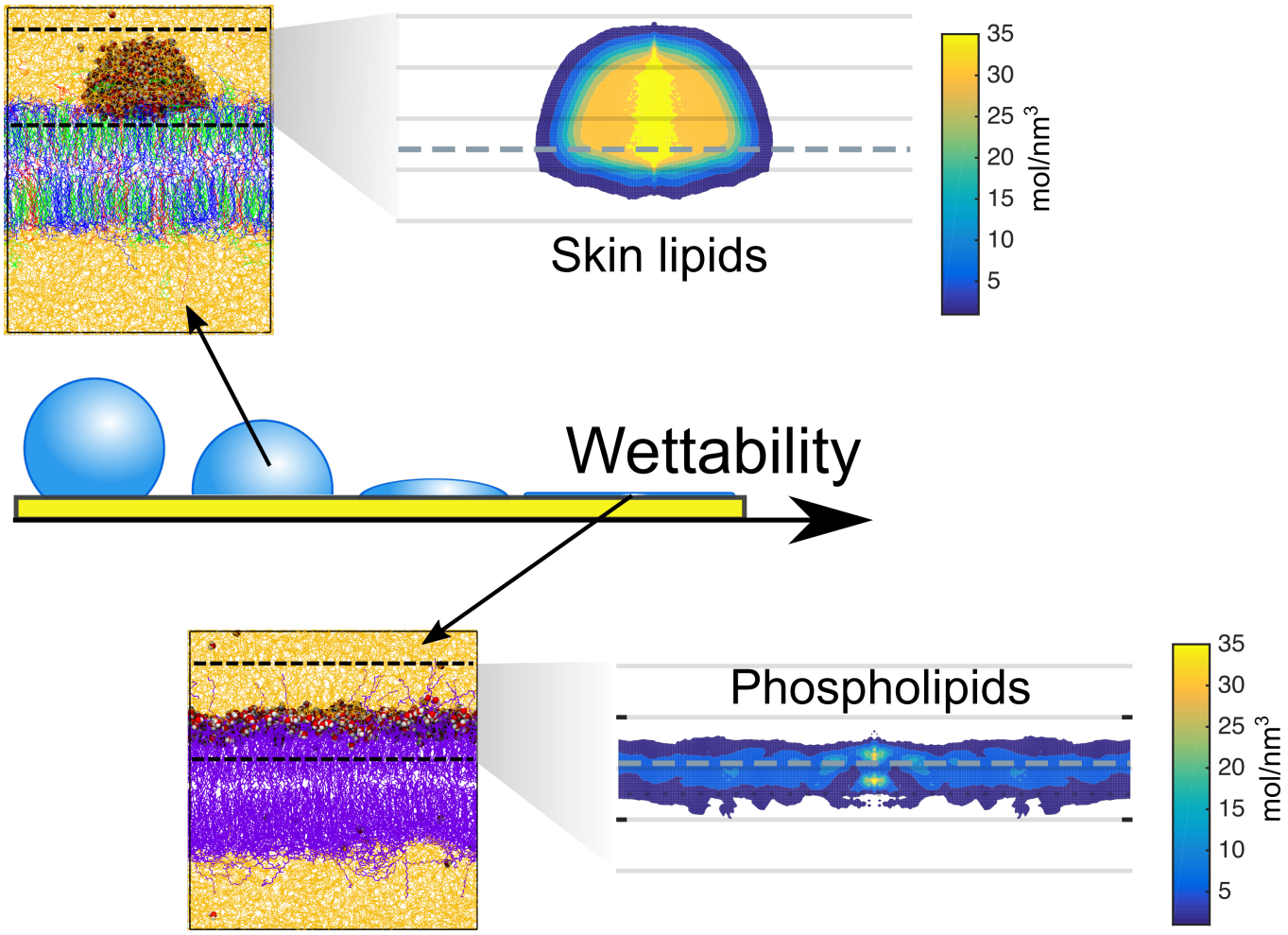


Figure 5.1: Title page figure for chapter 5

5.1 Introduction

The properties of the skin surface are crucial for understanding the spreading of topically applied drugs and cosmetics, where interfacial phenomena play a major role in the process of adsorption through the skin. Specifically, understanding the wetting properties of the skin surface is an important first step in determining the interactions of any agent with the skin surface and also in explaining the functionality of the skin barrier and its role of providing natural protection against infections[1].

The importance of studying the skin wettability to better understand the interactions of creams and cosmetics with the skin has been discussed in several studies[1, 2, 3, 4]. Mavon *et al.*[2] investigated the effect of sebum on the wettability of the skin surface against several solvents with different polarities and concluded that sebum rich areas of the body feature lower contact angles than sebum poor areas (≈55° *vs.* ≈88° for water on forehead (sebum rich) and water on forearm (sebum poor), respectively). The same study also reported a reduction of the wettability after ether treatments, resulting in an increased contact angle (from ≈55° to ≈89° on the forehead and from ≈88° to ≈101° on the forearm before and after the treatment). This study suggests that the sebum gives a hydrophilic character to the skin surface. This is supported by the works of Krawczyk[1] and Mavon *et al.*[4], which report similar contact angles for the skin surface with and without sebum. However, a recent and complete review on skin tribology by Derler and Gerhardt[5] highlights that a clear understanding of the impact of hydration on skin surface properties is still missing, as is a detailed explanation of the influence of the skin surface lipid film, mainly consisting of sebum.

Water is vital for human beings, hence loss of water must be carefully controlled[6]. The regulation function is performed by our skin, the largest organ of the human body, where water is distributed inhomogeneously, with a strong concentration gradient across the epidermis. Specifically, this water gradient is very sharp in the proximity of the outermost layer of the skin, the stratum corneum (SC). Here the amount of water drops dramatically (from 50% to 15% grams of water per total grams of wet tissue[7] - as displayed in Figure 5.2a), resulting in a low water content near the skin surface.

The low affinity between water and the stratum corneum lipids, which causes the low hydration of the SC layer, has been confirmed experimentally by Groen *et al.*[8], who found that an increase of the deuterated water (D_2O) concentration in the hydration buffer does not affect the neutron scattering signal, suggesting that a relatively low number of water molecules are present in the membrane sample. In previous experimental studies on the SC or SC model membranes it has been observed that, even at 100% humidity, the repeat distance of the lipid lamellae is almost insensitive to the level of hydration, suggesting that very little water is present in the head group region, between bilayers[9, 10, 11]. The exact location of the water within the SC layer is still an open question.

Subsection 1.2.3 and section 4.1 present a review of how MD simulations have been previously used to investigate the properties of SC lipids. In all the previous models, the SC lipids are modelled as a fully hydrated bilayer, neglecting the small amount of water actually present in the SC layer and its distribution. A recent publication by Del Regno and Notman[12] represents a first step in reproducing the level of hydration measured experimentally (2 water molecules per lipid[8]). Interestingly, despite the low amount of water between two SC bilayers, they found that the layer of water molecules placed at the interface tends to dewet from the bilayer interplane and accumulate into water pools. The location of water in the SC layer raises questions on the wettability of the SC lipid matrix.

In this chapter molecular dynamics simulations of a water droplet placed at the sebum-skin lipid interface are performed to microscopically investigate the wettability properties of the skin surface. Following on from our previous studies (presented in chapter 3 and chapter 4) the sebum was modelled by tri-cis-6-hexadecenoin, a representative sebum triglyceride (TG), while the SC membrane was assembled as a multicomponent bilayer made up of ceramide 2 (CR2), cholesterol (CHO) and a mayor component free fatty acid (lignoceric acid - FFA). The simulation results presented in this chapter show that the SC bilayer features a low affinity for water, with contact angles $\geq 90°$. The wettability analysis is also performed with a DPPC bilayer, which is fully wettable to water, with contact angles $= 0°$.

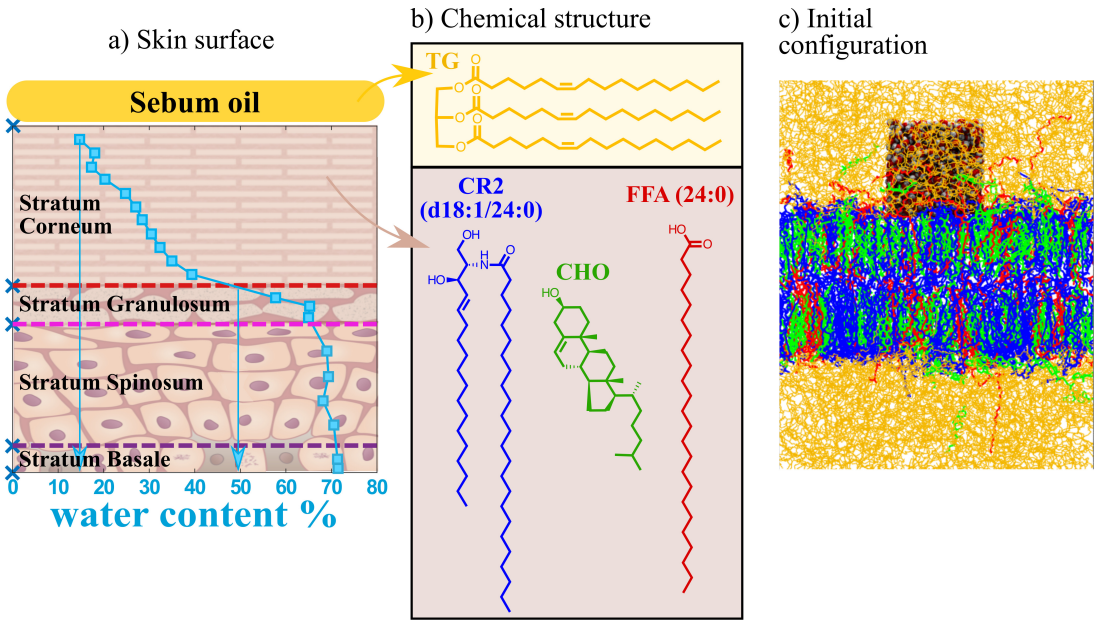


Figure 5.2: (a) The layers of the epidermis and the water content within them (the data for the water content are taken from ref. 7.) (b) Chemical structures of the lipids investigated in this work. The sebum is modelled by its most aboundant component, the triglyceride tri-cis-6-hexadecenoin (TG). The multicomponent model of the SC is made up of ceramide 2 (CR2), cholesterol (CHO) and lignoceric acid (FFA). (c) Initial configuration of the systems used in this analysis: a water cube is placed at the TG-SC interface. The colour code for the lipids is the same as the chemical structures in b. Water oxygen atoms are represented in red and the hydrogens in white.

5.2 Methods

5.2.1 Input configurations and simulations details

Section 4.2.1 presents a description of the set-up parameters employed to simulate an equilibrated TG-SC interface in the previous chapter. The same models of sebum triglycerides and SC lipids (see lipid structures in Figure 5.2b) were used in this chapter to give insight into the interactions between these two lipid mixtures and water and to investigate the wetting behaviour of the skin surface. Following the procedure presented in chapter 3, the water molecules were represented using the SPC model for consistency with the Berger lipid parametrisation and the water description employed in other simulations of epidermal and sebaceus lipids[13, 14, 15, 16].

To create the starting configuration (Figure 5.2c), 800 water molecules were first equilibrated in a cubic box with $L = 2.5\,\text{nm}$ for 10 ns at constant volume and temperature ($T = 310\,\text{K}$), using the velocity rescaling algorithm[17] with time coupling constant, τ, equal to 0.1 ps. Then,

starting from an equilibrated TG-SC interface taken from the end of the production run of the simulations described in chapter 4, the H_2O cube was placed just above the top leaflet of the SC bilayer, within the sebum layer, by removing all the overlapping TG molecules. Figure 5.2c shows the initial configuration with the water cube placed at the TG-SC interface. The evolution of the positions of the water molecules was monitored for 5 µs in the isothermal-isobaric ensemble (NPT). The temperature was kept constant at 310 K with the Nosé-Hoover thermostat, with $\tau = 1$ ps, and the pressure was kept at 1 bar, with the Parinello-Rahaman barostat, applied semi-isotropically with $\tau = 5$ ps and a compressibility of 4.5×10^{-5} bar^{-1}. The time-step of the simulations was set equal to 0.002 ps and periodic boundary conditions were applied in all directions. The intra-molecular bonds were constrained using the P-LINCS algorithm, van der Waals interactions were truncated at 1.2 nm and electrostatic interactions were computed in full using the Particle Mesh Ewald algorithm[18].

To compare the wettability of SC lipids to other phospholipids, abundant in other cell membranes in other parts of the body, the same procedure was repeated with a DPPC bilayer rather than the SC bilayer to generate a configuration containing DPPC, sebum and a water cube. The intra and inter molecular interactions of DPPC lipids were also described with the Berger forcefield[19]. The DPPC bilayer was first equilibrated in water (see procedure described in ref. 20) and then in TG for 500 ns in the NPT ensemble with the same simulation set-up described above for the SC lipids. Then, the H_2O cube was inserted just above the DPPC-TG interface, within the sebum layer by replacing the overlapping TG molecules. The evolution of the positions of the water molecules was monitored for 0.5 µs at constant T and p.

All the simulations were performed with the GROMACS/4.5.5 molecular simulation package[21].

5.2.2 Contact angle computations

Wetting refers to the study of how liquids, deposited on a solid (or liquid) substrate, spread out. A strong affinity between the liquid and the substrate results in full wetting, where the liquid is completely spread on the substrate; while a weaker affinity results in partial wetting. These two wetting regimes can be distinguished by evaluating the spreading parameter, S, which is

defined as:

$$S = [E_{substrate}]_{dry} - [E_{substrate}]_{wet} \tag{5.1}$$

where $E_{substrate}$ is the surface energy (per unit area) of the substrate under investigation[22]. Often the surface energy, defined as the energy needed to create a surface with unit area, is associated with the surface tension, γ, defined as a force acting along a line of unit length parallel to the surface but perpendicular to the line. These two quantities are equivalent for the liquid-liquid and liquid-gas interfaces[23]. S can be rewritten in terms of the surface tensions as $S = \gamma_{12} - (\gamma_{23} + \gamma_{31})$, where 1 often represents air or vapour, 2 the substrate and 3 the liquid (see examples in Figure 5.3). Following from equation 5.1, $S > 0$ corresponds to full wetting, while $S < 0$ corresponds to the partial wetting regime.

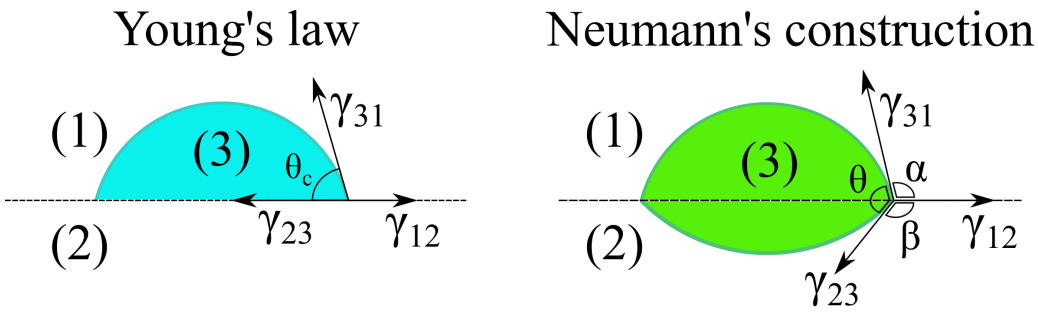


Figure 5.3: Left) Liquid droplet (3) on a rigid substrate (2) in equilibrium with phase (1). The equilibrium shape of the droplet with the contact line is determined by the configurational energy balance, described by Young's law. Right) Liquid droplet (3) at the interface of two immiscible liquids (1) and (2). The shape of the contact line separating the three fluids is determined by the balance of surface tension forces, γ_{ij}, which can be expressed according to Neumann's construction.

When the substrate is a smooth and rigid solid surface the wettability can be quantified from the contact angle of the liquid on the substrate (see Figure 5.3-left for the definition of the surface tensions and the contact angles), according to Young's equation:

$$\gamma_{31} \cos\theta_c = \gamma_{12} - \gamma_{23}. \tag{5.2}$$

Following the approximation of Young's law, S can be defined as a function of θ_c, by substituting equation 5.1 into 5.2, obtaining $S = \gamma_{31}(\cos\theta_c - 1)$. When $\theta_c > 90°$, the liquid (3) is assumed to be "mostly non-wetting" and "mostly wetting" otherwise. $\theta_c = 0°$ marks the transition to "full wetting" condition.

Under the approximation of solid surface, the substrate balances any projection of the capillary forces onto the vertical axis, without overcoming any deformation. The situation could be different for a soft substrate, which can deform due to the Laplace pressure associated to the curvature of the liquid droplet. For soft material Young's approximation needs to be reformulated and the solid surface can have a contribution to surface tension beyond the internal energy. The local deformation will be in this case dictated by a balance between the component of the surface tension of the liquid perpendicular to the surface and the ability of the soft surface to deform[24].

On the other hand, when we have a liquid substrate (2) immiscible with the droplet (3) (as in Figure 5.3-right), the surface will no longer be planar and will adjust itself to minimize the surface energy. In this case, Neumann's construction can be used to calculate the contact angles. Here, both the horizontal and vertical components of the capillary forces must be balanced, and considering the definition of the surface tensions in Figure 5.3-right we can write:

$$\vec{\gamma_{23}} + \vec{\gamma_{12}} + \vec{\gamma_{31}} = 0, \tag{5.3}$$

which leads to:

$$\begin{cases} \gamma_{31} + \gamma_{23}\cos\theta + \gamma_{12}\cos\alpha = 0 \\ \gamma_{31}\cos\theta + \gamma_{23} + \gamma_{12}\cos\beta = 0 \\ \gamma_{31}\cos\alpha + \gamma_{23}\cos\beta + \gamma_{12} = 0. \end{cases} \tag{5.4}$$

The system of equation 5.4 can be simplified by considering standard trigonometric relations in

$$\begin{cases} \gamma_{31}\sin\theta_A = \gamma_{23}\sin\theta_B \\ \gamma_{12} = \gamma_{31}\cos\theta_A + \gamma_{23}\cos\theta_A, \end{cases} \tag{5.5}$$

where $\theta_A = 180° - \alpha$, $\theta_B = 180° - \beta$ and $\theta_A + \theta_B = \theta$.

In this chapter, the shape of the water droplet was quantified by computing the density profile of the water molecules in cylindrical coordinates. The origin was placed at the centre of mass of the water droplet and the profile was computed in terms of the cylindrical radial coordinate,

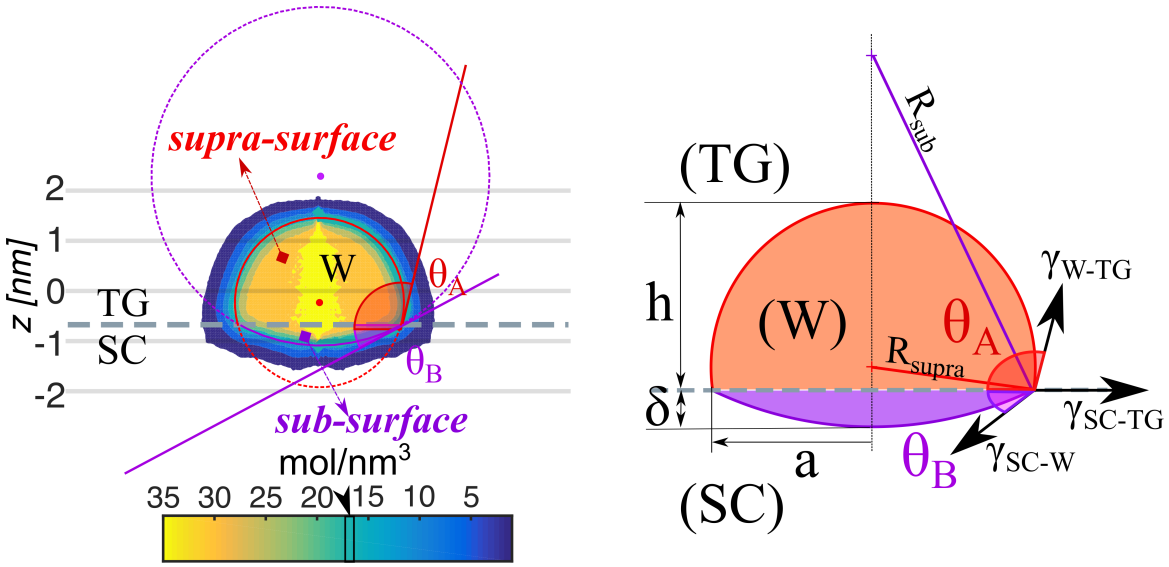


Figure 5.4: Left) Density profile of the water lens in cylindrical coordinates, (r,z). The contour $d = 16.7\,\mathrm{mol\,nm^{-3}}$ (see black arrow on colour bar) is used for the construction of the supra- (in red) and subsurface (in purple). The supra (θ_A) and sub (θ_B) contact angles are obtained from construction of the tangent lines to the two spherical caps at the intersection with the interface (shown as a grey dashed line). Right) Zoom on the two spherical caps defined by the supra (red) and sub (purple) surface. The definition of the droplet immersion, h, the spherical caps radii (R_{sub} and R_{supra}) and the penetration depth, δ, used to geometrically determine θ_A and θ_B, are shown.

r, and the z axis coordinate, by averaging over all azimuthal angles at a given vector (r,z). The density contours were obtained by averaging data taken over 20 ns along the 5 µs trajectory. The link between the contours and the numerical values of the density, measured as water molecules per nm^3, is expressed by the colour bar of Figure 5.4-left. The water droplet adopts the shape of a water lens. Its shape was described by fitting circular arcs to the iso-density line for $d_{\mathrm{H_2O}} = 16.7\,\mathrm{mol\,nm^{-3}}$ which corresponds to 50% of the bulk density of water (see black arrow on the colour bar of Figure 5.4-left). The position of the TG-SC interface was, instead, determined by fitting a straight line to the z coordinates of the SC atoms drawing the iso-density line for SC lipids at $d_{SC} = 20$ pseudo-particles nm^{-3}, where a pseudo-particle is a single united atom particle, which is either a single non-carbon atom or a carbon atom with its bonded non-polar hydrogens. All the SC atoms in the proximity of the water lens, whose $r \in [-2.5, 2.5]$nm, were excluded from the fitting. Considering the position of the interface, the water lens shape was fully characterised by approximating it to two spherical caps, namely a supra-surface (above the interface) and a sub-surface (below the interface) as illustrated in Figure 5.4-right. The *supra-* and *sub*-surface contours were fit to a circumference using a non-

linear square fit[25]. The two angles, θ_A and θ_B, were used to characterise the shape of the spherical caps and discuss the wettability of the interface (see section 5.3.4 below). They were computed by geometrical construction from:

- the spherical caps approximate radii, R_{supra} and R_{sub};
- the immersion of the droplet, h;
- the penetration depth, δ;

as:

$$\theta_A = \arccos\left(1 - \frac{h}{R_{supra}}\right) \quad \text{and} \quad \theta_B = \arccos\left(1 - \frac{\delta}{R_{sub}}\right). \tag{5.6}$$

See Figure 5.4-right for the description of the labels above.

5.3 Results

5.3.1 Characterising the water adsorption

To measure the wettability of the TG-SC interface the behaviour of a water droplet placed at the lipid interface was monitored during a 5 µs MD run. Interestingly, the size and the shape of the water droplet was fairly stable during the simulations. The stability of the water droplet was measured by computing the size of the droplet *vs.* time with the droplet size defined as the volume of the biggest 3D water cluster, v_{H_2O}. The water clusters were identified using the same 3D Voronoi tessellation algorithm as was employed for TG cluster determination presented in chapter 3 (see section 3.2.3) and ref. 16. The 3D Voronoi cells and their associated volumes were first determined with the open-source software voro++[26] and then the clusters were computed by considering the nearest neighbour Voronoi cells of water atoms in the 3D-space. The cluster volumes were obtained by summing up the individual Voronoi cell volumes for the cells in each cluster. The evolution of the water droplet volume, v_{H_2O}, with time is shown in Figure 5.5-top, along with the shape of the biggest water cluster, displayed in the six simulation snapshots

above the plot. v_{H_2O} slightly decreases until 2 µs and then plateaus at approximately 20 nm^3 for the rest of the simulation. The shape of the cluster is also fairly constant and well defined.

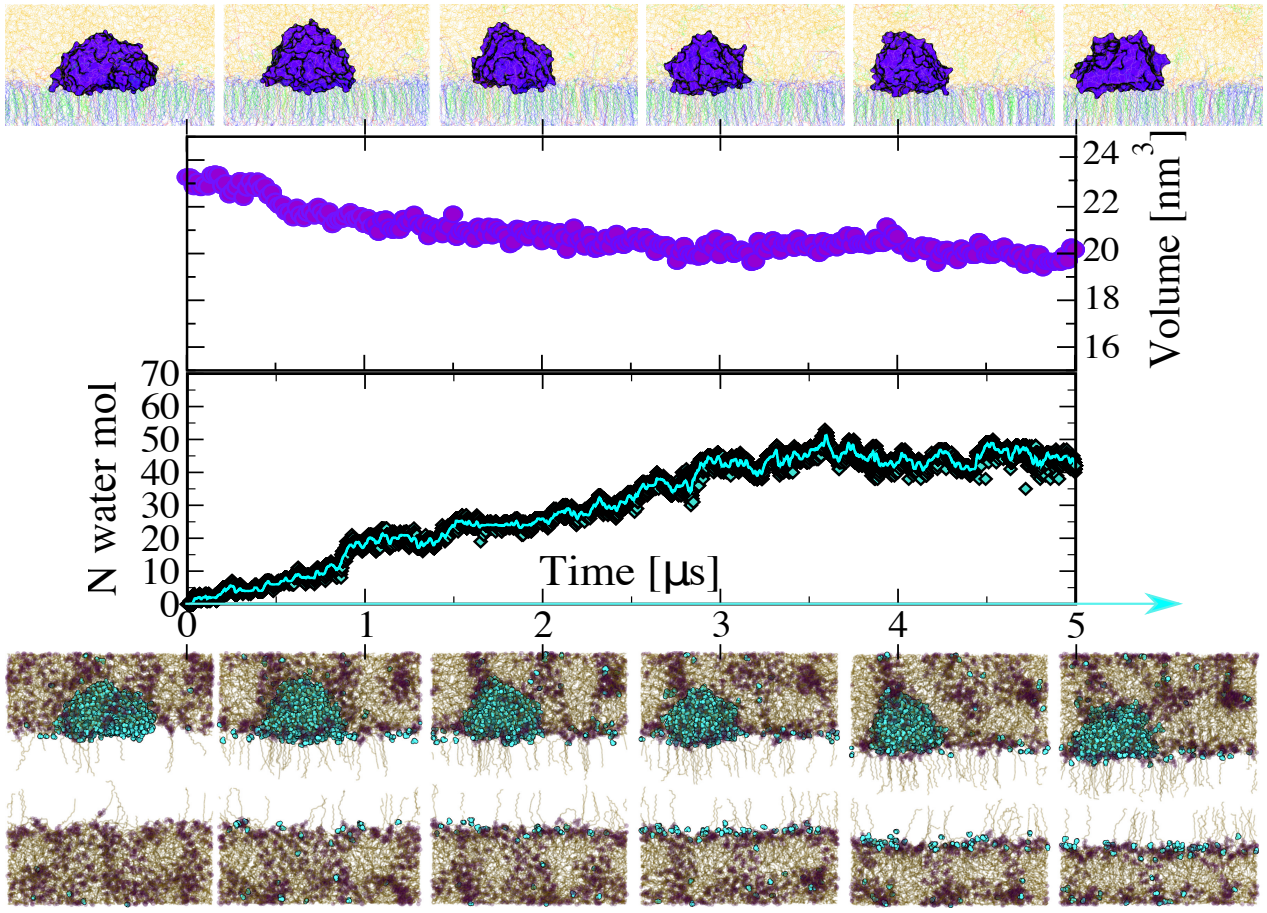


Figure 5.5: Top) Volume of the water droplet, v_{H_2O}, *vs.* time. On the top of the figure six snapshots showing the volume of the largest water cluster at 0 µs, 1 µs, 2 µs, 3 µs, 4 µs and 5 µs. Bottom) Number of water molecules at the lower TG-SC interface *vs.* time. The light-blue line is the running average of the computed results. The snapshots below the graph are six representative system configurations showing the evolution of the water droplet at the TG-SC interface (from left to right: at 0 µs, 1 µs, 2 µs, 3 µs, 4 µs and 5 µs). In the simulation snapshots water molecules are represented in light blue; TG molecules in yellow with the glycerol group beads highlighted in brown; SC lipids are not displayed for clarity.

Whilst most of the water molecules remain in the droplet, a few molecules leave the water droplet and diffuse within the sebum layer and some of these eventually reach the lower TG-SC interface. The position of the water molecules within the system is shown in six representative simulation snapshots in Figure 5.5-bottom. The number of water molecules reaching the lower TG-SC interface *vs.* time is also shown in the plot above these snapshots in Figure 5.5-bottom. After an initial slow and steady increase, the number of water at the lower interface equilibrates at a constant value of $\approx$ 45 water molecules. Considering that there are 256 lipids in each

leaflet, this water concentration corresponds to a value of ≈ 0.2 water molecules per SC lipid. This value of water per lipid concentration is much lower than the experimental one reported by Groen *et al.*[8] for short periodicity phase (SPP) SC model membranes, composed of CR2, CHO and FFA. In ref. 8 the proposed number of water molecules per lipid, evaluated by weighing a sample in the dry and hydrated states, measures 1.91 ± 0.42. However, in previous simulations of SC model membranes[12], starting with a level of hydration of 2 water molecules per lipid, it has been observed that, rather than forming a homogeneously distributed layer at the lipid interface, water forms flattened ellipsoidal pools. This suggests that the physiological number of water molecule per lipid might be lower than the one measured experimentally, which might actually correspond to a situation of over hydration with the water concentrated inhomogeneously in confined spaces between the lipid bilayers.

The main mechanism of water diffusion to reach the lower TG-SC interface is through the sebum phase and the periodic boundary conditions. The energy barrier for water to cross the SC bilayer has been previously computed by Das *et al.*[13], who concluded that the maximum in the excess free energy of water inside SC bilayers is twice that of water in phospholipid bilayers at the same temperature, resulting in a very low permeability for the SC membrane. On the other hand, TG molecules in bulk tend to form ordered structures, with the glycerol groups clustered together forming a percolating network, which facilitates the permeation of water[16]. The proximity to water or SC lipids has an impact on the formation of TG clusters, due to the reorientation of the TG molecules induced by the presence of an interface (as previously described in section 3.3.3 and section 4.3.1 for water and the SC bilayer respectively). The positions of the central carbon atom of the TG glycerol group (C2) within the system were used to estimate the number and shape of TG-head group clusters via 3D Voronoi tessellation and neighbour searching using the methodology described in the Methods section of chapter 3 (section 3.2.3). The obtained clusters are represented in the two simulation snapshots shown in Figure 5.6 at two different times, 1 µs (left) and 5 µs (right). Different clusters are shown here with different colours, the positions of C2 atoms in TG and water atoms are highlighted with yellow and cyan beads, respectively. By comparing the two snapshots, one can observe that TG clusters slowly join up into one single percolating cluster (multiple clusters can be observed

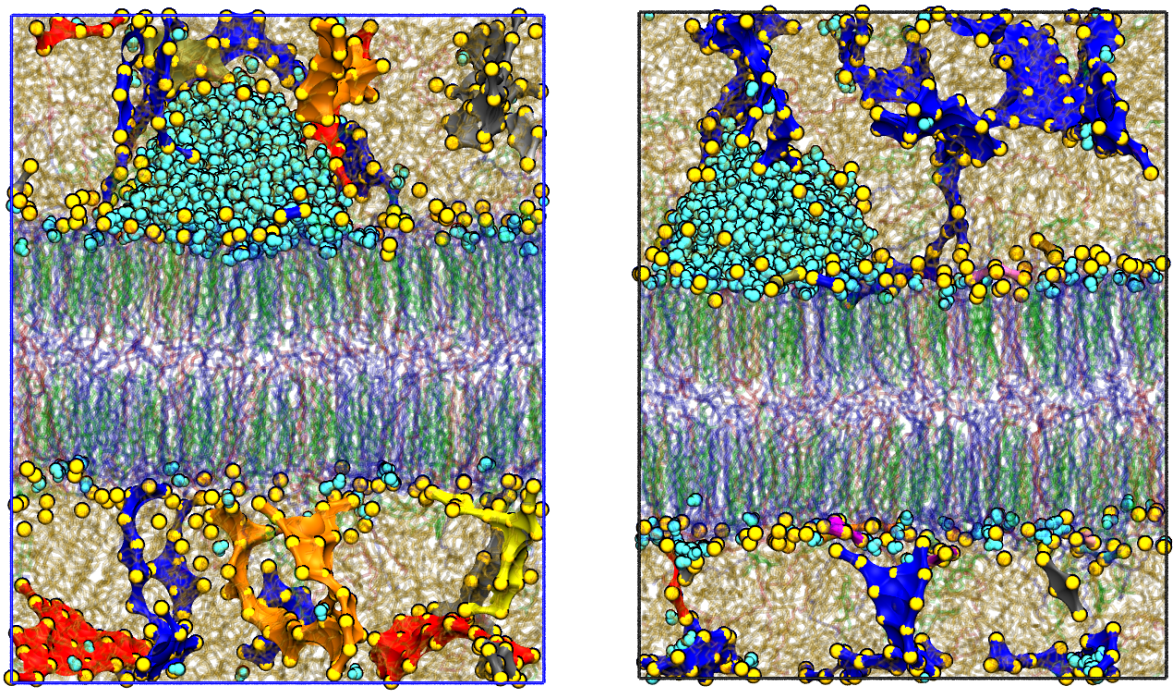
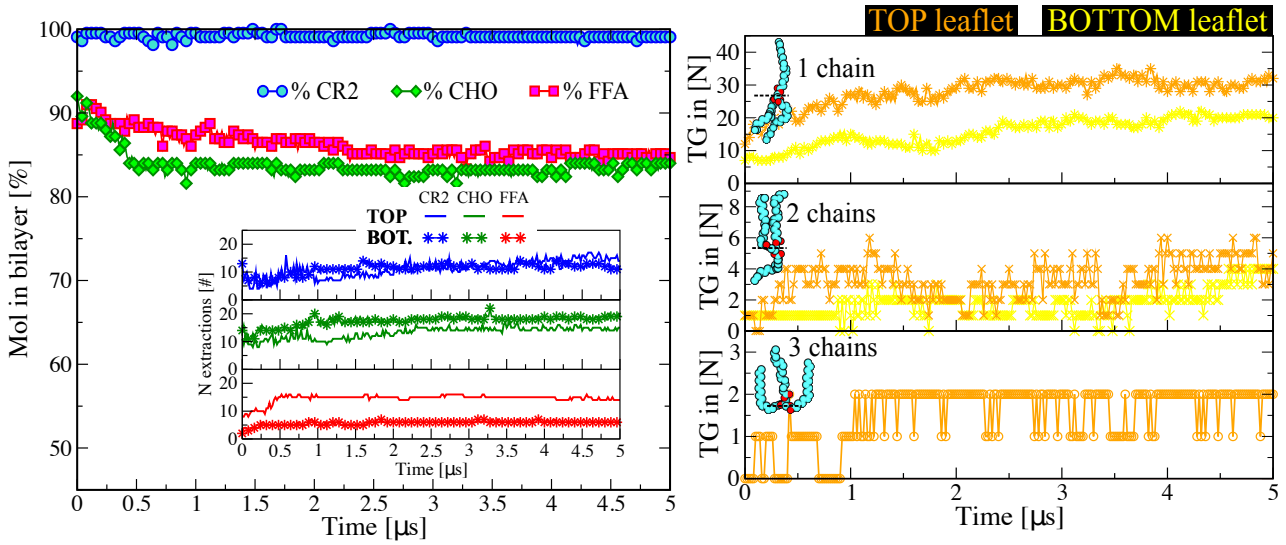


Figure 5.6: Two representative snapshots from the simulations at 1 µs (left) and 5 µs (right). The clusters of TG head groups are shown using different colours. The positions of the central carbon atom of TG molecules, C1, and water atoms are highlighted with yellow and cyan beads, respectively.

at 1 µs, and one main percolating cluster at 5 µs), forming a continuous network linking the water droplet and both sides of the bilayer. The percolating network is partially disrupted at the interface with the bilayer where TG interacts with SC lipids, demonstrating that at this interface TG has competing interests of forming clusters with other TG molecules or interacting with the SC lipids and possibly anchoring in the bilayer. In addition, all the water molecules dispersed in the sebum oil can be found within or in the proximity of the TG clusters, showing how water molecules move through the TG layer by interacting with the network of TG head groups (glycerol moiety) as previously reported in chapter 3 and ref. 16. This further supports the hypothesis suggested in chapter 3 that TG molecules might play a role in regulating the hydration of the outermost layer of the skin.

5.3.2 Structure of the SC bilayer

In this section the stability and the structural properties of the SC bilayer are presented and discussed. The results below can be set against the ones presented in section 4.3.2 of chapter 4 for a comparison with the anhydrous case. In the previous chapter, it has been concluded that



(a) Bilayer stability **(b)** TG insertion

Figure 5.7: a) Percentage of lipids remaining in the bilayer *vs.* time, blue circles refer to CR2 molecules, green diamonds to CHO and red squares to FFA. The inset shows the number of extraction events *vs.* time for the three lipid species, CR2, CHO and FFA (from top to bottom). The number of events, N, in the top leaflet are represented with a straight line and in the bottom leaflet with asterisks. b) Number of anchored TG molecules *vs.* time in the bottom (yellow curves and symbols) and top (orange curves and symbols) leaflet, with one (top-graph), two (middle-graph) or three (bottom-graph) anchored chains.

the sebum oil is able to extract SC lipids from the bilayer. In this chapter, it is important to address whether the presence of water has an impact on the mechanisms of extraction and on the bilayer stability. In Figure 5.7a, the percentages of lipids remaining in the lipid bilayer are reported as a function of time. After 5 µs, all the CR2 molecules remain inside the bilayer, with some of their chains ($\sim$ 10 in each leaflet) in the TG phase (as described previously as "CR2 split"). The lack of full CR2-extraction events may be due to the large activation barrier associated to this event. In addition, the presence of the few water molecules at the interface may enhance the hydrogen bond network between head groups and stabilise the bilayer. The overall percentage of CHO and FFA extraction follows, instead, the same behaviour as in the previous chapter, reaching a plateau at the same equilibrium concentrations of SC lipids in the

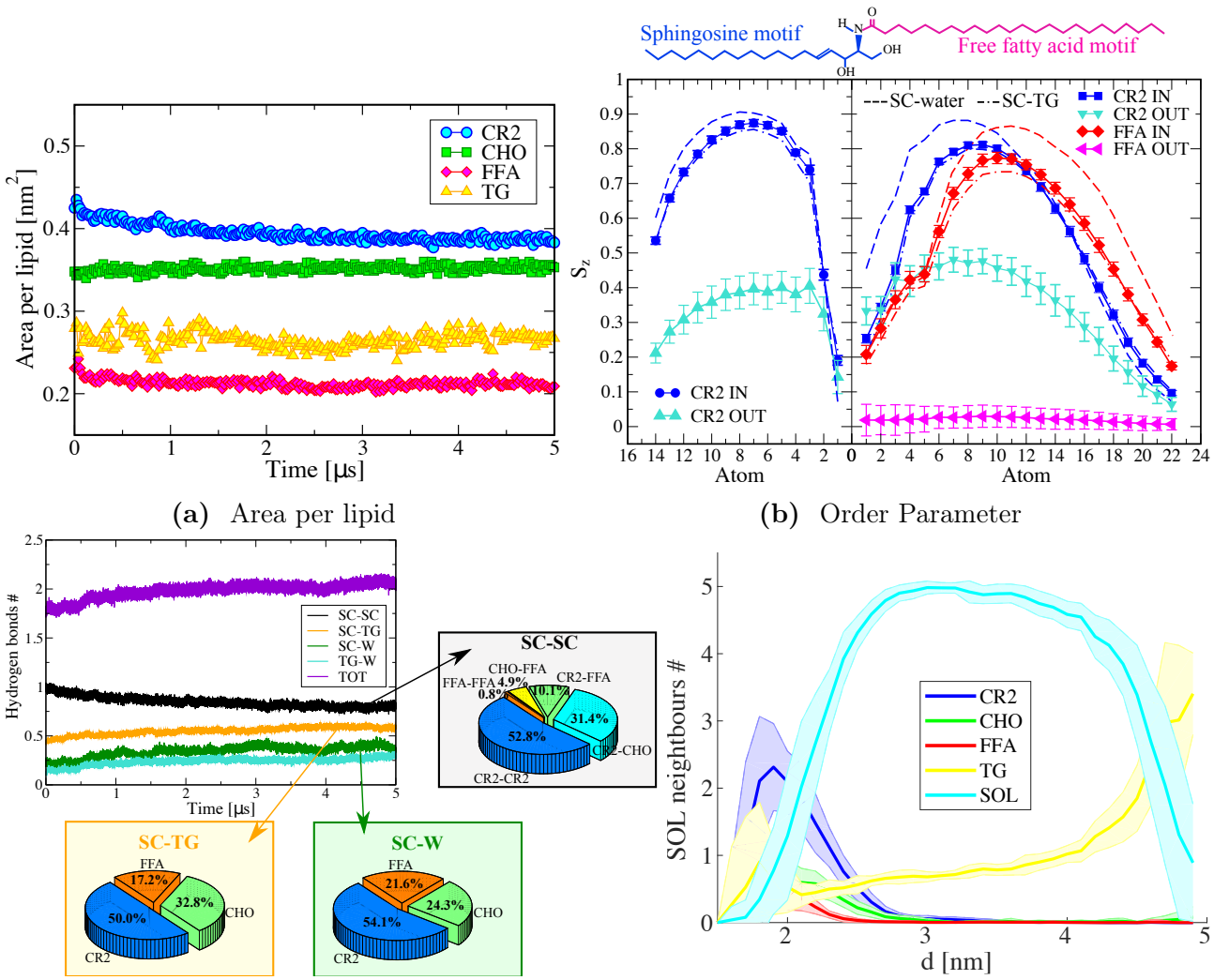
sebum phase as in the anhydrous case. The main difference from the anhydrous case is, instead, found in the number of TG insertions in the top leaflet, where the droplet is located. Figure 5.7b shows that TG molecules preferentially anchor to the top leaflet (∼30 *vs.* ∼20 events in the top and bottom leaflet respectively), indicating an higher tendency to insert in the bilayer in the presence of water (with an increase of 30% in the number of total events). This phenomena can also be observed in the snapshots of Figure 5.5-bottom, where the number of TG tails inside the top leaflet progressively increases and is significantly higher than in the bottom leaflet. To make space in the bilayer for the increase in TG insertions around the droplet, more FFA escapes from the top leaflet increasing the difference between top and bottom leaflets (and also slightly more CHO compared to the previous chapter, with a reduction in the extraction asymmetry). As discussed in the previous chapter, the extraction of FFA molecules is driven by the TG distribution near the bilayer surface, since the FFA head groups extract via the network of TG glycerol groups. This explains the local connection between the number of anchored TG and the FFA extraction events along the TG-SC interface.

A number of properties of the lipid bilayer, namely the area per lipid, hydrogen bonding and the order parameter, were computed, following the procedures explained in section 4.2.2, to measure the impact on the lipid bilayer of adding a water droplet at the TG-SC interface (see Figure 5.8). The evolution of the area per lipid of the different lipid species is shown in Figure 5.8a. As in the anhydrous case, the values of A_L are fairly constant during the simulation, confirming the stability of the bilayer. The differences in values with and without the water droplet (from the previous chapter) are minimal and within the error bar (see Table 5.1).

Lipid	A_L [nm^2]	Err [nm^2]
CR2	0.39382	0.07687
CHO	0.35176	0.05539
FFA	0.21220	0.03897
TG	0.26446	0.11692

Table 5.1: Mean value and standard deviation of A_L measurements. The data are averaged over the simulation length. The error represent the standard deviation of the measurement via Voronoi tessellation method (see section 4.2.2).

The computation of the acyl chain order parameter also shows a similar trend to the previ-



(a) Area per lipid **(b)** Order Parameter

(c) Hydrogen bond **(d)** H_2O neighbours

Figure 5.8: Properties of the SC lipid bilayer. a) Area per lipid *vs.* time for SC and sebaceus lipids, the different lipid contributions are highlighted with different colours and symbols, as specified in the legend. b) Order parameter, S_z, of the lipid tails *vs.* atom index. S_z for the C16 tails of CR2 molecules are reported on the left and S_z for the C24 tails of CR2 and FFA molecules on the right. Different curves correspond to the contribution of molecule inside (IN) and outside (OUT) the bilayer. The dashed curves represent the tail order parameters of a SC lipid bilayer in water, while the dot dashed curves represent the tail order parameters of a SC lipid bilayer in TG. c) Number of hydrogen bonds per lipid *vs.* time. The intra-bilayer (SC-SC), extra-bilayer (SC-TG), SC-water (SC-W) and TG-water (TG-W) contributions are represented separately. SC lipid contributions (expressed as %) are reported in the coloured boxes outside the main plot. d) Number of neighbours to water molecules *vs.* distance from the bilayer mid plane, *d*. The different lipid contributions are highlighted with different colours and symbols, as specified in the legend. Data for order parameter, hydrogen bond and number of neighbour were averaged over the last 1 μs of simulation.

ous chapter, with the values for the TG-SC-water system intermediate between the SC-water (dashed lines in Figure 5.8b) and TG-SC (dot dashed lines in Figure 5.8b) cases. The data is closer to the TG-SC system due to the low amount of water in the system under investigation. The total number of hydrogen bonds (purple curve in Figure 5.8c) is higher that in the anhydrous case (≈ 2 *vs.* ≈ 1.5 hydrogen bond per lipid), due to the formation of additional hydrogen bonds between lipids and water. Water makes hydrogen bonds in equal number with TG and SC lipids. Within the water-SC hydrogen bonds, the majority of those ($\approx 54\%$) are made with CR2 (for the possibility of this lipid molecules to serve as acceptor and donor in multiple sites).

In addition, by computing the number of neighbours to water atoms as a function of the distance from the bilayer mid-plane, d, it is possible to look specifically at the lipid composition around the droplet. The number of neighbours was computed by counting the surrounding Voronoi cells to water atoms. Neighbouring cells associated with different lipid species or water were accounted for, separately, and plotted in Figure 5.8d. The data are time-averaged on the last 1 µs of simulation. A distance, $\bar{d}$, of ~ 2.3 nm marks the boundary between the TG phase and the SC bilayer. For $d \geq \bar{d}$, within the TG oil phase, SOL molecules are mainly surrounded by other SOL molecules, especially within the droplet, or TG atoms. For $d \leq \bar{d}$ instead, the most abundant lipid neighbours are CR2 molecules. CHO, despite being present with the same concentration as CR2 within the bilayer does not feature a similar peak, probably due also to its position within the bilayer, lower than the SC-TG interface (as reported in chapter 4 and ref. 12). Interestingly, it is possible to observe a peak in the TG neighbour distribution below the TG-SC interface (for $d \leq \bar{d}$), suggesting that some inserted TG chains surround the water droplet.

5.3.3 Comparison between SC and DPPC

We discuss now the comparison between the wettability of SC and DPPC lipids. Specifically, a DPPC bilayer was chosen as a model to reproduce the properties of cell membranes in other biological tissues, as DPPC is one of the most commonly used phospholipids in experimental

and simulation studies. As specified above, a water droplet was placed at the interface between the sebum oil and the lipid bilayer and the evolution of the water shape was monitored to estimate the wettability regime of the system.

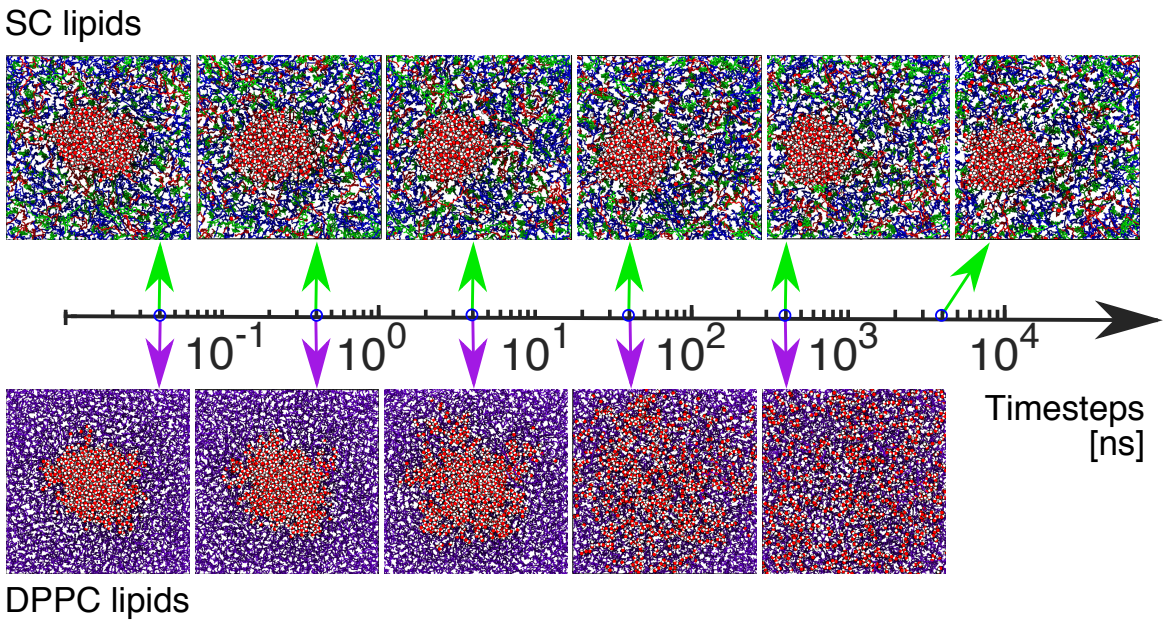


Figure 5.9: Footprint of the water droplet on the bilayer surface *vs.* time (on a logarithmic scale) for the SC-TG (top) and TG-DPPC (bottom) interfaces. Water oxygen and hydrogen are represented as red and white beads, respectively. CR2 lipids in blue, CHO in green, FFA in red and DPPC in purple.

Interestingly, the comparison between the two sets of simulations shows that SC and DPPC lipid membranes perform very differently in terms of water wettability. The droplet footprint at the interface for SC and DPPC lipids *vs.* time is shown in Figure 5.9. The time is here shown in log scale to capture the fast evolution of the water droplet in the DPPC-TG case. As already highlighted, the footprint of the water droplet at the SC interface is fairly constant and well defined over the entire simulation time scale. The situation at the DPPC interface is strikingly different, after just 4 ns the water starts spreading at the interface and after 0.4 µs the water molecules are completely spread out and the droplet shape has disappeared.

The system coordinates from the last 100 ns of the simulations were used to construct the 2D density profiles in cylindrical coordinates (r and z) presented in Figure 5.10, where the contributions of water, lipids (SC or DPPC) and TG are shown separately.

In the case of SC lipids (see graphs on the top of Figure 5.10) the density distribution of the water droplet has the shape of a liquid lens made up of two spherical caps. This is confirmed by

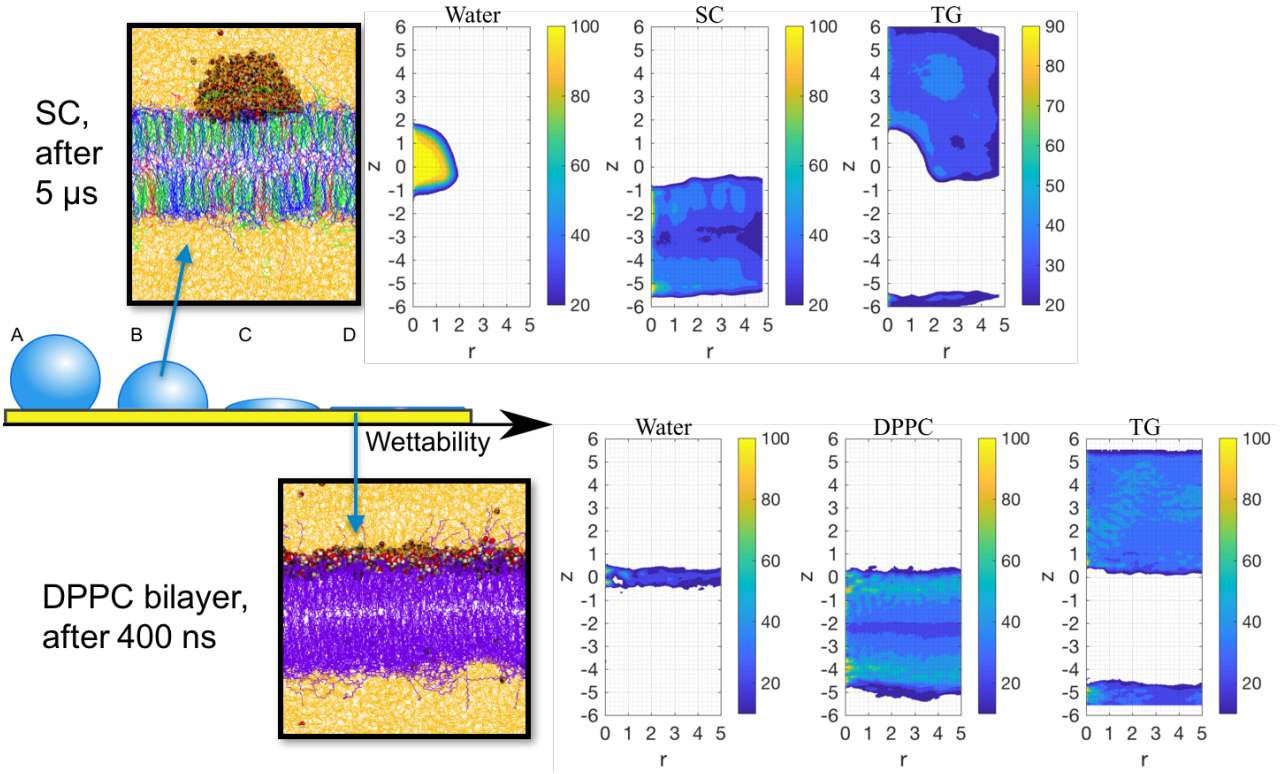


Figure 5.10: Representative snapshots from the simulations and density profiles in cylindrical coordinates, from left to right: water, lipid bilayer and TG density. SC bilayer on the top and DPPC bilayer on the bottom. The different wetting regimes of the two bilayers are illustrated by the schematic wettability scale. Atom positions are averaged over the last 0.1 µs of trajectory. z and r are expressed in nm, and all the densities are expressed in terms of pseudo-particle/nm^3 (as explained in the Method section). The link between colours and density values can be read on the colour bar nearby the plots.

the profiles of SC and TG lipids where the boundaries of the lipid and TG densities wrap around the profiles of the two spherical caps. The lipid profiles are not completely uniform, reflecting structural features of the lipid species. In the SC bilayer there is a reduction in density in the space between the two leaflets where just the longer C24 tails of FFA and TG interdigitate. In the TG plot, regions of higher density are due to the TG clustering via the glycerol groups. The well-defined shape of the water droplet corresponds to a spreading coefficient, $S < 0$, and consequently to the partial wetting regime (see next section for the full characterisation and discussion of SC water wettability).

As observed in Figure 5.9, when in contact with a DPPC bilayer, instead, the droplet collapses and all the water molecules cover quickly the DPPC-TG interface. In this case the simulations were stopped after 0.4 µs, with no need to extend them further. The water density profile shows an homogeneous rectangular profile, as does the boundary of the DPPC and TG density

contours (see graphs on the bottom of Figure 5.10). Therefore, in this case, the spreading factor, S, is greater than 0, corresponding to the regime of full wetting. This indicates that phospholipid membranes are significantly more hydrophilic than the SC membrane.

These results suggest that the water behaviour in TG-SC systems is due to the SC membrane composition, which confers special properties to this type of membrane. In addition, they confirm that the preservation of the water droplet and lack of full wetting in the SC-TG simulations is not due to an artefact of the simulation method used, since with DPPC, the droplet quickly speads out over the membrane surface reaching full wetting.

5.3.4 The wettability of the SC-TG interface

The shape of the water droplet, defined by the isodensity profile of water, was analysed to evaluate the contact angles θ_A and θ_B as explained in the Methods section (5.2.2).

Angles	value [°]
θ_A	$104° \pm 5°$
θ_B	$32° \pm 5°$

Table 5.2: Contact angles described by the two spherical caps used to characterise the water droplet. The angles are defined in Figure 5.4.

The values of θ_A and θ_B, presented in Table 5.2, were found to be fairly constant over time, confirming the stability of the water droplet at the interface, as discussed in section 5.3.1. Since the droplet maintains a well-defined shape, the spreading parameter, S, is lower than 0, and water at the SC interface is in the partial wetting regime. The suprasurface contact angle, θ_A, gives a measure of the wettability of the SC surface. The more hydrophobic a substrate is the higher the contact angle of the droplet placed on the surface, with angles $\geq 90°$ indicating a hydrophobic surface. From the computations presented in this chapter, θ_A measures $104° \pm 5°$ (see Table 5.2), highlighting the low affinity between the SC surface and water. In addition, the SC surface is a soft interface and it modifies its topology next to the water droplet, bending around the shape of the droplet. The subsurface angle, θ_B, captures this behaviour, describing the deformability of the SC surface. A similar angle can be extracted by analysing the work of Del Regno and Notman[12], where the interfacial water, initially evenly distributed between

two SC bilayers, equilibrates into water pools, resulting in a negative value for the spreading coefficient, S.

A close comparison to experimental results is challenging due to the method used to measure of the wettability of the skin surface in experiment, where a droplet of solvent is placed at the air/skin interface. Doing so, in a sebum rich area like the forehead, will mean that the measurement corresponds to the wettability of the skin surface lipid film, which is mainly composed of sebum, with a very low percentage of SC lipids. On the other hand, after treatment with ether to remove the sebum oil from the skin surface, the measured values of the contact angle (in the range 89°–101° [2]) compare quite well with our computations (≈ 104° - see Table 5.2). However, it is not clear what the microscopic effect of this treatment is on the skin surface and whether it really exposes the SC bilayer without also changing the properties and the integrity of the bilayer.

Several studies [24, 27, 28] have highlighted the challenge of employing continuum wetting theories to characterise the wettability of soft solid substrates.

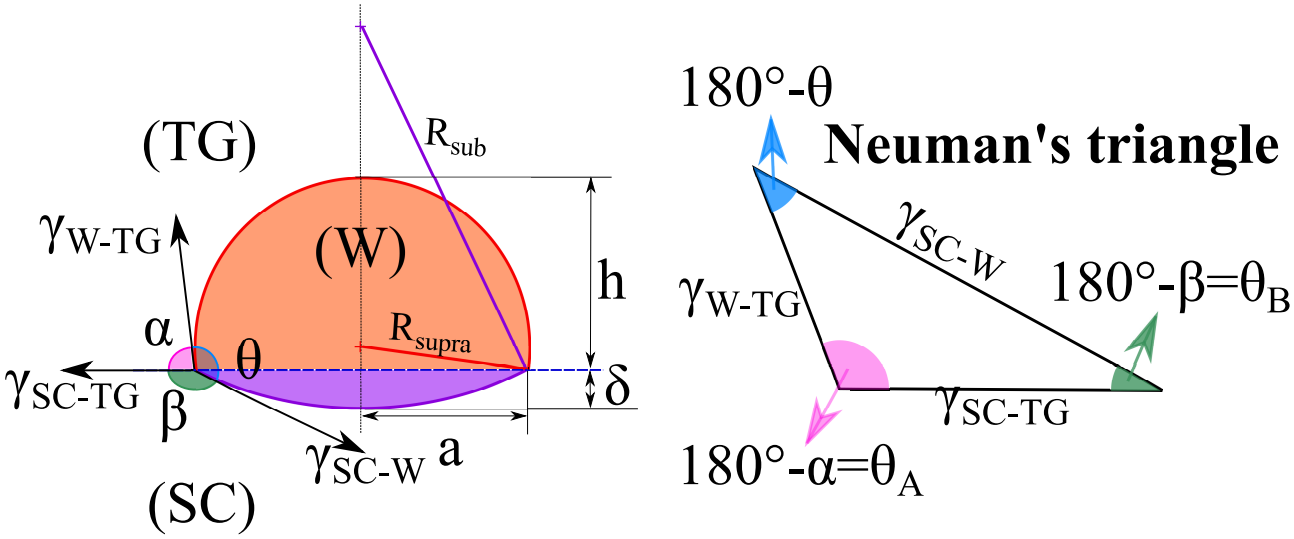


Figure 5.11: Left) Representation of the two spherical caps defined by the supra (red) and sub (purple) surface. The droplet immersion, h, the spherical caps radii (R_{sub} and R_{supra}) and the penetration depth, δ are indicated in the figure. Right) Surface tensions and angles relationships represented with the Neumann's triangle construction.

As a first consideration, one can think of the droplet in the limiting case of the wetting of a liquid on an hard surface. Under this approximation the balance of the surface tensions is

reformulated for a planar interface and gives

$$\Delta E_1 = \pi a^2 \left[(\gamma_{SC-W} - \gamma_{SC-TG}) + \gamma_{W-TG} \left(1 - \frac{h}{R_{supra}} \right) \right], \tag{5.7}$$

where $\left(1 - \frac{h}{R_{supra}}\right)$ is the geometrical derivation of the contact angle, θ_A, as shown in equation 5.6 and πa^2 is the area of the droplet footprint (see Figure 5.11-left for labels).

Following the work of Léonforte and Müller[24], when the wetting of the liquid involves a deformation of the substrate surface, δ, which is small in comparison to the droplet height, h, the free energy cost of the droplet can be written as

$$\Delta E_2 = -\pi a^2 S + \pi \left[\gamma_{W-TG} h^2 + \gamma_{SC-W} \delta^2 \right], \tag{5.8}$$

where the spreading parameter, S, can be written as $S = \gamma_{SC-TG} - \gamma_{SC-W} - \gamma_{W-TG}$.

The free energy gained by deforming the surface, can therefore be quantified from equations 5.7 and 5.8 as $\Delta E = \Delta E_2 - \Delta E_1$, resulting in the expression

$$\Delta E = \pi \left[h^2 \left(1 + \frac{a^2}{h R_{supra}} \right) \gamma_{W-TG} + \delta^2 \gamma_{SC-W} \right]. \tag{5.9}$$

If the hard solid substrate is replaced by a fluid, immiscible with the liquid droplet, then the angles (α, β, θ) at which the three fluid–fluid interfaces meet at the contact line between the three fluids (TG, W, SC) are governed by the mechanical force balance of the interfacial tensions[27] (γ_{W-TG}, γ_{SC-TG} and γ_{SC-W}), as derived in equation 5.5, which in the case of the TG-SC interface gives:

$$\begin{cases} \gamma_{W-TG} \sin\theta_A = \gamma_{SC-W} \sin\theta_B \\ \gamma_{TG-SC} = \gamma_{W-TG} \cos\theta_A + \gamma_{SC-W} \cos\theta_A, \end{cases} \tag{5.10}$$

with θ_A and θ_B which can be related to the three angles shown in Figure 5.11-left, as $\theta_A = 180° - \alpha$, $\theta_B = 180° - \beta$ and $\theta_A + \theta_B = \theta$. Equation 5.10 can be represented as a set of vectors in two-dimensions, forming a closed triangle known as Neumann's Triangle (see Figure 5.11-right).

Because fluids by definition have no reference configuration, configurational energy balance and force balance are indistinguishable. In the triangle, represented in Figure 5.11-right, the three edge lengths correspond to the three surface tensions, as indicated in the figure, and the three angles are the complement to 180° of α, β and θ (the colour of the angles in Figure 5.11-left match the colour of their complement to 180° in Figure 5.11-right).

When, instead of an immiscible liquid there is a soft material the contribution of the material resistance to deformations needs to be considered. To determine whether it is crucial to take into account this effect, one can look at the ratio between the surface energy and the Young's modulus of the substrate[27]. The latter will be high for solid substrates where the Young's approximation can be used and very low for liquid substrates where the Neumann construction should be used.

5.4 Conclusions and future work

In this chapter the wettability of the SC bilayer at the sebum interface has been investigated. A water droplet, placed at the interface, remains stable during microseconds simulations. The droplet stability has been assessed by looking at the water droplet volume over time and its footprint on the bilayer surface. During the long simulation time, some water molecules reach the lower SC-TG bilayer interface via the sebum phase, by interaction with the clustered network of TG head groups. The barrier for water penetration through the TG phase, measured in chapter 3 and ref. 16, is in fact much lower than the energy barrier for water to cross the SC bilayer, measured by Das *et al.*[13] ($\approx 9k_BT$ *vs.* $\approx 15k_BT$). From the number of water molecules at the lower interface, which stabilises after 2 µs, it is possible to estimate the number of water molecules per SC lipid, which measures 0.2. This number suggests a very low affinity between SC lipid and water. Being significantly lower that the number of water molecules per lipid estimated experimentally (*i.e.* 1.91 ± 0.42[8]), it also supports the suggestion that water is mainly located in water pools rather than homogeneously distributed across the SC head group region[12]. This hypothesis should be tested with new experiments.

The presence of the water droplet at the TG-SC interface increases the anchoring of TG molecules in the top leaflet, with a consequent increase in the number of FFA extractions. The total number of SC lipids found in the TG phase remains comparable with the results measured in chapter 4. The increased number of TG insertions occurs in the proximity of the water droplet, as shown by the peak in the water neighbours below the interface featured by TG atoms.

Given the stability of the droplet at the TG-SC interface, the corresponding spreading parameter of the interface, S_{TG-SC}, is lower than 0. This differs from the case of the DPPC-TG interface, where the water molecules from the droplet quickly spread to cover the interface demostrating that $S_{TG-DPPC} > 0$.

In the TG-SC case, the droplet shape was approximated with two spherical caps by considering the iso-density profiles of the water in cylindrical coordinates, above and below the interface. Two contact angles were measured accordingly, θ_A which measures $\approx 104°$ and corresponds to the wettability of the SC bilayer, and θ_B which measures $\approx 32°$ and gives an estimate for the deformability of the SC bilayer. From these two angles, the three angles required to construct the Neumann triangle can be derived, to test the continuum wettability theories[29]. However, while the value of γ_{TG-W} is available from the simulation presented in chapter 3, the value of γ_{SC-W} and γ_{SC-TG} needs to be estimated to address whether the continuum theory is effective in reproducing the properties of the bilayer interface. This will be the aim of future work.

Moreover, the same approach, used in this chapter to investigate the wettability of bilayers to water, could be reproduced to investigate the wettability properties of SC or other lipid membranes to other solvents such as urea, ethanol, glycerol, formamide. These solvents are often used in the formulation of creams for topical application, where the wettability of sebum and SC lipids is of high interest.

It needs to be noted that the spherical shape of the water droplet used in this work presents several limitations in estimating wettability properties of a substrate[24]. It would be interesting to repeat the simulations presented in this chapter with a different input configuration, where a cylinder of water is inserted at the interface, instead of the water droplet. This technique

has been previously employed in computer simulation studies of liquid lenses at liquid-liquid interfaces[30] in order to generate a three phase line with zero curvature.

Bibliography to chapter 5

[1] J. Krawczyk, *Skin Research and Technology*, 2015, **21**, 214–223.

[2] A. Mavon, H. Zahounani, D. Redoules, P. Agache, Y. Gall, P. Humbert, H. Zahouani, D. Redoules, P. Agache, Y. Gall and P. Humbert, *Colloids and Surfaces B: Biointerfaces*, 1997, **8**, 147–155.

[3] P. Capra, G. Musitelli and P. Perugini, *International Journal of Cosmetic Science*, 2017, **39**, 393–401.

[4] A. Mavon, D. Redoules, P. Humbert, P. Agache and Y. Gall, *Colloids and Surfaces B: Biointerfaces*, 1998, **10**, 243–250.

[5] S. Derler and L.-C. C. Gerhardt, *Tribology Letters*, 2011, **45**, 1–27.

[6] A. V. Rawlings, *Skin Barrier*, CRC Press, 2005, pp. 399–425.

[7] R. R. Warner, M. C. Myers and D. A. Taylor, *The Journal of Investigative Dermatology*, 1988, **90**, 218–224.

[8] D. Groen, G. S. Gooris, D. J. Barlow, M. J. Lawrence, J. B. Van Mechelen, B. Demé and J. A. Bouwstra, *Biophysical Journal*, 2011, **100**, 1481–1489.

[9] J. Bouwstra, G. Gooris, J. van der Spek, S. Lavrijsen and W. Bras, *Biochimica et Biophysica Acta (BBA)/Lipids and Lipid Metabolism*, 1994, **1212**, 183–192.

[10] T. J. McIntosh, M. E. Stewart and D. T. Downing, *Biochemistry*, 1996, **35**, 3649–3653.

[11] M. de Jager, W. Groenink, J. van der Spek, C. Janmaat, G. Gooris, M. Ponec and J. Bouwstra, *Biochimica et Biophysica Acta - Biomembranes*, 2006, **1758**, 636–644.

[12] A. Del Regno and R. Notman, *Physical Chemistry Chemical Physics*, 2018, **20**, 2162–2174.

[13] C. Das, P. D. Olmsted and M. G. Noro, *Soft Matter*, 2009, **5**, 4549–4555.

[14] R. Thind, D. W. O'Neill, A. Del Regno and R. Notman, *Chem. Commun.*, 2015, **51**, 5406–5409.

[15] A. Akinshina, C. Das and M. G. Noro, *Phys. Chem. Chem. Phys.*, 2016, **18**, 17446–17460.

[16] A. S. Tascini, M. G. Noro, R. Chen, J. M. Seddon and F. Bresme, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1848–1860.

[17] G. Bussi, D. Donadio and M. Parrinello, *Journal of Chemical Physics*, 2007, **126**, 014101.

[18] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *The Journal of Chemical Physics*, 1995, **103**, 8577–8593.

[19] O. Berger, O. Edholm and F. Jähnig, *Biophysical Journal*, 1997, **72**, 2002–2013.

[20] D. Peter Tieleman, H. J. Berendsen and M. S. Sansom, *Biophysical Journal*, 1999, **76**, 1757–1769.

[21] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *Journal of Computational Chemistry*, 2005, **26**, 1701–1718.

[22] P.-G. de Gennes, F. Brochard-Wyart and D. Quéré, *Capillarity and Wetting Phenomena*, 2004, p. 292.

[23] M. Nosonovsky and R. Ramachandran, *Entropy*, 2015, **17**, 4684–4700.

[24] F. Léonforte and M. Müller, *Journal of Chemical Physics*, 2011, **135**, 1–13.

[25] V. Pratt, *ACM SIGGRAPH Computer Graphics*, 1987, **21**, 145–152.

[26] C. H. Rycroft, *Chaos*, 2009, **19**, 041111.

[27] C.-Y. Hui and A. Jagota, *Soft Matter*, 2015, **11**, 8960–8967.

[28] L. Chen, E. Bonaccurso, T. Gambaryan-Roisman, V. Starov, N. Koursari and Y. Zhao, *Current Opinion in Colloid and Interface Science*, 2018, **36**, 46–57.

[29] L. Boruvka and A. W. Neumann, *The Journal of Chemical Physics*, 1977, **66**, 5464–5476.

[30] F. Bresme and N. Quirke, *The Journal of Chemical Physics*, 2000, **112**, 5985–5990.

Chapter 6

Nanoparticle interfaces

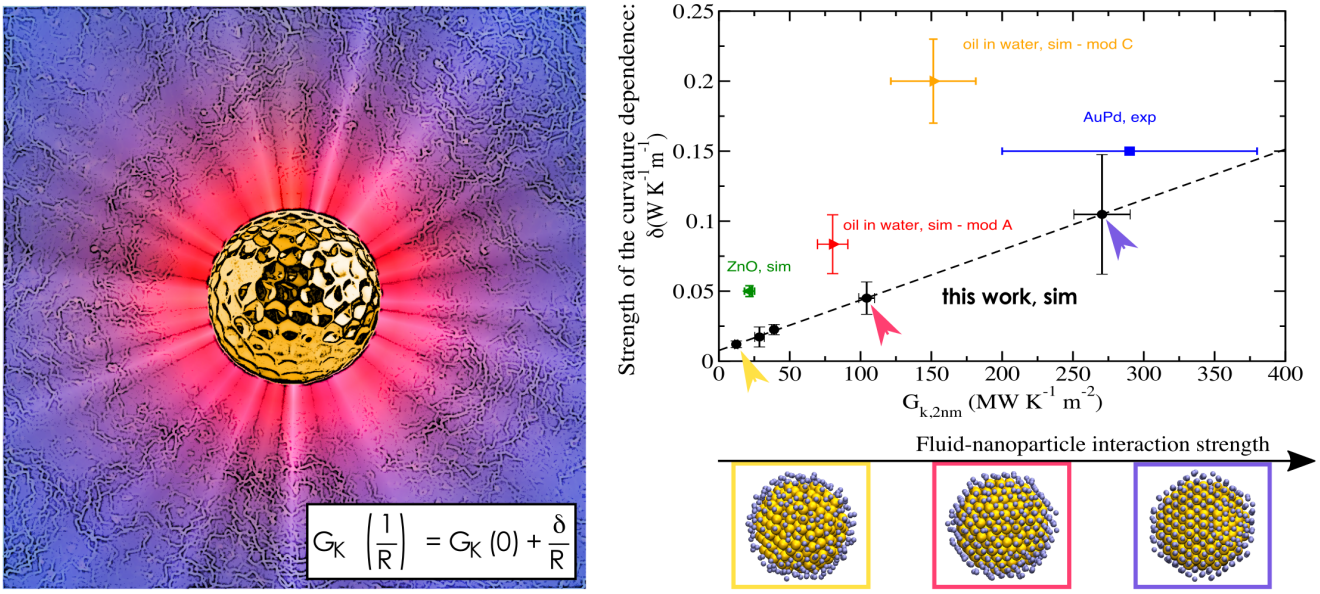


Figure 6.1: Title page figure for chapter 6. This picture has been used as Table of contents for our publication: Tascini *et al.*[1].

6.1 Introduction

Nanoparticles are assemblies of inorganic or organic building blocks, with size in range 1–100 nm. Such assemblies can exhibit exciting new properties different from those of corresponding bulk materials. Nanoparticles offer numerous opportunities in medicine[2], analytical chemistry[3] and technological application[4]. Often they are coated with a wide range of ligands, which provide specific functionality in catalysis[5, 6], biological compatibility[7, 8] or define the mechanical response of nanoparticle arrays[9].

Several nanoparticle applications involve thermal transport, *e.g.* light can be converted into heat by nanoheaters, made up of metal nanoparticles and metal nanostructures. This property is being exploited to set up nanoscale mass fluxes[10, 11] and in thermal therapy applications, where the nanoparticles can bind specifically to cells. Kotaidis *et al.*[12] measured a temperature increase next to the particle of hundreds of degrees, resulting in the explosive evaporation of the fluid, a phenomenon that takes place in subnanosecond timescales. Another way for using nanoparticles in thermal therapies is to exploit the magnetic properties of particles containing iron oxide cores, which can be heated by applying magnetic fields[13].

The description of heat flow from the nanoparticles to the surrounding fluid involves the full complexity of thermal transport at the nanoscale[14, 15]. Indeed, interfacial properties, *i.e.* surface free energies, interfacial curvature and interfacial structure, can strongly influence the heat transport. As a matter of fact there is evidence that nanoparticles influence the diffusion of water at the nanoparticle-water interface (see *e.g.* ref. 16). A deep understanding of the mechanisms that determine and control the transport of heat from the particle to the surrounding medium is crucial in the development of technological applications involving nanoparticle heating. This problem inevitably involves the investigation of thermal transport through the nanoparticle-fluid interface, and the quantification of the interfacial thermal conductance. The latter determines the temperature increase of the surrounding fluid as well as the characteristic time scale of the temperature relaxation towards the equilibrium temperature. Hence, the quantification of the thermal conductance of nanoparticle-fluid interface is a problem of evident practical significance.

The dependence of the thermal conductance on chemical composition has been shown by thermoreflectance experiments of monolayers adsorbed on solid substrates[14, 17]. Wilson *et al.*[17] show that hydrophilic interfaces feature higher thermal conductances than hydrophobic interfaces ($\sim 150\,\mathrm{MW\,m^{-2}\,K^{-1}}$ *vs.* $\sim 50\,\mathrm{MW\,m^{-2}\,K^{-1}}$ for alkane layers with polar terminal groups and alkane layers with methyl terminal groups, respectively). This dependence has been confirmed via MD simulations[18], with computed thermal conductances in quantitative agreement with the experimental data. In addition, previous studies on flat surfaces have suggested a linear dependence between the thermal conductance and the work of adhesion of the surface[19]. Such a correlation might be useful to predict heat transport in nanoparticle-fluid interfaces. This is a significant objective given the difficulty associated with experimental measurements at the nanoscale. However, the generality of such correlations has not been established yet, and their applicability to nanoparticles remains a challenging question, which will be addressed in this chapter.

Studies by Bresme and co-workers have considered the impact of curvature on thermal conductance, suggesting a strong dependence of the thermal conductance with nanoparticle curvature[20]. The thermal conductance of alkane nanoparticles immersed in water, computed using transient non-equilibrium MD simulations, has the order of magnitude of the conductance of hydrophilic surfaces. However, an extrapolation of the thermal conductances to those of a flat wall (zero curvature) predicted much lower conductances, with values in agreement with those reported for flat hydrophobic interfaces[14, 19]. More recently, the dependence of the thermal conductance with curvature has been confirmed in simulation studies of metal oxide nanoparticles immersed in octane[21], and in theoretical studies of nano-bubbles and nanodroplets[22, 23]. Computer simulations have further highlighted the importance of intermolecular interactions and interface structure on thermal conductance[24, 25, 26, 27], and the relevance of interfacial curvature on vapour nucleation around nanoparticles[28].

So far the interfacial composition has been identified as a key variable in interfacial thermal transport of flat surfaces in both experimental and theoretical studies. However, the experimental data for nanoparticles are still scarce and the estimates of the thermal conductance involve significant uncertainties. Overall, a good understanding on how the thermal conduc-

tance depends both on interface curvature and fluid-nanoparticle interaction is still lacking. Therefore, in this chapter, this problem is addressed using MD simulations in combination with generic nanoparticle models, to uncover the key correlations between nanoparticle curvature and nanoparticle-fluid interactions, and also how these correlations define the thermal conductance of nanoparticle-fluid interfaces.

The chapter is structured as follows. Firstly, the simulation models and the computational approach to study heat transport across nanoparticle-fluid interfaces is described. A discussion of the main results on the dependence of the thermal conductance of the nanoparticle-fluid interface with different wetting characteristics and nanoparticle curvature follows. This discussion also highlights the implications of our results for the interpretation of existing experimental and simulation data of thermal conductances. This chapter concludes with a final section devoted to a final remark and a discussion of further studies to address the interaction of nanoparticle with lipid membranes. The main outcomes and figures presented in this chapter were published in ref. 1 and reproduced from ref. 1 with permission from the PCCP Owner Societies.

6.2 Methods

6.2.1 Input configuration: nanoparticle model

As a first step to give insight into the thermodynamical properties of nanoparticles in biological systems, this chapter deals with a generic model of a nanoparticle immersed in a fluid (see Figure 6.2). A key quantity in the temperature diffusion across different phase and materials at the nanoscale is the appearence of a temperature jump at the interface, which can be characterised by a themal conductance, G_K. The model used in this work allows the systematic investigation of the interfacial thermal conductance with nanoparticle diameter and nanoparticle-fluid (NP-f) interactions. Due to the generality of the model, which can be applied to different solid nanoparticles and fluids, Lennard-Jones reduced units were here used to represent the results. All the properties given in these units are denoted with the superscript "*".

The procedure to build the system, made up of nanoparticles of different sizes immersed in a LJ fluid, included the steps below:

1. to generate a FCC solid, with lattice constant, l_c, defined as $l_c = (4/\rho_{NP})^{1/3}$, with the density set to $\rho^*_{NP} = \rho_{NP}\sigma^3 = 1$, σ being the atomic diameter.

2. to cut a spherical geometry from the FCC solid lattice, by selecting all the atoms inside a sphere of a given radius, R^*. Particles with different diameters, $2 \times R^*$, were generated to evaluate the dependence of the thermal conductance with the nanoparticle curvature. Specifically, a set of MD simulations were run for $R^* = R/\sigma$=5, 6, 7, 8 and 10 ($\sim$ 1.5-3.5 nm for atom diameters of the order of the gold atom).

3. to immerse the particle in a Lennard-Jones fluid, by adding fluid atoms, whose position were set randomly around the particle. The fluid was modelled with the Lennard-Jones (LJ) potential, and the density was set to $\rho^*_L = \rho_L\sigma^3 = 0.8$, which is close to the density of the Lennard-Jones fluid at triple point conditions[29]. Solid and fluid atoms had the same diameter, σ.

A typical simulation, generated with the protocol described above, consisted of $N \sim 2\times10^4-10^5$ particles in a cubic simulation box with box length L*=30–50, depending on the nanoparticle diameter. Periodic boundary conditions were applied in all directions.

The strength of the interatomic interactions between fluid atoms, ε_f, were set equal to $\varepsilon_f = 1$, while the solid ones, ε_{NP}, to $\varepsilon_{NP} = 10 \times \varepsilon_f$, consistently with the interaction strength employed in LJ models of metals[30]. To investigate the dependence of the interfacial conductance with nanoparticle wettability, the nanoparticle-fluid interaction was tuned using ε_{NP-f} =F$\sqrt{\varepsilon_{NP}\ \varepsilon_f}$. Different wetting states were obtained by varying the parameter F, from full drying to full wetting. The computation of the nanoparticle contact angle with the liquid-vapour interface was used to quantify the wetting state. This calculation should provide a link with experimental studies, which now allow the direct measurements of contact angles of very small nanoparticles[31].

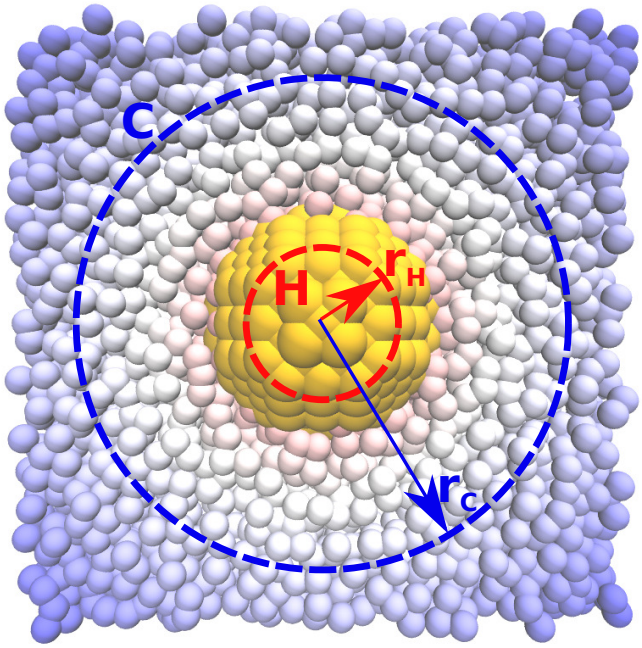


Figure 6.2: Snapshot of a representative configuration of a nanoparticle (yellow) of radius $R^* = 5$ immersed in the fluid (white-blue). The dashed circles represent the boundaries of the hot (H, red inner circle) and cold (C, blue outer circle) thermostats.

The nanoparticle-nanoparticle and fluid-fluid interactions were truncated at a cutoff $r_c^* = 2.5$. Although employing a different cutoff can give slightly different coexistence properties and transport coefficients, the general physical behaviour explored in this study remain the same. This has been proved by computing the temperature profile for different cutoff radii as presented in the results section (see Figure 6.6). Hence, $r_C^* = 2.5$ was chosen, as it offers a good compromise between accuracy and speed of computation.

6.2.2 Non-Equilibrium Molecular Dynamics

A temperature gradient was imposed through the nanoparticle-fluid interface using Boundary Driven Non-Equilibrium Molecular Dynamics (NEMD) simulations. The thermostats were applied to atoms belonging to two regions of the simulation box. The hot thermostat was set as a spherical volume of radius $r_H^* = R^* - 2$, located at the centre of the nanoparticle. The cold thermostat was defined by the region between a spherical surface at $r_C^* = R^* + 7$ and the edge of the simulation box. The shell thickness for the cold layer was equal to 3 (in LJ units) for all the simulations. The boundaries of the two thermostat defined by r_H and r_C are highlighted with

two dashed circles, red and blue respectively, in Figure 6.2. The location of the thermostats does not influence the thermal conductances reported in this chapter. All the atoms in the nanoparticle were tethered to their original positions (at the start of the simulation) using a harmonic potential with a force constant, $k^* = 100$. This potential prevented nanoparticle drifts during the simulation.

The NEMD simulations were implemented using the open source MD simulation code LAMMPS[32], with the following procedure:

1. **System equilibration:** canonical simulations were run using the Nosé-Hoover thermostat with a damping parameter of $\tau^* = 0.2$ for 10^6 steps. The average temperature of the simulations was set to $T^* = 0.72$, and is above the triple point temperature of our LJ model[33]. T^* was selected to generate the liquid-vapor interface used to quantify the wettability of the nanoparticle (see Results section). The equations of motion were integrated using the velocity-Verlet algorithm using $\delta t^* = 0.001$ as the integration timestep.

2. **Set up the NEMD:** The two thermostats were set at the temperatures $T^*_{HOT} = 0.92$ and $T^*_{COLD} = 0.72$, by rescaling the velocities of the atoms inside the thermostatted regions at every timestep . Following the rescaling, the linear momentum of the entire simulation box was reset to zero. To ensure the systems used in the analyses had reached the stationary state, we discarded the first 5×10^5 timesteps.

3. **Production run:** The production stage involved typically 5×10^6 timesteps, and it was used to compute the average properties reported below. All the errors reported in this work are given as one standard deviation of the mean and were obtained from the analysis of ten consecutive trajectories of 5×10^5 timesteps each.

6.2.3 Data analysis

To obtain spatial averages of the main properties of the system, such as density, temperature and energy fluxes, the system was subdivided into n_{BIN} spherical shells of thickness $\delta r^* = 0.2$ (see Table 6.1 for n_{BIN} *vs.* system size), centered around the centre of mass of the nanoparticle.

R*	5	6	7	8	10
$\mathbf{n_{BIN}}$	75	80	85	90	100

Table 6.1: Number of bins (n_{BIN}) corresponding to different particle sizes, R^*.

Heat flux evaluation

The heat flux across the interface is a key quantity in the computation of the interfacial thermal conductance. It can be obtained in two ways:

- from the continuity equation:

$$J_q = \frac{\langle \Delta U \rangle}{A(r)\delta t}, \tag{6.1}$$

  where

  ◇ $\langle \Delta U \rangle$ is the time average of the change in the internal energy associated with the thermostatting process in the hot or cold regions. The velocity rescaling is performed without modifying the particle coordinates, hence $\Delta U(t) = K_{new}(t) - K_{old}(t)$ at time t is obtained every timestep by calculating the difference between the kinetic energies after, K_{new}, and before, K_{old}, the rescaling;

  ◇ $A(r) = 4\pi r^2$, is the area of the surface normal to the heat flux direction, here the area of a spherical surface of radius r;

  ◇ δt is the timestep.

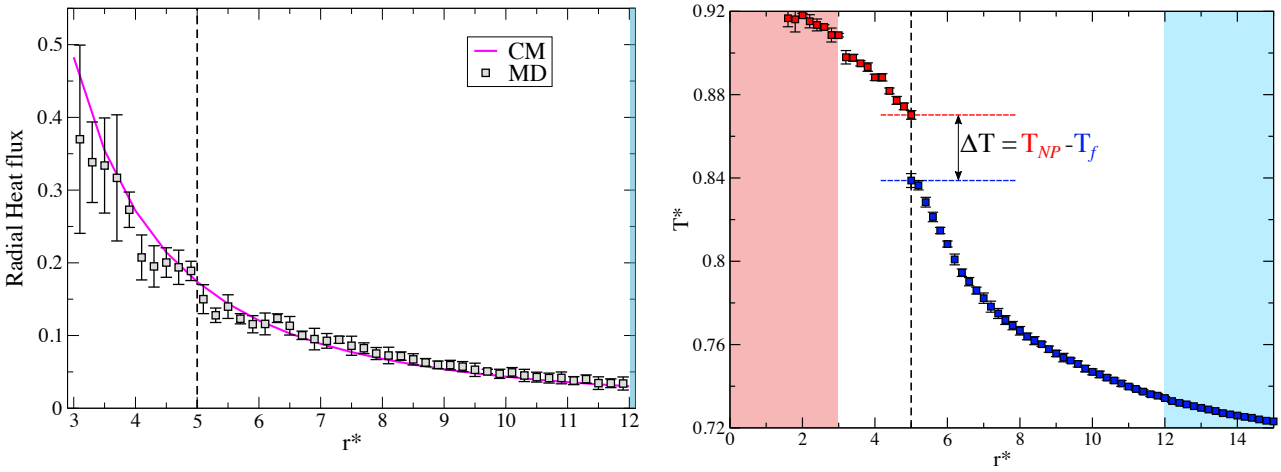
- from the statistical mechanics equation derived by Irving and Kirkwood[34]:

$$J_U(V) = \frac{1}{V} \sum_{i=1}^{N \in V} \left[\frac{1}{2} m_i \mathbf{v}_i^2 \mathbf{v}_i + \phi_i \mathbf{v}_i + \frac{1}{2} \sum_{j \neq i} (\mathbf{v}_i \cdot \mathbf{F}_{ij})\, \mathbf{r}_{ij} \right], \tag{6.2}$$

  where the sum runs over all the particles inside the control volume, V. This volume is the space defined by the spherical shell of radius r and thickness $\delta r^* = 0.2$. m_i and $\mathbf{v}_i$ represent the mass and velocity of particle i, respectively, ϕ_i, is the potential energy of particle i, $\mathbf{F}_{ij}$ the force of particle j on i, and $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$.

The consistency between equations (6.1) and (6.2) has been explored before for charged and polar fluids[35, 36, 37]. These approaches are also consistent in the case of the spherical geometry

considered here. Figure 6.3a shows the good agreement between the statistical mechanics equation (eq. 6.2 - represented with grey squares) and the continuity equation (eq. 6.1 - represented with a straight magenta line).



(a) Heat flux **(b)** Temperature profile

Figure 6.3: a) Radial heat flux for a nanoparticle of radius, $R^* = 5$. The full line (CM) represents the heat flux obtained from the continuity equation (6.1) and the symbols (MD) the internal energy flux from the MD simulations and equation (6.2). b) Representative temperature profile (red squares – nanoparticle, blue squares – fluid) along the radial distance from the centre of mass of a nanoparticle of radius $R^* = 5$ in a simulation box of length, $L_x^* = L_y^* = L_z^* = 30$. The coloured regions, red and blue, represent the location of the thermostats, hot and cold, respectively.

Temperature evaluation

The local temperature was computed using the equipartition formula, as

$$T(r) = \sum_{i=1}^{N \in r} m_i v_i^2 / (k_B N_f), \tag{6.3}$$

where the sum runs over all the particles i located in a shell with radius r, and N_f represents the number of degrees of freedom, which in this case is 3. Figure 6.3b shows a representative temperature profile (for $R^* = 5$) along the radial distance to the centre of the nanoparticle. The temperature profile shows a clear discontinuity at the nanoparticle-fluid interface, $r^* = 5$, which is connected to the interfacial thermal conductance, G_K.

G_K can be derived directly from the temperature gap, $\Delta T = T_{NP} - T_f$, where T_{NP} and T_f are the temperatures of the nanoparticle and fluid at the interface (as shown in Figure 6.3b) and

the energy flux, J_q, at $r = R*$, via

$$G_K = \frac{J_q}{\Delta T}. \tag{6.4}$$

The temperature profile can also be fitted to the solution of the one dimensional heat diffusion equation under stationary conditions[38]. The model was simplified by assuming that the thermal conductivities of the nanoparticle and the fluid are constant and independent of the radial distance to the nanoparticle centre of mass. The temperature profiles from MD simulations were fitted to Fourier's law of heat conduction, which at the steady state is given by $T(r) = \frac{A}{r} + B$, where A and B are constants that can be calculated by considering the boundary conditions of the heat diffusion equation.

By rewriting the heat diffusion equation in spherical coordinates, it can be significantly simplified solving the problem along one dimension, namely the radius of a sphere centred on the NP centre of mass. The heat diffusion equation can then be written as:

$$\frac{Q}{4\pi r^2} = -\lambda_\alpha \frac{dT}{dr}, \tag{6.5}$$

where Q is the heat transferred, $4\pi r^2$ is the corresponding isothermal area and λ_α is the thermal conductivity of phase α (NP or fluid). The differential equation 6.5 can be solved between radius r_1 at temperature $T = T_1$ and radius r at temperature T, using Dirichlet boundary conditions:

$$\frac{-Q}{4\pi} \int_{r_1}^{r} \frac{1}{r^2} dr = \lambda_\alpha \int_{T_1}^{T} dT, \tag{6.6}$$

From which the expression of the temperature profile as a function of r can be derived as:

$$T(r) = \frac{Q}{4\pi\lambda_\alpha} \left(\frac{1}{r} - \frac{1}{r_1} \right) + T_1 \;\approx\; T(r) = \frac{A}{r} + B. \tag{6.7}$$

By fitting the simulated temperature profiles to equation 6.7, it is possible to derive an estimate of the fluid and NP thermal conductivities, which follows directly from the integration constant, A, corresponding to the best fit:

$$\lambda_\alpha = \frac{Q}{4\pi A}. \tag{6.8}$$

In addition, the heat diffusion equation (equation 6.5) can be solved numerically, using a second order central finite difference method (FD)[39, 40]. In this way, the thermal boundary resistance at the interface ($r^* = R^*$) can be included to the problem resolution. When at the steady state, the Fourier's law in spherical coordinates (reduced to the one dimensional case) can be written as follows:

$$-\frac{1}{r^2}\frac{\partial}{\partial r}\left(\lambda_\alpha r^2 \frac{\partial T}{\partial r}\right) = 0,$$

and simplified to:

$$-\lambda_\alpha \frac{\partial^2 T}{\partial r^2} - \frac{2\lambda_\alpha}{r}\frac{\partial T}{\partial r} = 0. \tag{6.9}$$

Following the FD method, it is possible to perform a discretisation of the spatial domain in n nodes, using an homogeneous mesh (Δr). At node i, equation 6.9 can be reformulated as:

$$-\lambda_\alpha \frac{T_{i+1} - 2T_i + T_{i-1}}{\Delta r^2} - \frac{2\lambda_\alpha}{r_i}\frac{T_{i+1} - T_{i-1}}{2\Delta r} = 0.$$

and rearranged as:

$$T_{i+1}\left[-\frac{\lambda_\alpha}{\Delta r}\left(\frac{1}{\Delta r} + \frac{1}{r_i}\right)\right] + T_i\left[\frac{2\lambda_\alpha}{\Delta r^2}\right] + T_{i-1}\left[-\frac{\lambda_\alpha}{\Delta r}\left(\frac{1}{\Delta r} - \frac{1}{r_i}\right)\right] = 0 \tag{6.10}$$

The resulting system of n equations can be solved considering a linear system:

$$[A]\,\mathbf{T} = \mathbf{b}. \tag{6.11}$$

Where:

- $[A]\,(\lambda_{NP}, \lambda_L, G_K)$ is a tridiagonal matrix, defined according to equation 6.10, which includes: a constant term equal to $\left[2\frac{\lambda_\alpha}{\Delta r^2}\right]$ along the main diagonal of the matrix, an r-dependent term equal to $\left[-\frac{\lambda_\alpha}{\Delta r}\left(\frac{1}{\Delta r} + \frac{1}{r_i}\right)\right]$ along the first diagonal above the main one, and an r-dependent term equal to $\left[-\frac{\lambda_\alpha}{\Delta r}\left(\frac{1}{\Delta r} - \frac{1}{r_i}\right)\right]$ along the first diagonal below. The application of the FD procedure requires the modification of the row *1*, *k*, *k+1* (where k indicates the node i corresponding to the nanoparticle interface and $k+1$ the fluid one) and n to take into account the boundary conditions of the problem (see below);

- the vector $\mathbf{T}$ defines the temperature distribution along the radial distance r,
- $\mathbf{b}$ is a vector where all the elements are zero, except the first $\mathbf{b}[1]$ and last $\mathbf{b}[n]$ elements, which define the boundary conditions of the problems at $r^* = r_H$ and $r^* = r_C$, respectively (as described below).
- **Treatment of boundaries condition:** $r^* = r_H$ and $r^* = r_C$ denote the radial distances corresponding to the hot and cold thermostats. At the first node ($i = 1$ and $r^* = r_H$) and the last node ($i = n$ and $r^* = r_C$) the thermostatting process is described via *Dirichlet* boundary conditions, which results in the following conditions:

$$b(1) = T_H,\ A(1,1) = 1,\ A(1,2) = 0$$
$$b(n) = T_C,\ A(n,n) = 1,\ A(n,n-1) = 0$$

Robin boundary conditions are used to model the heat transport at the interface ($r^* = R^*$), and to take into account the temperature drop (ΔT) due to the thermal boundary conductance at the interface:

$$-\lambda \frac{\partial T}{\partial r} = G_K \Delta T. \tag{6.12}$$

This equation has to be included in the definition of the tridiagonal matrix $[A]$ in the three values of the two rows corresponding to the nanoparticle-fluid interface (here k and $k+1$):

$$\begin{aligned} A(k,k-1) &= \frac{\lambda_{NP}}{\Delta r} - G_K, \\ A(k,k) &= -\frac{\lambda_{NP}}{\Delta r}, \\ A(k,k+1) &= G_K; \end{aligned}$$

$$\begin{aligned} A(k+1,k) &= -G_K, \\ A(k+1,k+1) &= \frac{\lambda_L}{\Delta r}, \\ A(k+1,k+2) &= -\frac{\lambda_L}{\Delta r} + G_K. \end{aligned}$$

where k indicates the node i corresponding to the nanoparticle interface ($r^* = R^*$), whilst $k+1$ the fluid ones, at $r^* = R^* + \Delta r$.

The solution of the linear system 6.11 gives the temperature profile, $\mathbf{T} = [A]^{-1}\,\mathbf{b}$, at each node along the radial distance.

The transport coefficients (λ_{NP}, λ_L and G_K) were determined from this model by minimising (with a threshold of 10^{-5}) the mean squared error between the temperature profiles obtained in the MD simulations and the ones predicted by the FD scheme, as presented in the results section.

Nanoparticle wettability

The equilibrium contact angle of nanoparticles at the liquid-vapour interface was estimated via equilibrium simulations. In this set of simulations the vapour-fluid interface was firstly equilibrated at $T^* = 0.72$. The resulting coexisting densities, $\rho_l^* = 0.77$ and $\rho_v^* = 0.009$ for the liquid and vapour phase, are in good agreement with previous MD simulations of the same LJ model[33]. Then, the particle was placed in the vapour region and let equilibrate and absorb at the liquid-vapour interface. Following an equilibration of 10^5 steps, 2×10^3 steps of production were performed to compute the density profiles of the nanoparticle and the fluid using cylindrical coordinates (shown in Figure 6.4-left).

The contact angle, θ, was estimated from the density profile around the particle using the geometrical relation,

$$\cos\theta = \frac{h}{R}, \tag{6.13}$$

where h is the immersion of the particle in the liquid and R is the radius of the particle (the parameters definition is also reported in Figure 6.4-left). This approach has been previously applied for the evaluation of nanoparticles contact angle in previous works, as reviewed by Bresme and Oettel[41].

By tuning the interaction parameter, F, different wetting regimes can be reproduced, as shown in Figure 6.4-right. For $F > 0.25$, where $\varepsilon_{NP-f} > \varepsilon_f$, the nanoparticle completely immerses

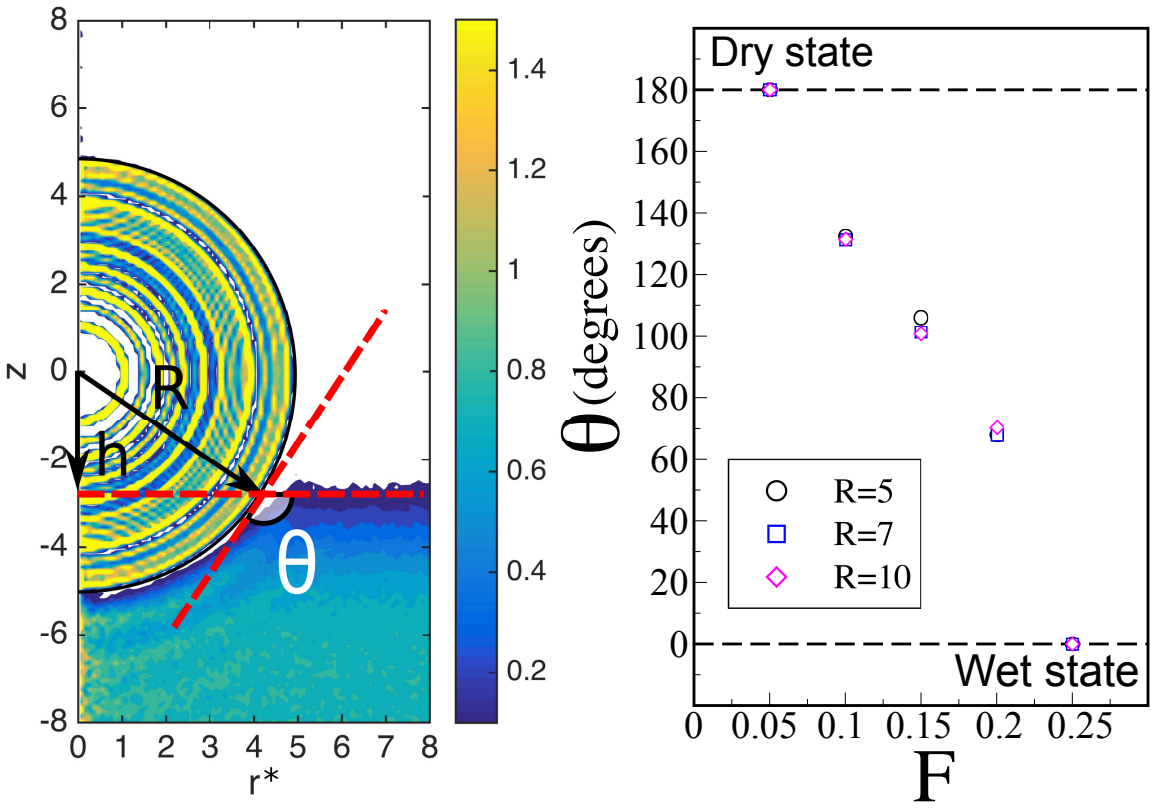


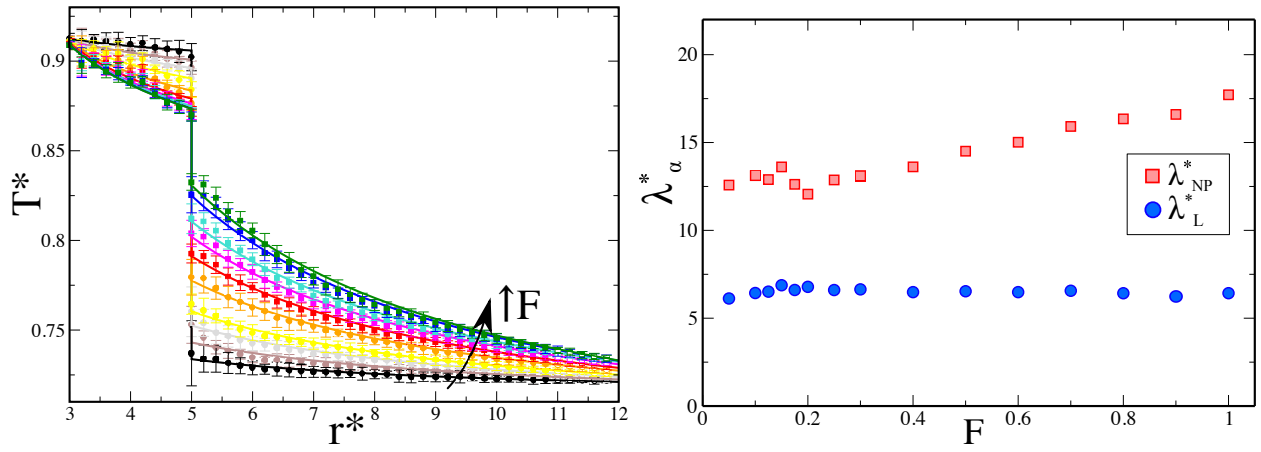
Figure 6.4: Left) Average density profile of the nanoparticle and the liquid-vapor interface. The definition of the contact angle, θ, in terms of the immersion h and particle radius, R, is also shown. The contact angle is given by $\cos\theta = h/R$. The colour map indicates different densities, ρ^* (see vertical bar for numerical values of the density). Right) Dependence of the nanoparticle contact angle with the interaction parameter F for different particle radii $R^* = 5, 7$ and 10.

itself in the fluid phase, corresponding to a full-wet regime. A full-dry state was instead reported for very hydrophobic particles, featuring $F < 0.05$. Within this range the particle was found in the partial wetting state, with the measured contact angle showing a linear relation with the parameter F, independently on particle radii. The thermal conductance is discussed below with reference to the wetting states obtained here.

6.3 Results

6.3.1 Temperature profiles around particle

Temperature profiles for a nanoparticle of radius $R^* = 5$ as a function of the fluid-particle interaction strength, F, are shown in Figure 6.5a. The temperature profiles around the nanoparticle



(a) Temperature profiles **(b)** Thermal conductivities

Figure 6.5: a) Temperature profiles as a function of the interaction strength, from $F = 0.05$ (black curve - bottom fluid region) to $F = 1.0$ (green curve - top fluid region) (see arrow in the plot and Table 6.2 for a list of the different values for F). Squares and circles represent systems with $\varepsilon_{NP} > \varepsilon_f$ and $\varepsilon_{NP} < \varepsilon_f$, respectively. The lines correspond to the solutions of the heat diffusion equation using the finite difference method. The results correspond to a nanoparticle of radius $R^* = 5$. b) Thermal conductivities of the nanoparticle (λ^*_{NP}) and the fluid (λ^*_L) evaluated using the FD approach.

features the characteristic temperature "jump", which can be used to quantify the thermal conductance of the nanoparticle-fluid interface. The magnitude of the temperature jump varies significantly with the nanoparticle-fluid interaction, F, decreasing as F increases, *i.e.* stronger fluid-nanoparticle interactions result in interfaces that are better heat conductors.

The resulting thermal gradients in the nanoparticle and fluid regions are determined by the different heat fluxes employed in the simulations and the different nanoparticle and fluid thermal conductivities. The sub-section 6.2.3 in the Methods section explores the possibility of integrating the Fourier's law to estimate the values of the thermal conductivities: the simulated temperature profiles are fitted to the solution of the one dimensional heat diffusion equation under stationary conditions[38]. Figure. 6.5a shows both the simulation (squares and circles) and the continuum model fitting (straight lines). The continuum model is satisfactory, showing that the approximation of constant thermal conductivities adopted for the nanoparticle and the fluid are reasonable for the temperature range considered here.

Numerical data for the thermal conductance and the conductivities extracted from this analysis are reported in Table 6.2 in LJ-reduced units. The same table, but in SI units is available in ref. 1. The estimates of the thermal conductance, G_K, obtained from the heat diffusion equation

	MD		FD		
F	$\mathbf{G^*_K}$	$\delta\mathbf{G^*_K}$	$\lambda^*_{\mathbf{NP}}$	$\mathbf{G^*_K}$	$\lambda^*_{\mathbf{L}}$
0.050 •	0.126	0.004	12.577	0.148	6.121
0.100 •	0.250	0.009	13.116	0.275	6.434
0.125	0.303	0.008	12.891	0.330	6.513
0.150 •	0.370	0.011	13.611	0.348	6.877
0.175	0.459	0.013	12.623	0.494	6.605
0.200 •	0.561	0.015	12.061	0.599	6.785
0.250	0.740	0.028	12.871	0.760	6.603
0.300 •	0.941	0.039	13.068	1.007	6.640
0.400 •	1.426	0.052	13.613	1.447	6.484
0.500 •	1.965	0.079	14.508	1.938	6.535
0.600 •	2.472	0.151	15.016	2.403	6.482
0.700	3.227	0.178	15.916	3.070	6.563
0.800 •	3.725	0.373	16.347	3.553	6.420
0.900	4.440	0.256	16.600	4.090	6.248
1.000 •	5.571	0.569	17.717	4.737	6.423

Table 6.2: Thermal conductances, G^*_K, and thermal conductivities of the nanoparticle and liquid, λ^*_{NP} and λ^*_L, as a function of the interaction strength, F. All the data correspond to a nanoparticle of radius $R^* = 5$. δG_K represents the uncertainty of the thermal conductance. The bullet points indicate the systems represented in Figure 6.5a and Figure 6.8a.

solution and the direct calculation of the temperature "jump" are consistent with each other. The thermal conductivities of the nanoparticle are of the order of the ones reported before for bulk solids[42, 43]. Figure 6.5b shows the thermal conductivities obtained from the FD approach as a function of the interaction parameter F, highlighting a slight increase in the NP thermal conductivity with interaction strength. It can be associated to the strong adsorption of the fluid on the nanoparticle in the wetting regime, for $F > 0.25$, as further discussed in the following sections.

In addition, the temperature profiles around the nanoparticles were also computed for two different cut-off radii ($R^*_c = 2.5$ and 4.4), for three different fluid-nanoparticle interaction strengths ($F = 0.1$, 0.5 and 1.0). Employing a different cut-off might result in slightly different coexistence properties and transport coefficients and affect the reliability of the results presented. However, Figure 6.6 shows that in this analysis the differences in the two cases are negligible. Therefore we can assume that the general physical behaviour explored in this study (where $R^*_c = 2.5$) remains the same.

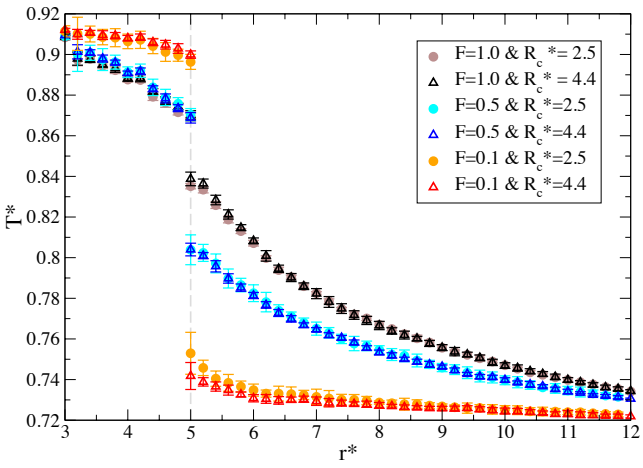


Figure 6.6: Temperature profiles of nanoparticles of radius $R^* = 5$ and different fluid-nanoparticle interaction strengths (see legend). Results for two cut-offs, 2.5 and 4.4 are shown.

6.3.2 Thermal conductance - dependence on F

The thermal conductance was computed from equation (6.4) and it is represented in Figure 6.7-left as a function of the nanoparticle-fluid interaction strength, F. The computation uses the temperatures at both sides of the temperature "jump", which is located at the particle surface, R^* (see Figure 6.5a). The "jump" is very well defined in all the systems and the thermal conductance can be determined accurately. As shown in Figure 6.7-left, G_K features a strong dependence on the interaction strength, F, without any distinctive discontinuous changes in G_K at the drying ($F = 0.05$) or wetting ($F = 0.25$) transitions. The extrapolation of the thermal conductances to $F \to 0$ indicates that the resulting interface has a very low thermal conductance, $G_K = 0.017 \pm 0.007$, which corresponds to $\sim 0.939\,\mathrm{MW\,K^{-1}\,m^{-2}}$ in SI units. This thermal conductance is similar to the values reported for liquid-vapour interfaces[44] and is consistent with the fact that at $F = 0$ the system is in the drying region and the nanoparticle would be surrounded by a vapour layer. This result supports the validity of this model in predicting thermal conductances that are consistent with existing results.

It has been previously reported by Shenogina *et al.*[19] that in simulation of self assembled monolayers in contact with water the thermal conductance varies linearly with the wetting characteristics of the surface. This dependence follows from an analysis of the work of adhesion

required to separate water from the monolayer:

$$\gamma_m + \gamma_w - \gamma_{mw} = (1 + \cos\theta)\gamma_{mw} \tag{6.14}$$

where γ_m, γ_w and γ_{mw} are the surface free energies of the monolayer-air, water-air and monolayer-water interfaces. From equation 6.14, the water contact angle, θ, can be derived from the surface free energies as $\cos\theta = (\gamma_m - \gamma_{mw})/\gamma_w$.

For nanoparticles the analysis of the correlation between the work of adhesion and the nanoparticle wetting properties is much more complex, since the free energy of nanoparticle adsorption at the liquid-vapour interface, $\mathcal{F}$, needs to be quantified. Following a previous work by Bresme and Oettel[41], for spherical particles this free energy can be written as:

$$\mathcal{F} = (\gamma_{p2} - \gamma_{p1})A_{p2} - \gamma_{12}A_{st}, \tag{6.15}$$

where $\gamma_{p\alpha}$ is the interfacial tension of the nanoparticle with phase α, either liquid or vapour, γ_{12} is the liquid-vapour surface tension, A_{p2} the area of the nanoparticle in contact with phase 2, and A_{st} the cross sectional area of interface "12" removed by the nanoparticle. It should be noted that in equation 6.15 the line tension effects was neglected. This assumption is consistent with other free energy computations of LJ nanoparticles adsorbed at liquid-vapour interfaces. These studies showed that the line tension has a small effect on the particle wettability for the nanoparticle radii considered in this work, *i.e.* $R^* \geq 5$[41, 45].

Following Aveyard and Clint[46] equation 6.15 can be re-written as:

$$\bar{\mathcal{F}} = -\frac{1}{4}(1 - \bar{h}^2) + \frac{1}{2}\cos\theta(1 - \bar{h}), \tag{6.16}$$

where $\bar{\mathcal{F}} = \mathcal{F}/(\gamma_{12}A_p)$, $A_p = 4\pi R^2$ is the total area of the spherical particle with radius R, and $\bar{h} = h/R$ quantifies the nanoparticle immersion. When the centre of mass of the particle is located at the interface $\theta = \pi/2$ and $h = 0$.

In the partial wetting regime, $0.05 < F < 0.25$, where the contact angle, θ, can be estimated

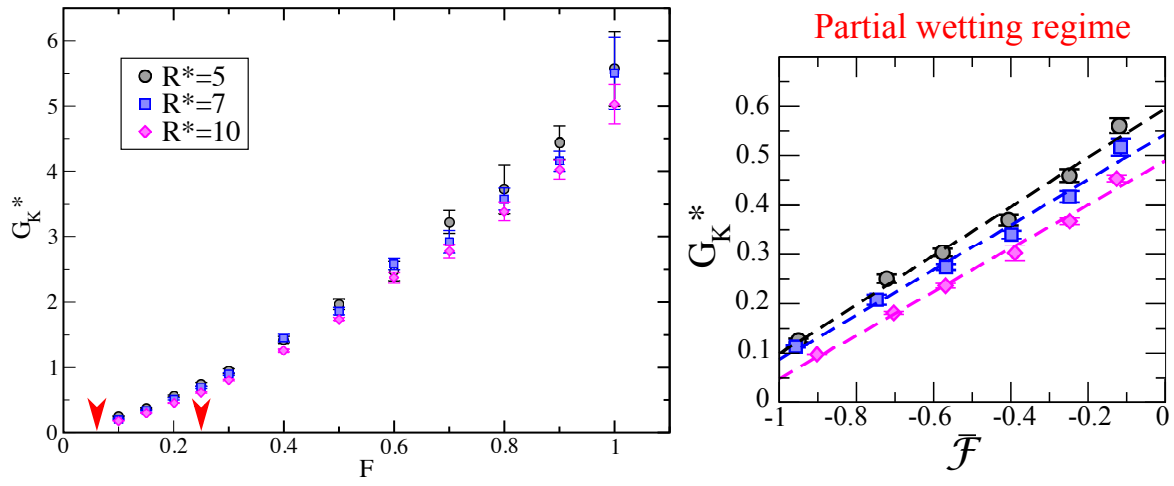
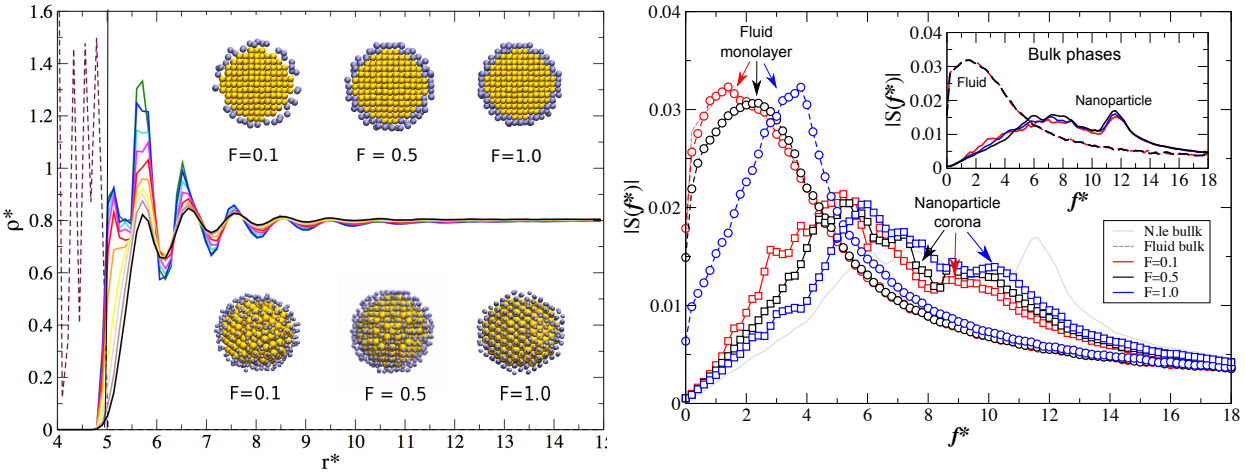


Figure 6.7: Left) Dependence of the thermal conductance with the interaction strength, F. The red arrows on the x-axis indicate the drying ($F = 0.05$) and wetting ($F = 0.25$) transitions. Right) Dependence of the thermal conductance on the free energy of adsorption, $\bar{\mathcal{F}}$, for thermodynamic states in the partial wetting regime, *i.e.* between the arrows o in the left plot. The different lines and symbols in the two plots represent the results for different particle radii, as in legend.

as discussed in the Methods section, it is possible to estimate the dependence of the thermal conductance on the free energy of adhesion of the nanoparticle with the fluid, $\bar{\mathcal{F}}$, derived from equation 6.16. This dependence is shown in Figure 6.7-left. G_K varies linearly with the nanoparticle work of adhesion for the different nanoparticle radii. The linear dependence is consistent with the trends observed in ref. 19 for liquid-monolayer interfaces, and it is remarkable that a similar trend is observed in the case of nanoparticles when the appropriate adsorption energy (see equation 6.16) is considered. A linear regression to $G_K^* = a + b\bar{\mathcal{F}}$ using all the particle sizes, gives a slope of $b \sim 0.5$. This value corresponds in real units to $28\,\mathrm{MW\,K^{-1}\,m^{-2}}$ (using Lennard-Jones units for argon, $\sigma = 0.3405\,\mathrm{nm}$, $\varepsilon = 1\,\mathrm{kJ\,mol^{-1}}$ and $m = 39.94\,\mathrm{g\,mol^{-1}}$). The estimated slope for self assembled monolayers in contact with water is $85\,\mathrm{MW\,K^{-1}\,m^{-2}}$ [19]. These results show the difficulty in mapping different substrates (surfaces, nanoparticles) and fluids to a general master curve that would relate the thermal conductance to the work of adhesion. Indeed, the data collected in this work reveals deviations across systems, which correlate with nanoparticle size. These deviations are connected to the underlying dependence of the thermal conductance with curvature, as discussed in the following section.

Nanoparticles in the full wetting regime, $F > 0.25$, feature a strong dependence of the thermal conductance on nanoparticle-solvent interaction strength. Such a dependence is consistent

with experimental measurements of the thermal conductance of alkane monolayer-water planar interfaces[47], which shown that the thermal conductance increases with the hydrophilicity, *i.e.*, the strength of the monolayer-water interactions. Smaller internal energy transfer rates are expected when the liquid-solid interaction is weak, a notion that has been confirmed in several simulation studies [19, 48, 49, 50]. The data reported here for the thermal conductance of the high solvent-nanoparticle interaction regime show that a similar principle applies to nanoparticle-fluid interfaces. Indeed, beyond the wetting transition, $F > 0.25$, the thermal conductance increases significantly, by about 5-6 times for $F \sim 1$, with values corresponding to the high end of thermal conductances ($\sim 250\,\mathrm{MW\,K^{-1}\,m^{-2}}$). These values are of the order of the ones found in experiments of AuPd nanoparticles ($\sim 4 - 22$ nm diameter)[51], which in our units would correspond to radii $R^* \sim 5 - 30$.



(a) Density profiles - F effect **(b)** DoS

Figure 6.8: (a) Density profiles of the solvent around a nanoparticle of radius $R^* = 5$ (the vertical line indicates the particle radius) for partial wetting and wetting states. The different lines correspond to different solvent-particle interaction strengths. Focusing on the main peak $r^* \sim 6$ the lines correspond to, from bottom to top, $F = 0.05$, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1. The snapshots show the structuring of the fluid around the nanoparticle for three representative interaction strengths. (b) Density of states for nanoparticles with nanoparticle-fluid interactions corresponding to a partial wet state ($F = 0.1$) and to wet states ($F = 0.5, 1$). The main panel shows the contributions from the atoms on the surface of the nanoparticle (squares), corresponding to the nanoparticle corona, and the fluid monolayer (circles) in direct contact with the nanoparticle. The "||" symbol refers to the amplitude of the Fourier transform and f^* is the frequency in Lennard-Jones units. The inset shows the contribution from the remaining particles, which correspond to the "bulk" nanoparticle and "bulk" fluid. The bulk contribution for the nanoparticle (grey line and area) and the fluid (brown dashed line and area) are also shown in the main panel ("N.le" and "Fluid"), to facilitate the comparison with the interfacial contributions.

In the full wetting regime, the correlation with the contact angle employed to construct Figure 6.7-right can not be used. In order to shed light on the large increase of the thermal conductance with the solid-liquid interactions in this regime, the fluid density profiles around the nanoparticles, ρ^*, were computed. Figure 6.8a shows the behaviour of ρ^* for different interaction strengths (from $F = 0.05$ to $F = 1$), in the case of a nanoparticle of radius $R^* = 5$. In addition, representative snapshots of the nanoparticle with the surrounding fluid atoms (until $R^* \sim 6$) are shown as insets in the Figure for a partial wetting situation ($F = 0.1$) and two wetting states ($F = 0.5$ and 1). A first large difference could be noticed in the density profiles between the partial wetting and the full wetting regimes ($F > 25$). At low interaction strengths the fluid develops a depletion layer around the nanoparticle and there is little evidence for adsorption on the nanoparticle surface. At high interaction strengths, $F = 0.5$ and 1, the fluid develops a strong structure around the nanoparticle. The layering induced by the nanoparticle extends to long distances inside the bulk fluid, $r^* \sim 10$. The density profiles for highly interacting particles feature a double adsorption peak ($r^* \sim 5$, 5.7) which is connected to the underlying faceted structure of our nanoparticle, which results in two characteristic radial distances for adsorption. The intensity of these peaks increases with F. The fluid particles adsorb strongly on the nanoparticle surface, which acts as a template for the adsorbed layer (see structure of the monolayer in the simulation snapshots). Heat transfer across this interface is more favourable and this explains the strong increase of the thermal conductance with the nanoparticle-fluid interaction strength.

The structural changes discussed above must result in a distinctive modification of the dynamic properties of the fluid adsorbed on the nanoparticle surface and the nanoparticle atoms in direct contact with this fluid. To check this hypothesis, the vibrational density of states (DoS) of NP and fluid atoms was computed. The DoS has been considered before in theories of interfacial conductance by Swartz and Pohl[52], which focused on the density of states of the two bulk phases in contact. Computer simulation offers the possibility of splitting the interfacial and bulk contributions to the density of states. The density of states can be obtained from the power spectra of the velocity autocorrelation function (VAF):

$$\tilde{S}(\omega) = \int_0^{t_{max}} \left[\frac{\mathrm{VAF}(t)}{\mathrm{VAF}(0)}\right] \exp(i\omega t)\ \mathrm{d}t, \tag{6.17}$$

where t_{max} is the maximum time used to compute the velocity autocorrelation function, which is obtained from

$$\mathrm{VAF}(t) = \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} \langle \mathbf{v}_i(t) \cdot \mathbf{v}_i(0) \rangle, \tag{6.18}$$

where N_α is the number of particles of type α, *i.e.*, nanoparticle, nanoparticle corona, bulk fluid or interfacial fluid. To compute the interfacial contributions to the VAF, two region were defined:

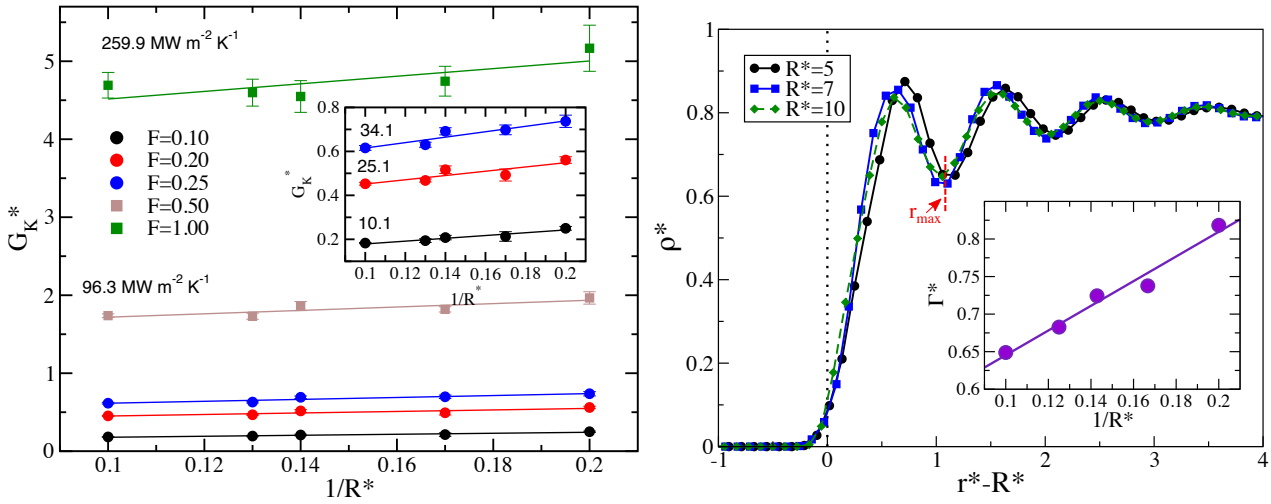
- interfacial fluid monolayer, which includes all the atoms inside a spherical shell of radius, $r^* = 6.2$, around the nanoparticle;
- nanoparticle corona, which includes all the nanoparticle atoms outside a spherical shell of radius $r^* = 4.5$.

The maximum correlation time for the velocity autocorrelation function was $t^*_{max} = 5$ (corresponding to 5000 time steps). Only those particles that remained in the shells defined above, during the entire sampling time, *i.e.* from $t_0 = 0$ to $t_{max} = 5$, where included in the analysis of the interfacial contributions. Particles outside these two layers were employed to compute the "bulk" contributions for the nanoparticle and the fluid.

The power spectrum was computed by discretising the velocity autocorrelation function in the frequency domain, using the MATLAB library *fft*. The sampling frequency was $F^*_S = 1/t^* = 1000$ (in reduced units). Figure 6.8b shows the DoS of the "bulk" and interfacial contributions. The DoS of the nanoparticle shows a structure similar to the phonon density of states of extended three dimensional solids[53]. The DoS of the "bulk" phases, nanoparticle and fluid, are independent of the fluid-particle interaction, hence it is not possible to establish a direct correlation between these DoS and the thermal conductance. However the interfacial contribution to the DoS features a significant dependence, for both the surface atoms of the

nanoparticle and the interfacial fluid. The DoS of the fluid shifts to higher frequencies, leading to a better overlap of the DoS of the nanoparticle and fluid atoms. This observation is consistent with an enhancement of the thermal conductance that is connected to a better vibrational coupling of the interfacial atoms.

6.3.3 Thermal conductance - dependence on curvature and coupling effect



(a) Dependence on curvature **(b)** Density profiles - $1/R^*$ effect

Figure 6.9: a) Thermal conductance of the nanoparticles as a function of the particle curvature, $1/R^*$, and interactions strength, F. The numbers indicate the thermal conductance for $1/R^* = 0.1$ in SI units, using $\sigma = 3.405$ Å and $\varepsilon_f = 0.996\,\mathrm{kJ\,mol^{-1}}$ as conversion factors for the effective diameter and interaction strength of the fluid. The inset shows the dependence of the thermal conductance in the partial-dry/wet states, $0.1 < F < 0.25$. b) Density profiles of the fluid around nanoparticles of different radii and interaction strength $F = 0.1$. The inset shows the dependence of the adsorption, Γ^* (see text and equation 6.20 for details) with the curvature of the nanoparticle, $1/R^*$. The error associated to the data is of the order of the symbol size. r_{max} indicates the integration limit used in equation 6.20.

So far, the analysis presented in this chapter has been focused on the impact of the nanoparticle-fluid interactions on the interfacial thermal conductance. In this section instead, the dependence of the thermal conductance with particle size is discussed, addressing the impact of the interfacial curvature on the thermal conductance.

A proof of the existence of a relationship between interfacial thermal conductance and curvature was reported for the first time in the MD simulation study by Lervik *et al.*[20], where the

interfacial conductance of an oil-water interface was found to exhibit a strong dependence on the oil droplet radius. This question has been also addressed on oxide-alkane interfaces by Hu *et al.*[21] and on droplets in coexistence with vapour[23].

The relationship between G_K and curvature found in this analysis is presented in Figure 6.9a, where the thermal conductance is represented as a function of the particle curvature, $1/R^*$. The different curves correspond to different fluid-particle interaction strengths. The dependence of the thermal conductance of the nanoparticles can be fitted to a linear equation:

$$G_K\left(R\right) = G_K(0) + \frac{\delta}{R}, \tag{6.19}$$

where $G_K(0)$ is the thermal conductance of the corresponding flat surface, and δ quantifies the change of the thermal conductance with particle curvature. δ features positive values, irrespectively of the interaction strength, with all the thermal conductances increasing with increasing curvature. It follows that curvature enhances the thermal conductance, with highly curved interfaces being better thermal conductors. To better understand this result, the adsorption of fluid on the nanoparticle has been computed as a function of the particle size. The adsorption corresponding to the first fluid layer was estimated from:

$$\Gamma^* = N_l/A^* = R^{*-2}\int_0^{r_{max}} r^2\rho^*(r)dr, \tag{6.20}$$

where N_l is the number of fluid beads around the nanoparticle, A^* represents the area of the nanoparticle of radius R^* and r_{max} is the distance at the first minimum in the density profile, $\rho^*(r)$. Figure. 6.9b, shows the density profiles for $F = 0.1$ at several nanoparticle radii, $R^* = 5, 7, 10$. To better compare the results for different particle sizes, all the profiles have been translated of $-R^*$. The position of the first minimum is indicated in figure with a red label (r_{max}). Interestingly, the adsorption, evaluated from equation 6.20, decreases with decreasing curvature, approximately in a linear fashion, as reported in the inset of Figure 6.9b. This result shows that the curvature dependence of the thermal conductance is correlated with the adsorption, or in other words, with the surface density of the fluid layer in direct contact with the nanoparticle. This result thus provides a microscopic explanation of the curvature

dependence in terms of the fluid-nanoparticle correlations.

In addition, the variation of the thermal conductance with curvature, quantified by the parameter δ defined in equation 6.19, depends on the characteristic fluid-nanoparticle interactions. Wet states (stronger interactions) result in a stronger dependence than in partially dry states (see δ value in Table 6.3). Following the linear dependence in equation (6.19), it is possible to predict the thermal conductance of a flat surface from the thermal conductances of particles of varying sizes. The extrapolated conductances, $G_K(0)$, are also reported in Table 6.3. The consistency of these predictions was tested by performing additional computations of a fluid in contact with a flat surfaces of varying interaction strengths, F. The thermal conductance, $G_k(0)$, was computed again from the temperature jump at the interface and the heat flux imposed in the non-equilibrium simulation. The results for the flat surfaces show very good agreement with the ones extrapolated from equation (6.19) (see Table 6.3), and confirm the validity of this equation.

F	$G_K^*(0)$	$G_K(0)$ (MW K^{-1} m^{-2})	δ (W K^{-1} m^{-1})	$G_{K,fs}$ (MW K^{-1} m^{-2})
0.10	0.114±0.025	6.25±1.12	0.0121±0.0024	9.33±0.48
0.20	0.370±0.069	20.01±3.39	0.0173±0.0072	
0.25	0.499±0.039	27.68±1.75	0.0225±0.0037	28.3±1.4
0.50	1.483±0.118	81.84±5.53	0.0450±0.0116	80.3±11.2
1.00	3.916±0.366	218.1±19.9	0.1048±0.0427	216.1±47.8

Table 6.3: Dependence of the thermal conductance with the nanoparticle-fluid interaction strength, F. $G_K(0)$ is the thermal conductance for a flat surface ($1/R = 0$) obtained from the extrapolation of the linear regression represented in Figure 6.9a and equation 6.19. The quantities with the $*$ denote reduced units. $\delta = dG_k/d(1/R)$ is the slope of the linear regressions represented in Figure 6.9a, and $G_{K,fs}$ is the thermal conductance obtained from computer simulations of flat surfaces in contact with the fluid.

Thermal conductance measurements in experimental studies of gold nanoparticles reveal large variations depending on the chemical composition and the solvent employed. For example, the thermal conductance of alkanethiol coated AuPd (3–5 nm diameter)-toluene interfaces is low, $\sim 15\,\mathrm{MW\,K^{-1}\,m^{-2}}$, while in water is one order of magnitude higher[17, 51], $250 \pm 90\,\mathrm{MW\,K^{-1}\,m^{-2}}$ (3–5 nm), $145 \pm 55\,\mathrm{MW\,K^{-1}\,m^{-2}}$ (3–5 nm) and $230 \pm 50\,\mathrm{MW\,K^{-1}\,m^{-2}}$ (22 nm), most likely connected to the hydrophilic character of the ligands. The values reported in our work cover this wide range of thermal conductances. For high G_K values, $\sim 200\,\mathrm{MW\,K^{-1}\,m^{-2}}$, our model predicts that the thermal conductance would increase by $\sim 25\,\mathrm{MW\,K^{-1}\,m^{-2}}$ in decreasing

the particle diameter from 22 nm to 3 nm. The analysis of the available experimental data highlights that a low sensitivity and precision in the reported data (see uncertainties in the thermal conductances quoted above), which makes difficult to resolve the dependence with particle size with enough precision. Nonetheless, some of the experimental data for nanoparticles can be compared with existing data for flat surfaces, where, according to our model, differences in thermal conductances are expected to be larger. The thermal conductance of AuPd particles stabilized with thioalkylated ethylene glycol is $250\pm90\,\mathrm{MW\,K^{-1}\,m^{-2}}$ [51]. Thermoreflectance experiments of polyethylene glycol silane monolayers on silver have shown that the thermal conductance is of the order of $180 \pm 30\,\mathrm{MW\,K^{-1}\,m^{-2}}$ [14]. Assuming the nanoparticle and flat thermal conductances for these systems are comparable, given the similar chemical composition of the ligands, one could estimate the curvature dependence of the thermal conductance. Using the experimental values for the flat interface and the nanoparticle (assuming $\sim$ 2 nm radius) coefficient determining the curvature dependence, $\delta = dG_K/d(1/R)$, measures $\sim 0.15\,\mathrm{MW\,K^{-1}\,m^{-1}}$. This result is of the same order as the one we obtain for nanoparticles with large thermal conductances (see δ values in Table 6.3).

Additional predictions of the thermal conductance can be found in computer simulation studies of nanoparticles. Recent simulations of small nanodroplets immersed in water predicted conductances of the order of $G_K \sim 100\,\mathrm{MW\,K^{-1}\,m^{-2}}$ for $1/R \sim 0.1$ nm^{-1} [20], which in our model correspond to the values for a medium interaction strength between the fluid and the particle. The coefficient determining the curvature dependence can be estimated from the simulation data to be of the order of $\delta \sim 0.2$ or $0.1\,\mathrm{MW\,K^{-1}\,m^{-1}}$ depending on the model employed in the analysis of the simulation data (see models A and C in reference [20]). Computer simulations of oxide zinc nanoparticles ($\sim$ 2 nm size), immersed in an alkane solvent predicted low thermal conductances, $\sim 25\,\mathrm{MW\,K^{-1}\,m^{-2}}$ for $1/R \sim 0.15\,\mathrm{nm}$ [21]. The curvature dependence that can be inferred from analysis of the results reported in ref. 21 is consistently lower, $\delta \sim 0.05\,\mathrm{MW\,K^{-1}\,m^{-1}}$. These results for the curvature dependence are of the order predicted by our model.

The available simulation and experimental data, described above, have been collected in Figure 6.10. The figure shows the dependence of the curvature dependence, δ, with the thermal

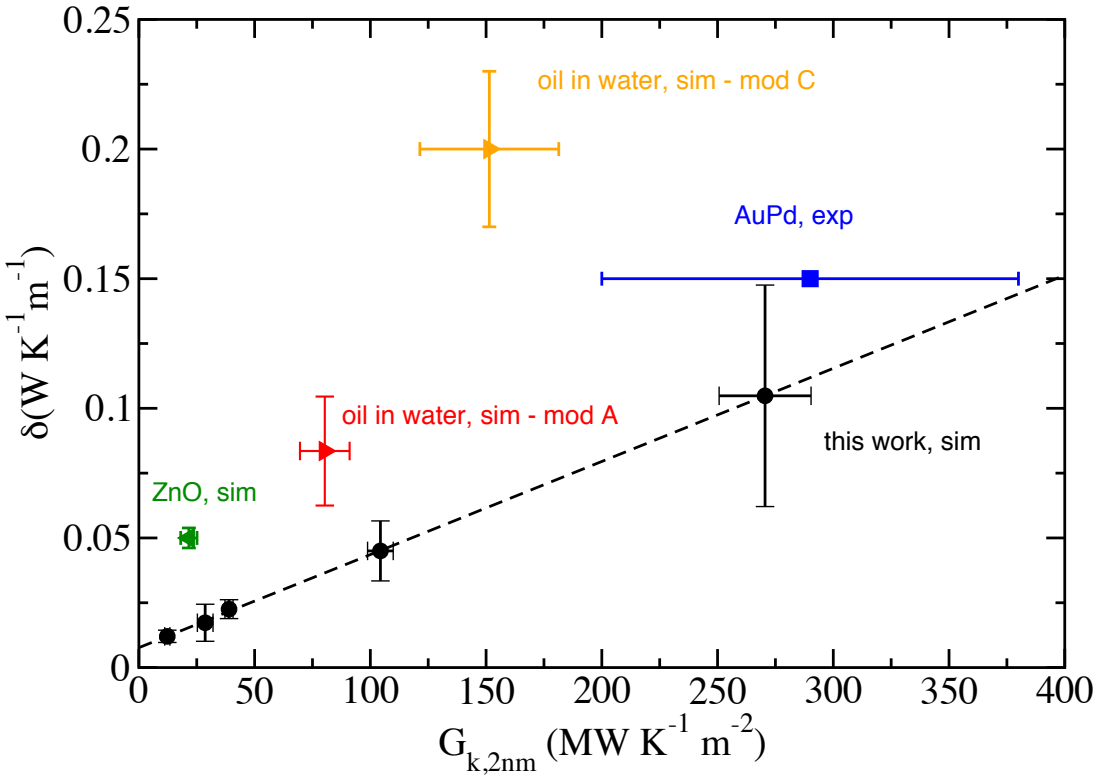


Figure 6.10: Dependence of the parameter δ on the thermal conductance for nanoparticles of radius 2 nm. The circles and dashed line are results obtained with the model introduced in this work. The other symbols have the following meaning. Red right triangle: simulation results for oil nanodroplets in water (taken from ref. 20, "mod A" and "mod C" refer to the different models employed in that work). Green left triangle: simulation results for ZnO nanoparticles in tetradecane (from ref. 21). Blue square: thioalkylated ethylene glycol-stabilized AuPd nanoparticles in water (from ref. 51).

conductance, using as a reference thermal conductances of 2 nm radius nanoparticles, available for both experimental and simulation estimates. It must be noted that the current experimental estimates involve significant uncertainties. The general trend, however, for both simulation and experiments is consistent with the predictions of our model. These data support the main conclusion of this chapter that higher thermal conductances involve stronger variations of the conductance with curvature.

6.4 Conclusions

Using MD simulations and a generic model of a nanoparticle immersed in a fluid, the dependence of the thermal conductance with both nanoparticle-fluid interaction strength and nanoparticle size (*i.e.* interfacial curvature) has been investigated and quantified. This analysis reveals a

subtle interplay of nanoparticle size and interfacial interactions that defines the thermal conductance of nanoparticle interfaces over a wide range of length scales, from small nanoparticles, 1 nm, to the flat interface limit.

For interaction strengths corresponding to the nanoparticle-fluid partial wetting regime there is a linear correlation between the thermal conductance and the free energy of adsorption of the nanoparticle. This result is of practical significance, since it connects the thermal conductance of nanoscale interfaces with an experimentally measurable property, the contact angle. Now the contact angles of truly small nanoparticles (few nanometers diameter)[31] can be measured using neutron probes, providing a route to estimate adsorption free energies. These recent advances may provide an experimental approach to verify the correlations presented in this chapter. Beyond the partial wetting regime, in the full wetting region, the thermal conductance increases with the fluid-nanoparticle interaction strength. This observation can be understood in terms of the enhanced adsorption of the fluid on the nanoparticle surface, and the increased overlap between the vibrational density of states of the fluid layer adsorbed at the nanoparticle and the density of states of the nanoparticle surface.

The thermal conductance increases with interfacial curvature for a wide range of fluid-nanoparticle interaction conditions. Small nanoparticles are better heat conductors than larger ones. The dependence of the thermal conductance on curvature can be correlated to the change in fluid adsorption with particle radius. For the same interaction strength, in fact, small particles feature a stronger adsorption, namely, the surface density of the fluid layer in direct contact with the substrate is higher. This notion provides a microscopic explanation of the observed curvature dependence. In addition, the curvature dependence of the thermal conductance can be fitted to a linear equation, $G_K(1/R) = G_K(0) + \delta/R$, which depends on the nanoparticle curvature, $1/R$, the thermal conductance in the limit of a flat surface, $G_K(0)$, and the parameter δ, which measures the "strength" of the curvature dependence. It has been found that strong interactions (resulting in large thermal conductances) lead to larger values of δ. These predictions has been tested against experiments of AuPd nanoparticles and simulations of oxide nanoparticles, which follow the main predictions of our model.

Most probably the analysis proposed in this chapter will motivate new experiments and simulations. The synergy between these two approaches should play an important role to quantify thermal conductances at the nanoscale and to address, experimentally, the dependence of the thermal conductance on particle size and composition. The connection of thermal conductance and nanoparticle wettability/adsorption may provide new routes to design experiments that exploit this correlation, and that may help to reduce the large level of uncertainty in the measurements obtained with current methods, based on optical absorption. Ultimately, a better understanding of heat transport at the nanoscale and its dependence on nanoparticle curvature will help to advance the use of nanoparticles in technological and medical applications.

6.5 Future work: nanoparticles interacting with SC-lipid bilayers

Previous experimental studies [54, 55, 56] have investigated the capability of nanoparticle to cross the dermal barrier. Furthermore computational studies (*e.g.* ref 57) attempted to understand the mechanisms by which gold nanoparticles (1–6 nm in diameter) penetrate and disrupt the skin lipid matrix. The relevance of studying the heat transferred from a heated particle to a surrounding skin lipid bilayer stems from the possible use of metallic nanoparticle to fight skin cancer[58, 59]. Metallic nanoparticle embedded within tumors can generate localised heat in response to laser light. Understanding the mechanism of heat transfer at the nanoscale is crucial to estimate the efficiency of photothermal therapy. In addition TiO_2 and ZnO nanoparticle are currently added to sunscreen creams formulation. A study of how heat propagates to the surrounding SC lipids is still missing and will be the object of our future works.

The study proposed in this chapter represents a first step in understanding the thermal properties of therapeutic nanoparticle and how they exchange heat with surrounding media. This work provides a solid foundations for studying the impact of a heated nanoparticle in the proximity of a stratum corneum (SC) lipid membrane. The initial configuration of the future set of simulations is shown in Figure 6.11.

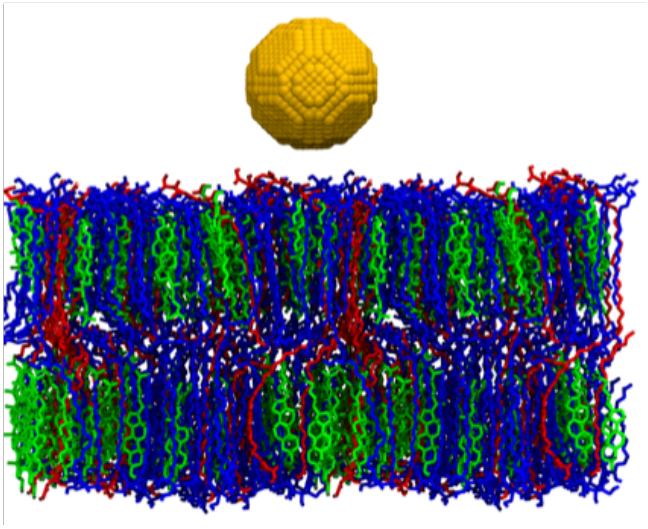


Figure 6.11: Initial configuration for the study of the interaction of nanoparticles and SC lipids

Bibliography to chapter 6

[1] A. S. Tascini, J. Armstrong, E. Chiavazzo, M. Fasano, P. Asinari and F. Bresme, *Phys. Chem. Chem. Phys.*, 2017, **51**, 13725–13730.

[2] W. Lin, *Chemical Reviews*, 2015, **115**, 10407–10409.

[3] J. B. Edel, A. A. Kornyshev, A. R. Kucernak and M. Urbakh, *Chem. Soc. Rev.*, 2016, **45**, 1581–1596.

[4] X.-M. Lin, Y. Sun, E. V. Shevchenko, S. K. R. S. Sankaranarayanan, D. John, I. Fedin, F. Bresme, H. Mohwald, P. Moriarty, C. M. Sorensen and B. M. Law, *Chem. Commun.*, 2015, **51**, 13725–13730.

[5] J. D. Wilton-Ely, *Dalton Trans.*, 2008, 25–29.

[6] S. M. Rogers, N. Dimitratos, W. Jones, M. Bowker, A. G. Kanaras, P. P. Wells, C. R. A. Catlow and S. F. Parker, *Phys. Chem. Chem. Phys.*, 2016, **18**, 17265–17271.

[7] R. W. Cheyne, T. A. Smith, L. Trembleau and A. C. Mclaughlin, *Nanoscale Research Letters*, 2011, **6**, 423.

[8] K. J. Si, Y. Chen, Q. Shi and W. Cheng, *Advanced Science*, 2018, **5**, year.

[9] Z. Jiang, J. He, S. A. Deshmukh, P. Kanjanaboos, G. Kamath, Y. Wang, S. K. R. S. Sankaranarayanan, J. Wang, H. M. Jaeger and X.-M. Lin, *Nat. Mater.*, 2015, **14**, 912.

[10] J. E. Reiner, J. W. F. Robertson, D. L. Burden, L. K. Burden, A. Balijepalli and J. J. Kasianowicz, *J. Am. Chem. Soc.*, 2013, **135**, 3087–3094.

[11] C. R. Crick, P. Albella, B. Ng, A. P. Ivanov, T. Roschuk, M. P. Cecchini, F. Bresme, S. A. Maier and J. B. Edel, *Nano Lett.*, 2015, **15**, 553–559.

[12] V. Kotaidis, C. Dahmen, G. von Plessen, F. Springer and A. Plech, *J. Chem. Phys.*, 2006, **124**, 184702.

[13] D. Jaque, L. Martinez Maestro, B. del Rosal, P. Haro-Gonzalez, A. Benayas, J. L. Plaza, E. Martin Rodriguez and J. Garcia Sole, *Nanoscale*, 2014, **6**, 9494–9530.

[14] D. G. Cahill, P. V. Braun, G. Chen, D. R. Clarke, S. Fan, K. E. Goodson, P. Keblinski, W. P. King, G. D. Mahan, A. Majumdar, H. J. Maris, S. R. Phillpot, E. Pop and L. Shi, *App. Phys. Rev.*, 2014, **1**, 11305.

[15] M. Hu, X. Wang, G. V. Hartland, V. Slagueirino-Maceira and L. M. Liz-Marzan, *Chem. Phys. Lett.*, 2003, **372**, 767–772.

[16] E. Chiavazzo, M. Fasano, P. Asinari and P. Dacuzzi, *Nature Comm.*, 2014, **5**, 3565.

[17] O. M. Wilson, X. Hu, D. G. Cahill and P. V. Braun, *Phys. Rev. B*, 2002, **66**, 224301.

[18] H. A. Patel, S. Garde and P. Keblinski, *Nano Lett.*, 2005, **5**, 2225–2231.

[19] N. Shenogina, R. Godawat, P. Keblinski and S. Garde, *Physical Review Letters*, 2009, **102**, 15113–15118.

[20] A. Lervik, F. Bresme and S. Kjelstrup, *Soft Matter*, 2009, **5**, 2407.

[21] M. Hu, D. Poulikakos, C. P. Grigoropoulos and H. Pan, *Journal of Chemical Physics*, 2010, **132**, 164504.

[22] O. Wilhelmsen, D. Bedeaux, S. Kjelstrup and D. Reguera, *J. Chem. Phys.*, 2014, **140**, 24704.

[23] O. Wilhelmsen, D. Bedeaux and S. Kjelstrup, *Phys. Chem. Chem. Phys.*, 2014, **16**, 10573–10586.

[24] L. Xue, P. Keblinski, S. R. Phillpot, S. U.-S. Choi and J. A. Eastman, *J. Chem. Phys.*, 2003, **118**, 337–339.

[25] K. M. Stocker, S. M. Neidhart and J. D. Gezelter, *J. App. Phys.*, 2016, **119**, 25106.

[26] S. Merabia and K. Termentzidis, *Phys. Rev. B*, 2012, **86**, 94303.

[27] K. Issa and A. Mohamad, *Phys. Rev. E*, 2012, **85**, 31602.

[28] S. Merabia, S. Shenogin, L. Joly, P. Keblinski and J.-L. Barrat, *Proc. Natl. Acad. Sci.*

USA, 2009, **106**, 15113–15118.

[29] E. A. Mastny and J. J. de Pablo, *J. Chem. Phys.*, 2007, **127**, 104504.

[30] H. Heinz, R. A. Vaia, B. L. Farmer and R. R. Naik, *Journal of Physical Chemistry C*, 2008, **112**, 17281–17290.

[31] J. Reguera, E. Ponomarev, T. Geue, F. Stellacci, F. Bresme and M. Moglianetti, *Nanoscale*, 2015, **7**, 5665–5673.

[32] S. Plimpton, *Sandia National Laboratories*, 2007.

[33] A. Trokhymchuk and J. Alejandre, *Journal of Chemical Physics*, 1999, **111**, 8510–8523.

[34] J. H. Irving and J. G. Kirkwood, *The Journal of Chemical Physics*, 1950, **18**, 817.

[35] B. Hafskjold, T. Ikeshoji and S. Ratkje, *Molecular Physics*, 1993, **80**, 1389–1412.

[36] F. Bresme, B. Hafskjold and I. Wold, *The Journal of Physical Chemistry*, 1996, **100**, 1879–1888.

[37] J. Muscatello and F. Bresme, *Journal of Chemical Physics*, 2011, **135**, year.

[38] F. P. Incropera, D. P. DeWitt, T. L. Bergman and A. S. Lavine, *heat and mass transfer - Incropera 6e*, 2007.

[39] C. Johnson, *Numerical solution of partial differential equations by the finite element method*, 1987, vol. 32, p. 278.

[40] J. Peiro and S. Sherwin, *Handbook of Materials Modeling*, Springer Netherlands, Dordrecht, 2005, pp. 2415–2446.

[41] F. Bresme and M. Oettel, *Journal of Physics Condensed Matter*, 2007, **19**, year.

[42] K. V. Tretiakov and S. Scandolo, *The Journal of Chemical Physics*, 2004, **120**, 3765–3769.

[43] H. Kaburaki;, J. Li;, S. Yip, H. Kaburaki, J. Li and S. Yip, *Mater. Res. Soc. Symp. Proc.*, 1999, **538**, 503–508.

[44] J. M. Simon, S. Kjelstrup, D. Bedeaux and B. Hafskjold, *Journal of Physical Chemistry B*, 2004, **108**, 7186–7195.

[45] F. Bresme and N. Quirke, *Physical Review Letters*, 1998, **80**, 3791–3794.

[46] R. Aveyard and J. H. Clint, *Journal of the Chemical Society, Faraday Transactions*, 1996, **92**, 85.

[47] Z. Ge, D. G. Cahill and P. V. Braun, *Physical Review Letters*, 2006, **96**, year.

[48] J.-L. Barrat and F. Chiaruttini, *Molecular Physics*, 2003, **10111**, 1605–1610.

[49] L. Xue, P. Keblinski, S. R. Phillpot, S. U. S. Choi and J. A. Eastman, *Journal of Chemical Physics*, 2003, **118**, 337–339.

[50] A. Giri and P. E. Hopkins, *Applied Physics Letters*, 2014, **105**, year.

[51] Z. Ge, D. G. Cahill and P. V. Braun, *Journal of Physical Chemistry B*, 2004, **108**, 18870–18875.

[52] E. T. Swartz and R. O. Pohl, *Reviews of Modern Physics*, 1989, **61**, 605–668.

[53] P. Heino, *Phys. Rev. B*, 2005, **71**, 144302.

[54] H. I. Labouta, L. K. El-Khordagui, T. Kraus and M. Schneider, *Nanoscale*, 2011, **3**, 4989.

[55] M. Schneider, F. Stracke, S. Hansen and U. F. Schaefer, *Dermato-Endocrinology*, 2009, **1**, 197–206.

[56] J. Lademann, H. Richter, A. Teichmann, N. Otberg, U. Blume-Peytavi, J. Luengo, B. Weiß, U. F. Schaefer, C. M. Lehr, R. Wepf and W. Sterry, *European Journal of Pharmaceutics and Biopharmaceutics*, 2007, **66**, 159–164.

[57] R. Gupta and B. Rai, *Journal of Physical Chemistry B*, 2016, **120**, 7133–7142.

[58] I. Salas-García, F. Fanjul-Vélez, N. Ortega-Quijano, A. Lavín-Castanedo, P. Mingo-Ortega, M. López-Escobar and J. L. Arce-Diego, Proc. SPIE 8092, Medical Laser Applications and Laser-Tissue Interactions V, 2011, pp. 809204–809204.

[59] I. Salas-García, F. Fanjul-Vélez, N. Ortega-Quijano, J. Poliak, L. Hudcová, O. Wilfert and J. L. Arce-Diego, Proceedings of 22nd International Conference Radioelektronika 2012, 2012, pp. 1–4.

Chapter 7

Antimicrobial peptides

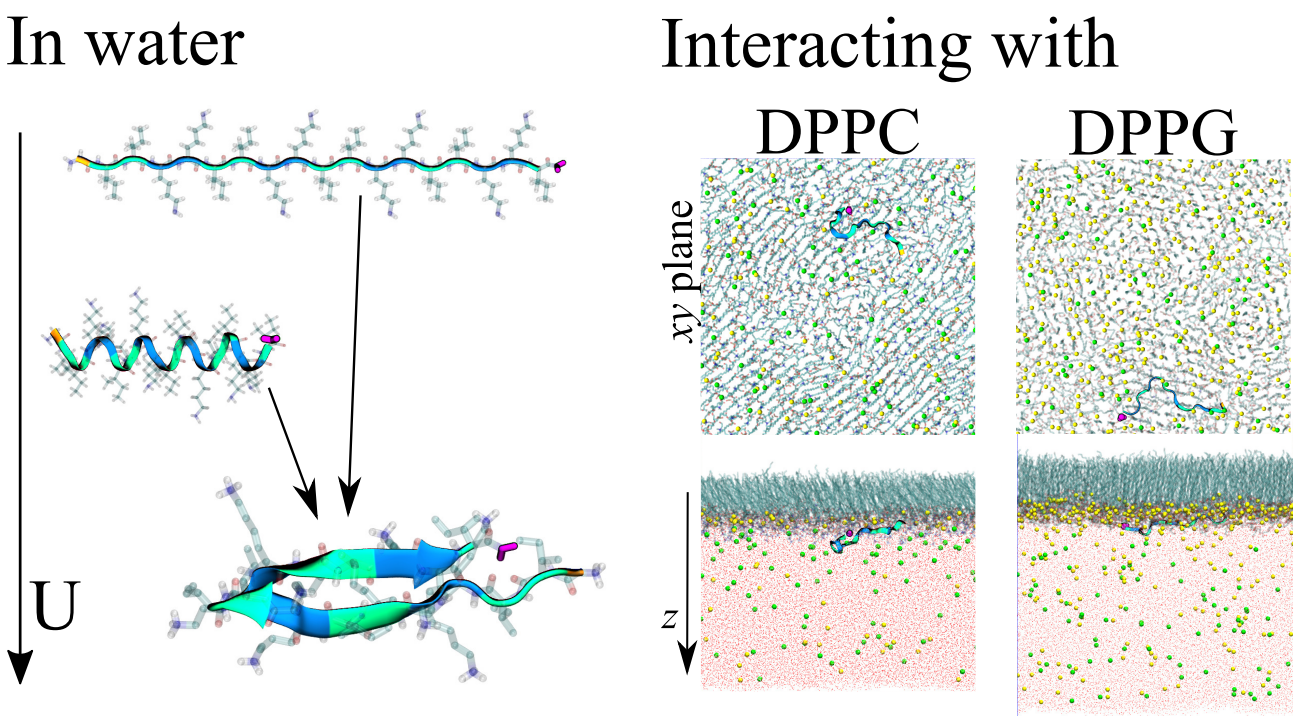


Figure 7.1: Title page figure for chapter 7.

7.1 Introduction

Natural antimicrobial peptides (AMPs) are promising players against bacterial infections. They are produced by many tissue and cell types in a variety of invertebrates, plants and animal species, and represent an alternative to conventional antibacterial drugs. The problem with conventional antibacterial drugs is that their mechanisms of action can be easily defeated by bacteria mutations in chromosomal genes and by horizontal gene transfer[1, 2]. In the last few years, the number of antibiotic-resistant bacteria has dramatically increased[3], causing a global health crisis due to the lack of available alternatives to treat infections. Data reported by the European Centre for Disease Prevention and Control (ECDC)[4] reported a high percentages of bacteria with resistance to key antimicrobial groups, which makes antimicrobial resistance a serious threat to public health in Europe, especially in southern and south-eastern areas (see Figure 7.2).

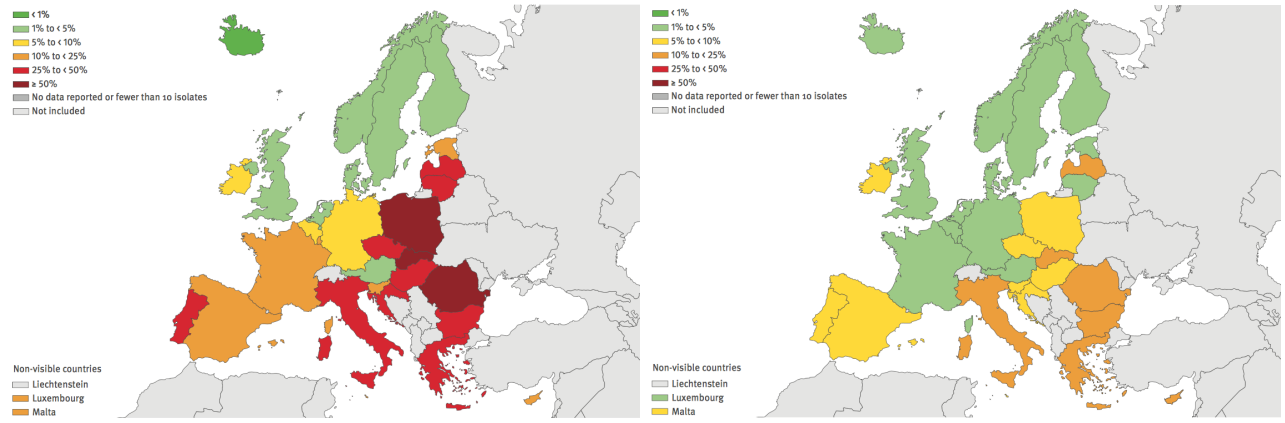


Figure 7.2: Percentage of *Klebsiella pneumoniae* (left) and *Escherichia coli* (right) bacteria with combined resistance to third-generation cephalosporins, fluoroquinolones and aminoglycosides, in EU/EEA countries in 2016. These figures were reproduced from the ECDC public record, see ref. 4.

The mechanism of action of AMPs is disrupting bacterial membrane structures. The structural features of AMPs which allow them to selectively bind to charged bacterial membranes are a net positive or negative charge and a high ratio of hydrophobic amino acids. Generally, the binding to membranes leads to non-enzymatic disruption[5]. The disruption can be caused by several causes, *e.g.* pore formation and inhibition of bacteria growth[6]. At a certain threshold concentration, AMPs permeabilise the membrane and induce cell killing. Three main models of membrane permeabilisation exist:

- Barrel-stave model: the hydrophobic region of the peptide aligns with the lipid tails, while the hydrophillic region forms the interior of the pore.
- Carpet model: the peptides bind to the surface of the membrane by covering it in a carpet-like manner, and destroy the membrane in a detergent-like way, with fragments of the membrane captured by peptide micelles.
- Toroidal-pore model: AMP helices insert into the membrane causing the tilting of lipids around the pore, so that the pore is lined by both the peptide and the lipid head groups.

Peptides can also be rationally designed to selectively bind to anionic lipid surfaces, mimicking the properties of natural AMPs. They usually have short chains and a well-defined structure, which makes them ideal to be used in mechanistic studies. A promising example of this is represented by the class of cationic, α-helical peptides designed by Jian Lu and co-workers[7] called Gn peptides. These peptides feature short sequences, $\mathrm{G\,(IIKK)}n\,\mathrm{I{-}NH_2}$ (with α-helical repeats, $n = 1$–4). They have been shown to kill both Gram-positive and Gram-negative bacteria and to possess potent antitumour activity against cancer cell lines. Specifically, Hu *et al.*[7] found that Gn peptides predominantly adopt unfolded extended conformations in aqueous solution, remaining unfolded when exposed to zwitterionic liposomes (mimicking mammalian host cell membranes) and forming α-helical structure upon exposure to negatively charged vesicles, e.g. 1,2-dipalmitoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DPPG) liposomes (mimicking bacterial cell membranes). This result was also confirmed by Ciumac *et al.*[8] by studying the interaction of G4 (the peptide of that family with $n = 4$) with 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) and DPPG monolayers, using combined measurements from Langmuir trough, Brewster angle microscopy and neutron reflectivity.

The investigation of the interactions between AMPs and lipid assemblies and the process of aggregation and pore formation improves the understanding of the AMPs biological effects. Molecular dynamics simulations offer an approach for developing a dynamical view of various molecular processes at the atomic level[9], and have often been used to investigate the interactions of peptides with lipid bilayers[10, 11, 12, 13], and describe the mechanism of pore formation

by AMPs[14, 15]. In this chapter, MD simulations were used to investigate the properties of G4 in water and the interactions with lipid monolayers, with different charges, at the atomic level.

The preliminary results presented in this chapter represent the first step of a collaboration with Manchester University (Jian Lu group) to understand how the antimicrobial peptide G4 interacts with lipid monolayers and bilayers. The comparison of MD results and those acquired using experimental techniques will enable a better interpretation of experimental data and a more detailed understanding of the G4 mechanism of action. Some of the simulations of the G4 peptide in water were run by Daniela Ciumac, under my supervision, during her 1-month secondment to Imperial College. The aim of this secondment was to teach the fundamentals of molecular dynamics to the experimental partner of the collaboration.

7.2 Methods

7.2.1 G4 in aqueous solutions

The sequence of the peptide was obtained from the experimental collaborators and it reads $G(IIKK)_4-I-NH_2$, where G stands for glycine, I for isoleucine and K for lysine. The structure of the G4 peptide involves four repetitions of the IIKK amino acid sequence with a glycine, G, before the N-terminus and an isoleucine, I, before the C-terminus, which is amidated with an NH_2 group. The amidation of the C-terminus is a common feature in most wild-type peptides. All the Gn series of peptides are C-terminal amidated and at neutral pH they have a net electrostatic charge of +9. However, the specific role of the C-terminal amidation on the peptide molecular mechanism remains unclear[16]. It is believed that it enhances antimicrobial efficacy by stabilising the α-helical conformations. In addition, C-terminal amidated peptides have a higher positive charge than those with a free C-terminus. Therefore the increased charge could also play a role in their higher activity, although charge alone would not be sufficient to explain the difference[8].

The input configurations of the G4 peptide were generated using the open source software

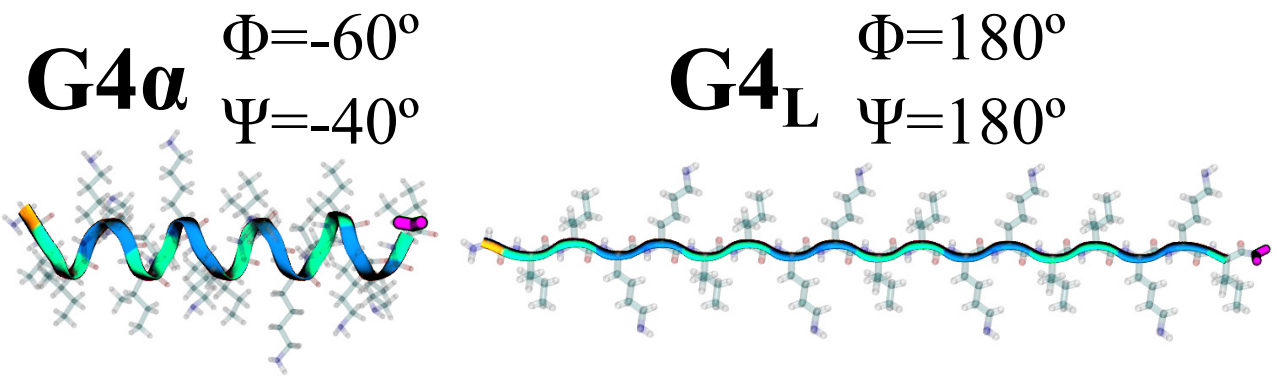


Figure 7.3: Initial configurations of the G4 peptide with α-helix (G_α on the left) and linear conformation (G_L on the right). Φ and Ψ are the initial values of the backbone dihedral angles, used to determine the secondary structure. In the amino acid sequence GLY is represented in orange, ILE in green and LYS in light blue. The NH_2-cap at the C-terminus is highlighted in purple.

Avogadro[17]. Avogadro has a built-in function which enables users to construct a peptide by specifying its amino acid sequence, the preferred input secondary structure (which is determined by the two backbone dihedral angle Φ and Ψ), and how to cap the N and C-termini.

The conformation of the peptide in water was investigated by starting from two different initial conformations: an α-helix ($G4_\alpha$) and an extended chain configuration ($G4_L$), as represented in Figure 7.3. By doing so, the simulations started from two different points in the phase space to sample different independent trajectories. The N-terminus at physiological pH is NH_3^+, while the COO^- on the C-terminus was amidated by the NH_2 group.

As mentioned in the chapter 2, several sets of forcefield parameters are available to describe the interactions of proteins and of antimicrobial peptides. In this study, the G4 peptide was described with the united atom forcefield GROMOS97-43a1[18]. This choice was supported by the analysis proposed by Best *et al.*[19] which compares a number of frequently used forcefields, namely Amber03, Amber99SB, AmberGS, CHARMM27/cmap, OPLS-aa/L, Gromos53a6 and Gromos43a1, against J-coupling NMR experiments. In general all the forcefields were found to sample the α-helix region of the Ramachandran plot too much in comparison to experiments. Among the best scoring forcefields, GROMOS43a1 was chosen since the use of UA forcefield can drastically reduce the computational time, especially in the study of the interactions between peptides and lipid monolayers and bilayers. Other analysis, *e.g.* Cino *et al.*[20], confirm this choice. The G4 peptide was placed in the centre of a cubic box of 8 nm length, to minimise

interactions between peptide periodic images, due to periodic boundary conditions, which were applied in all directions. The box was filled with $\sim$ 15000 water molecules with the GROMACS built-in tool `genbox`. The water was described with the SPC/E model, which was used to parametrise the GROMOS forcefields. 9 negative ions, *i.e.* Cl^- ions, were added using the GROMACS built-in tool `genion` to balance the positive charge of the peptide. To analyse the impact of salt concentrations on the peptide conformations, a set of simulations in which the peptide was equilibrated in a salt solution, mimicking the experimental concentration of 137 mM, was run by adding an equivalent number of Na^+Cl^- ions to the simulation box with `genion`. The ions were modelled with the GROMOS43a1 forcefield[18].

The energy of all the initial configurations was minimised using the steepest descent algorithm[21], until the maximum force between two atoms was lower than 1000 kJ mol^{-1} nm^{-1}. The system was then equilibrated for 100 ps in the canonical (NVT) ensemble and for 200 ps in the isothermal-isobaric (NPT) ensemble, using the velocity rescaling algorithm[22] with $\tau_T = 0.1$ ps to keep the temperature constant at 295 K and the Parrinello-Rahman barostat applied isotropically and set at 1 bar with $\tau_P = 2$ ps and compressibility, $\beta = 4.5 \times 10^{-5}$bar^{-1}. The timestep was set equal to 2 fs. van der Waals interactions were truncated at 1.0 nm and electrostatic interactions were computed with the Ewald summation technique[23].

System	**type**	$\mathbf{N_{WATER}}$	$\mathbf{N_{IONS}}$	$\mathbf{N_{LIPID}}$	**Ensemble**	$\mathbf{t_{RUN}}$
G4α water	classic MD	16997	9 Cl	0	NPT	175 ns
G4α salt sol.	classic MD	16913	42 Na 51 Cl	0	NPT	175 ns
G4$_L$ water	classic MD	16997	9 Cl	0	NPT	175 ns
G4$_L$ salt sol.	classic MD	16913	42 Na 51 Cl	0	NPT	175 ns
G4α water	solv$_{0.1}$	16997	9 Cl	0	NPT	100 ns
G4α salt sol.	solv$_{0.1}$	16913	42Na 51Cl	0	NPT	100 ns
G4$_L$ water	solv$_{0.1}$	16997	9 Cl	0	NPT	100 ns
G4$_L$ salt sol.	solv$_{0.1}$	16913	42 Na 51 Cl	0	NPT	100 ns
G4 water	32 x REMD	9005	9 Cl	0	NPT	50 ns each
G4 + DPPC salt sol.	classic MD	28478	70 Na 79 Cl	256	$N\gamma T$	500 ns
G4 + DPPG salt sol.	classic MD	28222	326 Na 79 Cl	256	$N\gamma T$	700 ns

Table 7.1: List of peptide simulations with simulation parameters. N_{WATER}, N_{IONS} and N_{LIPID} represent the number of water, ions and lipids molecules, respectively. t_{RUN} is the time length of the production run.

Table 7.1 summarises the set of simulations reported in this chapter. The equilibrated G4$_\alpha$

and $G4_L$ peptide conformations in water and 137 mM NaCl solution were run for 175 ns in the NPT ensemble, coupled to a Nosé-Hoover thermostat ($T = 293\,\mathrm{K}$, $\tau_T = 1\,\mathrm{ps}$) and a Parrinello-Rahman barostat ($p = 1\,\mathrm{bar}$, $\tau_P = 2\,\mathrm{ps}$, $\beta = 4.5 \times 10^{-5}\mathrm{bar}^{-1}$).

In addition, as described in chapter 2, several techniques can be used to accelerate the conformational dynamics of peptides. The strategy, proposed by Lin and Tuckerman[24], which involves changing the relative masses of the solvent and solute to decouple their relative motion, was used here, by reducing the solvent masses by 10 times. These simulations (labelled as $\mathrm{solv}_{0.1}$ in Table 7.1) were run for 100 ns in the NPT ensemble, with the same simulation parameters. The time step of 2 fs was small enough to account for the higher frequency induced by the mass scaling, as further discussed in the results section. One configuration taken at the end of the $G4_\alpha$ simulation in water was used to run 32 T-REMD (temperature replica exchange molecular dynamics) simulation in the temperature range 295 K–350 K. To reduce the computational cost of these simulations the conformation was first solvated again and equilibrated in a smaller simulation box (cubic, with size 6.5 nm) with 9005 water molecules. The choice of the temperature distribution in replica exchange simulations is a key parameter to ensure a constant probability of exchange, p_{ex}, across the replicas (as discussed in detail in section 2.3.2). The chosen temperatures followed an exponential distribution and were set according to the procedure proposed in ref. 25 to obtain a $p_{ex} \sim 0.25$, as: [295.0, 296.7, 298.4, 300.1, 301.8, 303.5, 305.2, 306.9, 308.6, 310.4, 312.0, 313.8, 315.5, 317.3, 319.1, 320.9, 322.7, 324.5, 326.3, 328.1, 330.0, 331.8, 333.7, 335.5, 337.4, 339.3, 341.1, 343.0, 345.0, 346.9, 348.8, 350.7] K. Each replica was first equilibrated at its temperature and then run in the NPT ensemble (parameters as before) for 50 ns.

7.2.2 G4 interaction with DPPC and DPPG monolayers

The experimental results of Ciumac *et al.*[8] highlight the selective interactions between the G4 peptide and lipid monolayers with different head group charges. Specifically, neutron reflectivity experiments show that the peptide bound 4 times more strongly onto a DPPG monolayer than onto a DPPC monolayer, increasing the ability of the peptide to disrupt charged membranes.

Therefore, here the interactions of the G4 peptide with DPPC and DPPG monolayers were studied to unravel the mechanisms of interaction at the molecular level.

The Berger forcefield[26] was used to model the phospholipid molecules. The initial DPPC monolayer configuration was generated from an equilibrated configuration taken from ref. 27, which was replicated 4 times in the xy plane to obtained a monolayer of 256 lipids. The DPPC lipid molecules were mutated to DPPG by replacing the terminal $CH_2N(CH_2)_3^+$ atoms by CHOHCHOH to generate the initial DPPG monolayer configuration (as shown in Figure 7.4).

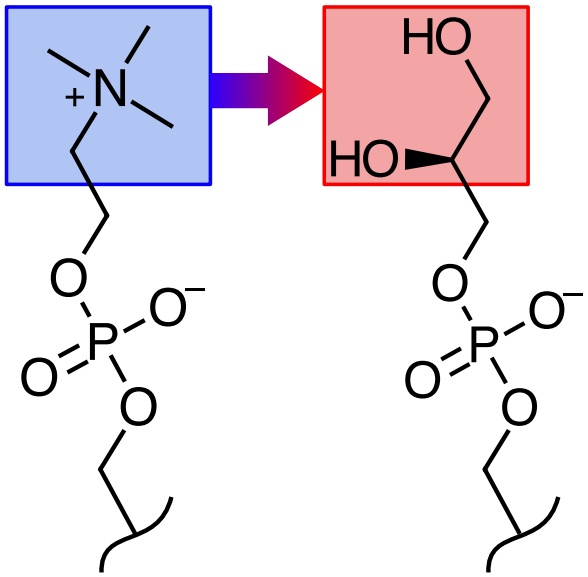


Figure 7.4: Mutation of DPPC head group to DPPG head group. The terminal $CH_2N(CH_2)_3^+$ atoms are replaced by CHOHCHO.

Approximately 28 500 water molecules (see exact numbers in Table 7.1) were added below the lipid head groups, which were partially immersed in water. To balance the negatively charged DPPG head group, 256 positively charged Na^+ ions were added by replacing some of the water molecules in the DPPG-water system. Na^+Cl^- ions were also added to both systems to reach the experimental salt concentration of 137 mM (the precise number of ions can be read in Table 7.1). The simulation box was elongated in the z direction to create two vacuum regions above and below the water-lipid system, which was set parallel to the xy plane. Periodic boundary conditions were applied in all directions. The size of the vacuum region was big enough to ensure the absence of spurious effects connected to the interactions between the molecules in the central box and their periodic images along z. Bresme and co-workers[27, 28]

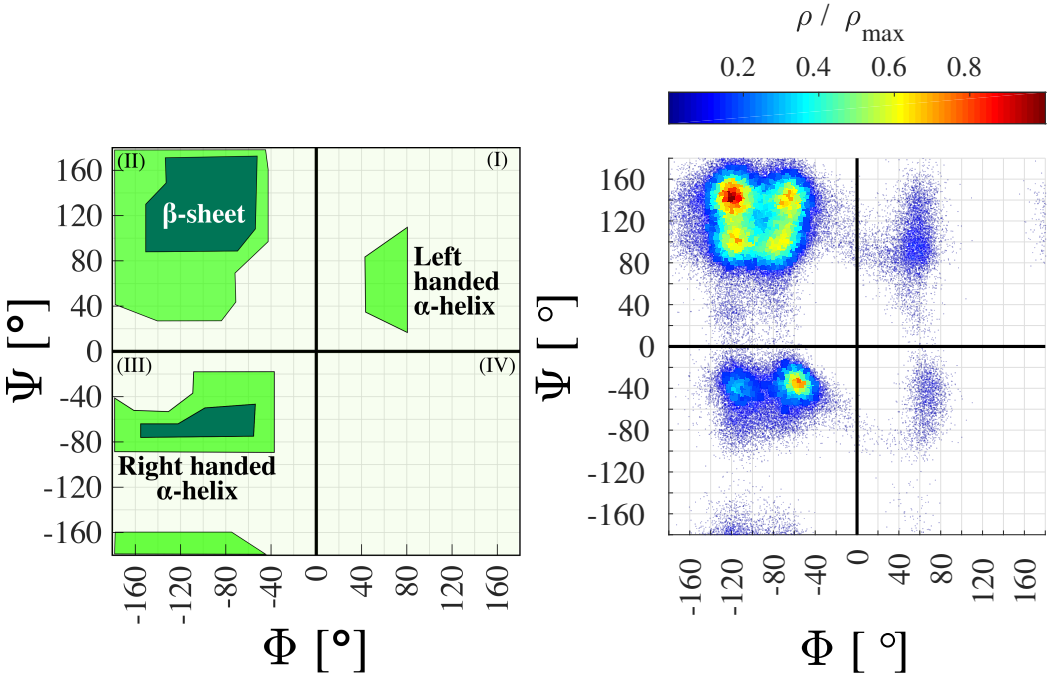
showed that the monolayer pressure-area isotherms are sensitive to the chosen cut-off radius, $r_{cut-off}$, of short-range Lennard–Jones interactions. Therefore, $r_{cut-off}$ was set equal to 1.7 nm for a better representation of the surface pressure. The long-range Coulombic interactions were computed using the Particle Mesh Ewald summation method[23], including dipole corrections for the Ewald Sum[29]. Water molecules were described by the SPC/E water model. Although different water models, *e.g.* TIP4P/2005[30] and OPC[31], were found to better reproduce the water surface tensions and therefore give more realistic pressure area isotherms[27, 32], the SPC/E model was used throughout this analysis for its compatibility with the lipid and peptide forcefields (*i.e.* Berger and GROMOS97-43a1[18] respectively). The resulting input configurations were firstly equilibrated in the NVT ensemble for 10 ns, at 293 K using the Berendsen thermostat, with $\tau_T = 0.1$ ps. Then, the monolayers were run for 80 ns at constant temperature and surface tension parallel to the xy-plane with $\gamma_t = 1000$ bar nm. Temperature and pressure were equilibrated with the Berendsen algorithm with coupling constants of 0.1 ps and 1.0 ps, respectively.

Finally, a representative conformation of the peptide taken from the $G4_L$ simulation in a 137 mM NaCl water solution was solvated in the water solution under the lipid monolayer, with an initial distance of 5 nm between the peptide centre of mass and the head group region of the DPPC and DPPG monolayers. 9 negatively charged Cl^- ions were added to balance the peptide charge. The system was equilibrated in NVT for 20 ns, at 293 K using the Berendsen thermostat, with $\tau_T = 0.1$ ps. The equilibrated G4-monolayer configurations were run for 600 ns at constant temperature and surface tension, $\gamma_t = 1000$ bar nm, parallel to the monolayer (xy-plane), with the same set-up as before.

7.2.3 Ramachandran plot

Section 2.5.2 presents a description of how the secondary structure of a peptide can be determined by the 2D diagram of the backbone angles Φ and Ψ, known as a Ramachandran plot (RP) (as shown in Figure 7.5a). In this chapter, the RP for the G4 peptide is plotted and used to give a first estimation of the tendency of the peptide to assume a linear coil or an α-helix

conformation.



(a) Ramachandran plot schematic **(b)** Ramachandran plot for $G4_L$

Figure 7.5: Ramachandran plots showing a schematic plot with secondary structures (a) and an example of the computed Ramachandran plot for $G4_L$ peptide (b).

The GROMACS built in function g_rama was used to compute the dihedral backbone angles of each residue. The energetically favourable regions for G4 conformations, expressed in terms of Φ and Ψ were then represented on the 2D RP, as shown in Figure 7.5b. Each point in the graph corresponds to one average conformation of a G4 residue in the 2D space. The colour and size of the points reflects the relative number density probability of each conformation, using a mesh-grid with size of 1°. The correspondence between colour and the probability density (computed as ρ/ρ_{max}) is given in the colour bar above the graph. The conformational distribution in each of the four quadrants was assigned to a type of secondary structure:

- The conformations in the first quadrant (I) ($0° < \Phi < 180°$ and $0° < \Psi < 180°$) were counted as left-handed α-helix conformations ($p_{LH\alpha h}$ in the tables below);
- The conformations in the second quadrant (II) ($-180° < \Phi < 0°$ and $0° < \Psi < 180°$) were counted as extended coil conformations (p_β the tables below);
- The conformations in the third quadrant (III) ($-180° < \Phi < 0°$ and $-180° < \Psi < 0°$)

were counted as right-handed α-helix conformations ($p_{RH\alpha h}$ in the tables below);

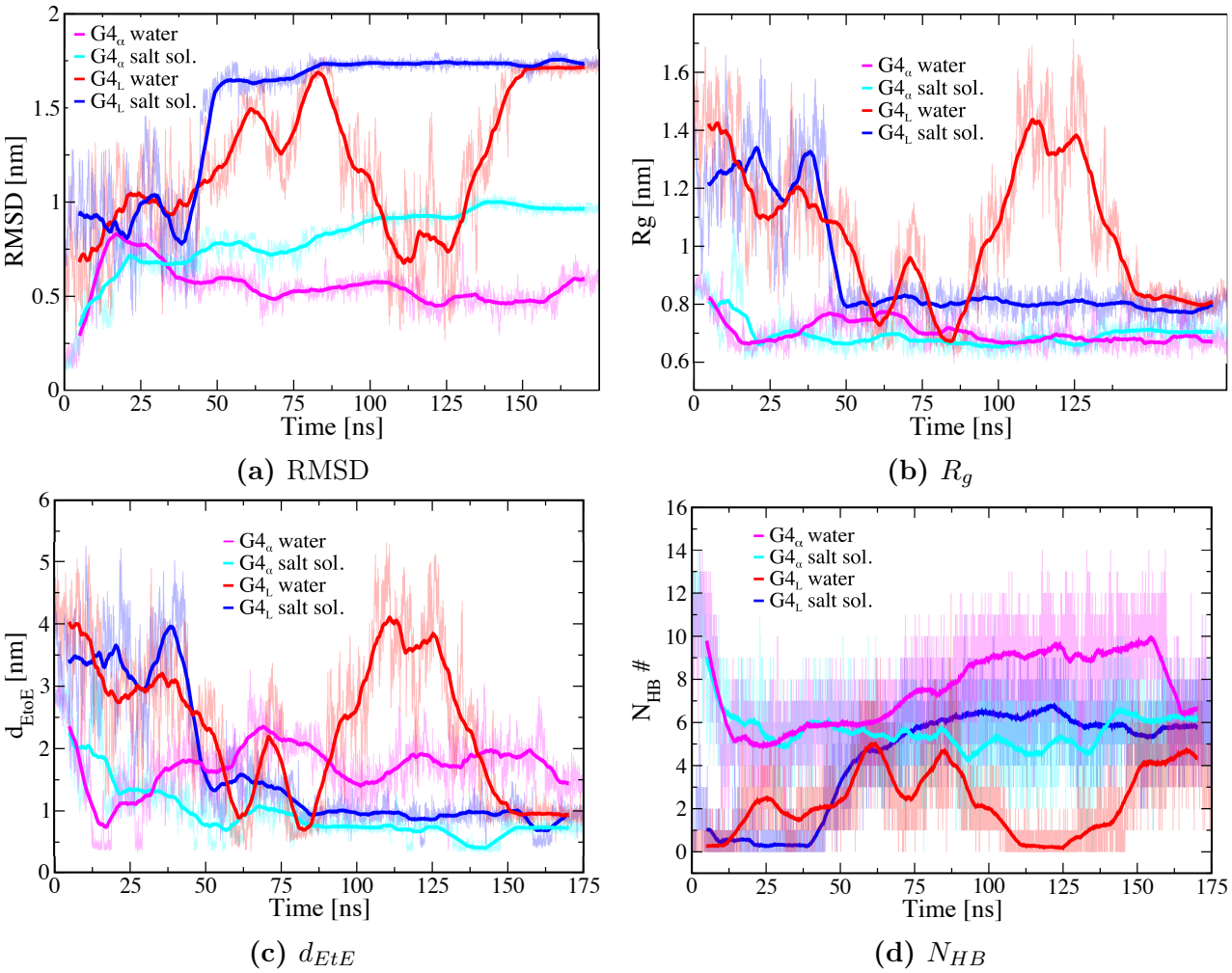
- The conformations in the fourth quadrant (IV) ($0° < \Phi < 180°$ and $-180° < \Psi < 0°$) was recorded as p_{none}. These conformations are very rare due to steric clashes between peptide atoms.

In the results sections below the conformation probabilities are given both as averages considering the whole peptide secondary structure (see Table 7.2 and 7.4) and considering the secondary structure conformation around each residue, individually (see Figure 7.8, 7.11, 7.13 and 7.17). The residue conformation was described as α-helix when $(p_{LH\alpha h} + p_{RH\alpha h}) \geq 55\%$ (reported in the figures as a red square), as random coil when $(p_{LH\alpha h} + p_{RH\alpha h}) \leq 45\%$ (reported in the figures as a green square). An orange square was used when the probabilities of the two conformations were similar ($45\% < (p_{LH\alpha h} + p_{RH\alpha h}) < 55\%$).

7.3 G4 in water

Circular dichroism experiments have shown that the G4 peptide is unfolded in aqueous solution. However, the native conformation of the G4 peptide is not available from experimental studies. Hence, the simulations presented here were started from two different conformations (*i.e.* α-helix and fully-extended - as shown in Figure 7.3) and track the evolution of these conformations in time. The evolution of the peptide conformation was tracked by computing the parameters plotted in Figure 7.6.

As discussed in section 2.5.2, the RMSD gives a measure of the average distance between backbone atoms of superimposed structures, with higher RMSD values corresponding to greater variation from the reference conformation. The RMSD values shown in Figure 7.6a measure the difference between the initial configuration and the configuration at time, t. G4 peptides starting from the fully-extended conformation (see blue and red curves in Figure 7.6a) have a larger RMSD. This could be easily understood considering that a perfectly straight peptide conformation is very unstable. Interestingly the $G4_L$ peptide equilibrates quicker in salt solution than in water, and its RMSD curve becomes constant after ~ 50 ns. Despite a longer



(a) RMSD (b) R_g

(c) d_{EtE} (d) N_{HB}

Figure 7.6: a) Radial Mean Square Displacement (RMSD) *vs.* time. b) Radius of gyration *vs.* time. c) End to End distance *vs.* time. d) Hydrogen bonds *vs.* time. Different curves correspond to different initial configurations: magenta for G4$_\alpha$ water, cyan for G4$_\alpha$ salt solution, red for G4$_L$ water and blue for G4$_L$ salt solution. The thicker lines are running averages over 100 time steps.

equilibration time ($\sim$ 150 ns), the final values of the RMSD for G4$_L$ in water and in salt solution converge towards the same value. Conversely, the RMSD values for the peptides starting from the α-helix conformation (magenta and cyan curves in Figure 7.6a), do not converge after 175 ns. Although in this case the variation from the initial configuration is smaller than in the G4$_L$ case, the RMSD values are still too high to be interpreted as indicating conformational stability. Note that generally, an RMSD higher than 0.3 nm indicates two structures with very different conformations[33].

Figure 7.6b shows the evolution of the radius of gyration, R_g (see equation 2.30). Smaller R_g values correspond to more compact structures. All the simulations result in a smaller value of R_g than the corresponding initial configurations, with an equilibrium value which converge for

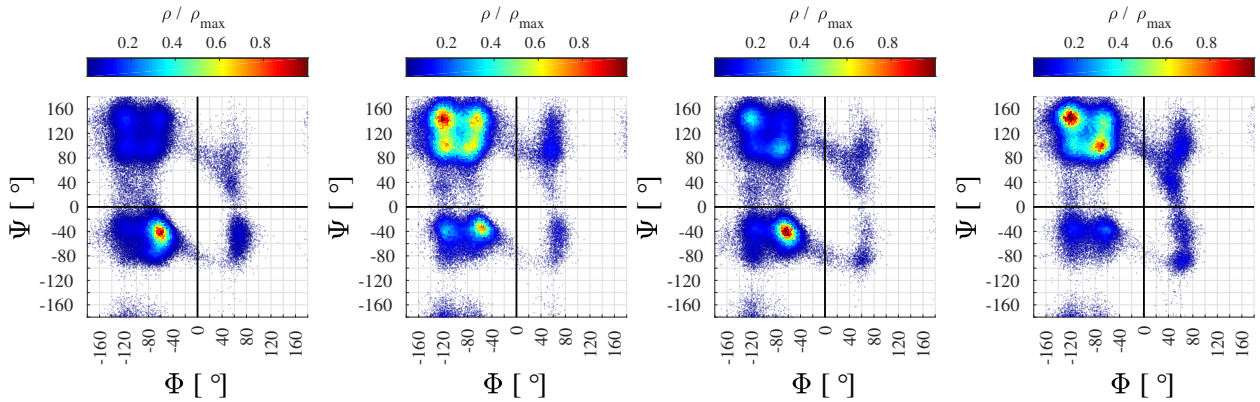
all the four curves (with $G4_L$ in water featuring a longer equilibration time). Interestingly, the end to end distance, defined as the distance between the C_α of the first (GLY) and last (ILE) residue, also decreases with time and converges towards a similar value in all four cases (see Figure 7.6c).

The structural stability of β-sheet or α-helix structures is the result of hydrogen bonding between the amide proton (the hydrogen atom bonded to the nitrogen of the peptide bond) and the carbonyl oxygen (the oxygen atom bonded to the carbon atom in the amide plane). The number of hydrogen bonds within the peptide backbone was computed *vs.* time and represented in Figure 7.6d. In the initial extended conformations G4 cannot form any backbone hydrogen bonds, while in the initial α-helix conformation 14 hydrogen bonds are formed. After 175 ns both configurations present ~ 5 hydrogen bonds. Irrespective of the initial configuration, the presence of salt stabilises the number of hydrogen bonds and, thus, the peptide secondary structure (see blue and cyan curves in Figure 7.6d).

To give insight into the secondary structure of G4 in water and salt solution, the Φ and Ψ torsional angles were computed and represented in Figure 7.7. The Ramachandran diagrams were used to estimate the probability of α-helix or coil conformations, which are reported in table 7.2. As expected, starting from an ideal α-helical conformation results in a higher probability of occurrence of this conformation (see diagrams 7.7a and 7.7c). Most likely, time scales of hundreds of nanosecond or microseconds would be required for a complete unfolding of this conformation. In addition, a factor to keep in mind for the interpretation of these results is that most of the current protein forcefield overpopulate the α-helix region (as shown by the study performed by Best *et al.*[19]).

A second clear message from this analysis is that adding Na^+Cl^- ions to the water increases the probability of coil conformations (see diagrams 7.7c and 7.7d and table 7.2). The impact of ion concentration on peptide secondary structure has been discussed in previous MD studies[34, 35, 36]. In agreement with the findings of our analysis, Fedorov and co-workers found that NaCl solutions favour the formation of extended conformations[35, 36]. Ioannou *et al.*[34] observed that the presence of electrolytes can have a bivalent effect, promoting helix nucleation,

but reducing helix propagation, and has, therefore, a different impact depending on the length of the amino acid chain. In general the effect of salt on the peptide conformations can be quite complex, involving ion hydration, ion pairing, and specific ion peptide group interactions, and further analysis would be required to shed light on this question[37, 38].



(a) G4$_\alpha$ water **(b)** G4$_L$ water **(c)** G4$_\alpha$ salt solution **(d)** G4$_L$ salt solution

Figure 7.7: Ramachandran diagrams for the G4 peptide. The four plots correspond to different input configurations: (a) α-helix conformation in water; (b) extended conformation in water; (c) α-helix conformation in 137 mM NaCl solution; (d) extended conformation in 137 mM NaCl solution.

Initial configuration	p_β [%]	$p_{RH\alpha h}$ [%]	$p_{LH\alpha h}$ [%]	p_{none} [%]
G4$_\alpha$ water	37.3	56.6	2.1	4.0
G4$_\alpha$ salt solution	51.7	43.9	3.0	1.4
G4$_L$ water	73.8	19.7	4.9	1.6
G4$_L$ salt solution	78.5	11.0	7.2	3.3
G4$_\alpha$ water - solv$_{0.1}$	71.1	19.0	3.0	6.9
G4$_\alpha$ salt solution - solv$_{0.1}$	55.0	34.9	3.5	6.6
G4$_L$ water - solv$_{0.1}$	82.4	11.2	2.7	3.6
G4$_L$ salt solution - solv$_{0.1}$	83.2	6.2	7.1	3.4
G4$_\alpha$ water - REMD ($T = 295$ K)	73.9	15.6	6.1	4.4

Table 7.2: Probabilities of occurrence of G4 secondary structures. solv$_{0.1}$ denotes the simulations with mass rescaling. See section 7.2.3 for a description of how the probabilities were defined and computed.

To gain a better understanding of the peptide secondary structure the structure around each residue was computed individually. The result is shown in Figure 7.8. From the figure it appears that the helix conformations start to unfold from the N and C termini rather than from the middle. In particular the residues in the proximity of the amidated C-terminus are most likely to assume coil conformations within the simulation time.

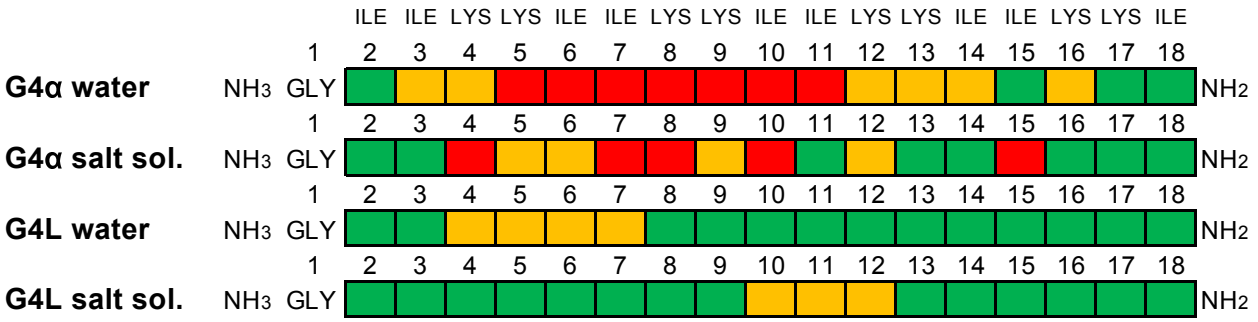


Figure 7.8: Probability of secondary structure for the different G4 residues. The probability is represented by coloured squares. The square is green when the most probable secondary structure is the random coil (II and IV quadrants of the Ramachandran plot - see section 7.2.3). The square is red when the most probable secondary structure is the the α-helix (I and III quadrants of the Ramachandran plot - see section 7.2.3). An orange square was used when the two % were similar (in range the 45–55%).

Figure 7.9 shows two representative G4 conformations for the peptide starting from the helix (left) and extended (right) conformations in 137 mM NaCl solution. The conformations were obtained by applying the Daura clustering algorithm[39] to the backbone of the peptide in the last 25 ns of the trajectory. This algorithm compares the position of protein backbone atoms between trajectory snapshots and groups them together based on their root mean square displacement, and can be computed via the GROMACS built-in function g_cluster. Using a cutoff threshold of 2 Å, only one cluster for each set of simulations was obtained and its conformation is shown in Figure 7.9. Both $G4_\alpha$ and $G4_L$ peptides after 175 ns assume an hairpin β-sheets conformation, with two strands.

7.3.1 Acceleration techniques

In this section two techniques were used to better sample the peptide phase space and accelerate the conformational sampling, namely mass rescaling (proposed by Lin and Tuckerman[24]) and replica exchange molecular dynamics.

Mass rescaling

As mentioned previously, the choice of solvent masses (used here to reduce the solvent viscosity and improve conformational sampling) does not critically affect the equilibrium average of a

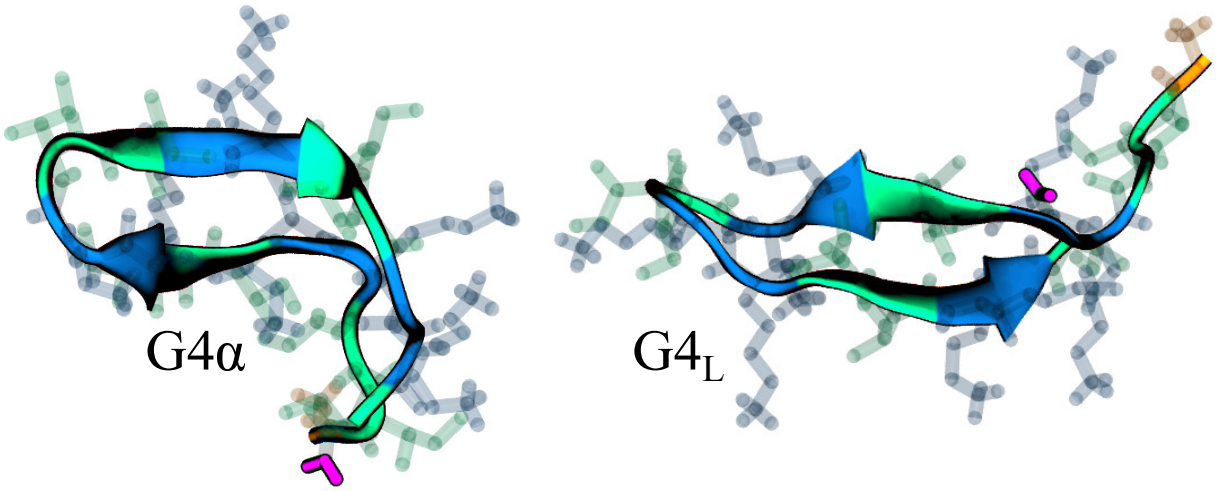


Figure 7.9: Configurations of the G4 peptide after 175 ns, computed with the clustering algorithm proposed by Daura and co-workers[39], using data from the last 25 ns of trajectory. In the amino acid sequence GLY is represented in orange, ILE in green and LYS in light blue. The NH_2-cap is highlighted in purple.

position-dependent function. Some examples of quantities whose averages do not depend on the choice of masses (average energies, temperatures, and pressures) are tabulated in Table 7.3, and the differences in these quantities over the various simulations are negligible.

	$G4_\alpha$	$G4_\alpha$ $solv_{0.1}$	$G4_L$	$G4_L$ $solv_{0.1}$
E_{TOT} [kJ mol^{-1}]	−589272(1029)	−591883(1017.66)	−593378(1036)	−596006(1035)
U [kJ mol^{-1}]	−718950(857)	−721552(844)	−718964(868)	−721591(863)
T [K]	295(1.3)	295(1.2)	295 (1.2)	295(1.2)
p [bar]	2(98)	2(98)	2(99)	2(98)
	$G4_\alpha$NaCl	$G4_\alpha$NaCl $solv_{0.1}$	$G4_L$NaCl	$G4_L$NaCl $solv_{0.1}$
E_{TOT} [kJ mol^{-1}]	−624588(1017)	−627145(1020)	−624585(1031)	−627198(1027)
U [kJ mol^{-1}]	−749866(854)	−752412(853)	−749865(858)	−752458(857)
T [K]	295(1.3)	295(1.2)	295(1.3)	295(1.3)
p [bar] (bar)	1(98)	3(97)	−1(98)	0(99)

Table 7.3: Average total and potential energies (E_{TOT} and U), temperatures (T) and pressures (p) in the simulation with standard and modified ($solv_{0.1}$) solvent masses. The standard deviations of the mean values are expressed in parentheses.

The differences between the standard and $solv_{0.1}$ simulations can be observed by comparing Figure 7.10 and Figure 7.6. Applying the mass rescaling technique results in a higher degree of convergence for all the peptide properties. In addition, the equilibration time scale significantly decreases, with most of the properties converging within 25 ns. This result is well reflected in the analysis of the peptide secondary structure. The probability of helix conformations

dramatically decreases after decreasing the solvent masses by a factor of 10 (*e.g.* from 58.7% to 22% for the G4$_\alpha$ peptide in water), as shown in Table 7.2. Figure 7.11 shows the most probable conformations for the peptide residues. Most of these are characterised by coil conformations, especially in the proximity of the two termini, confirming the results from the simulations without mass re-weighting.

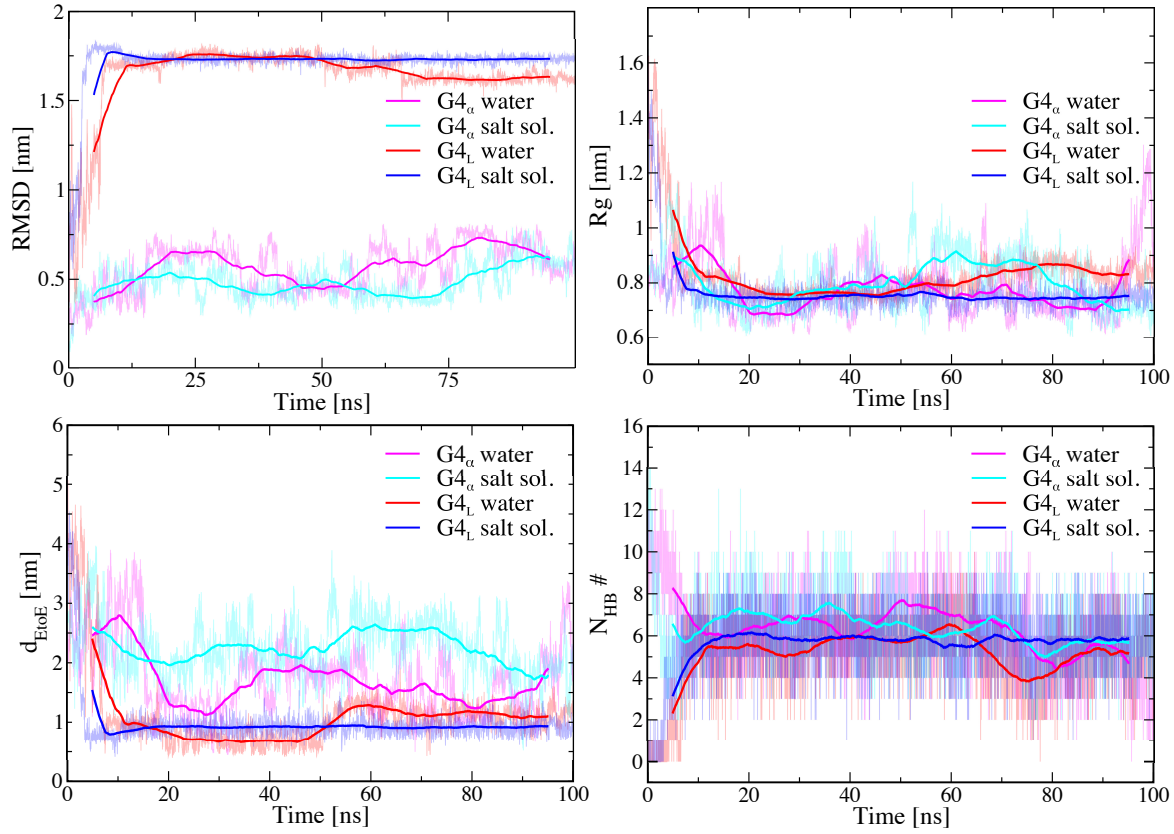


Figure 7.10: solv$_{0.1}$ simulations: a) Radial mean square displacement (RMSD) *vs.* time. b) Radius of gyration (R_g) *vs.* time. c) End to end distance (d_{EtoE}) *vs.* time. d) Number of hydrogen bonds (N_{HB}) *vs.* time. Different curves correspond to different initial configurations: magenta for G4$_\alpha$ in water, cyan for G4$_\alpha$ in salt solution, red for G4$_L$ in water and blue for G4$_L$ in salt solution. The thicker lines are running averages over 100 time steps.

Replica Exchange Molecular Dynamics

Each replica of the REMD simulations needs to be able to escape from its local energy minimum and exhibit a random walk in the temperature and energy space. Firstly, potential energy histograms at each target temperature were computed to see whether the overlap of the potential energy of consecutive replicas was sufficient. Figure 7.12-left shows that the chosen temperature

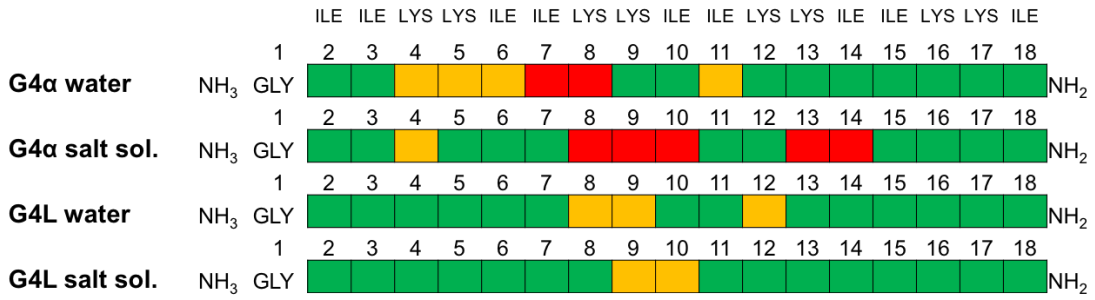


Figure 7.11: Probability of secondary structure for the different G4 residues with modified solvent masses. The probability is represented by coloured squares. See the caption of Figure 7.8 for the colours legend.

distribution does ensure sufficient overlap. The GROMACS Perl script demux was then used to demultiplex the replicas (*i.e.* to reconvert the multiple trajectories back into the original separate and unrelated ones) and write out the location of a given conformation over time. Figure 7.12-right shows the evolution of the 32 temperature indices over the first 150 ps of the simulation, showing a random walk in the temperature space, which suggests the REMD simulations were performed properly.

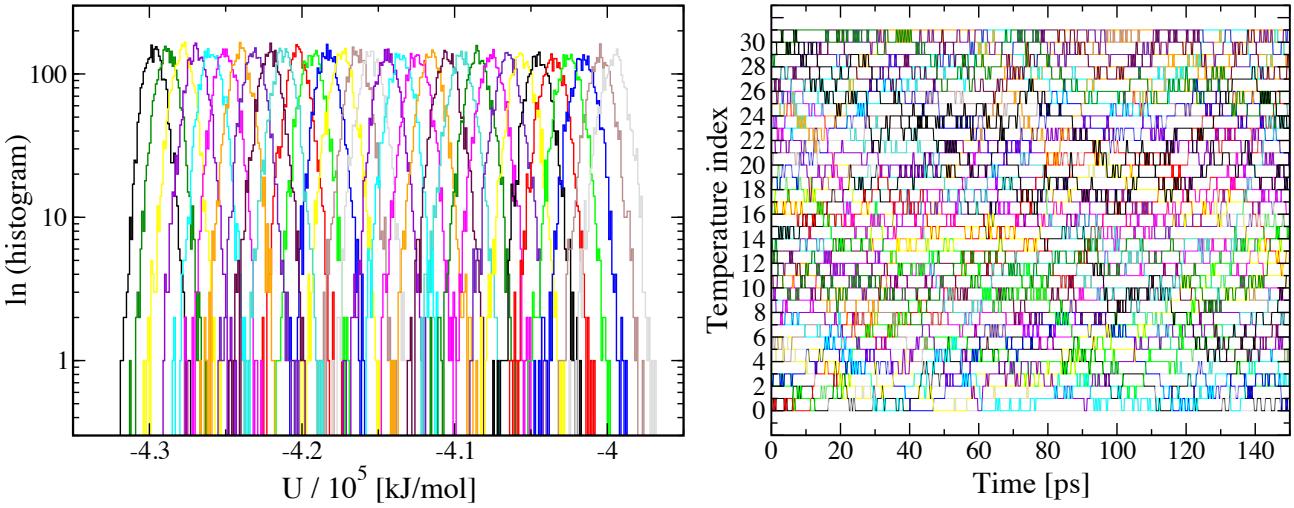


Figure 7.12: Left) REMD potential energy histogram overlap. Right) A demultiplexed set of 32 REMD replicas that propagated various peptide conformations across multiple replicas. The first 150 ps of simulation are displayed for clarity.

The replica at $T = 295\,\mathrm{K}$ was further analysed to determine its secondary structure. The probabilities of each conformation are reported in Table 7.2. The starting point of the REMD simulation was the $G4_\alpha$ conformation from the end of the 175 ns classic MD run. After the REMD simulations the probability of the coil conformations increased significantly (from 37.3% to 73.8%). The corresponding Ramachandran diagram is shown in Figure 7.13, showing higher

population of the II quadrant of the graph (see Figure 7.7 for comparison). Also the secondary structures around individual residues show higher probabilities for coil conformations (see Figure 7.13-right). This shows how the REMD simulations helped the peptide to escape from a local minimum conformation in the phase space.

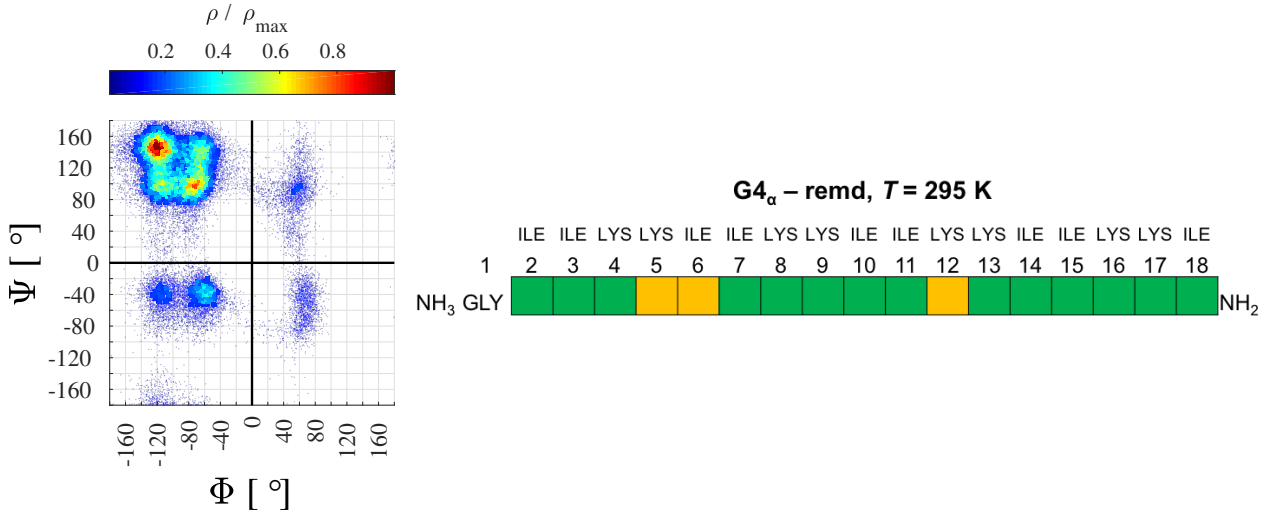
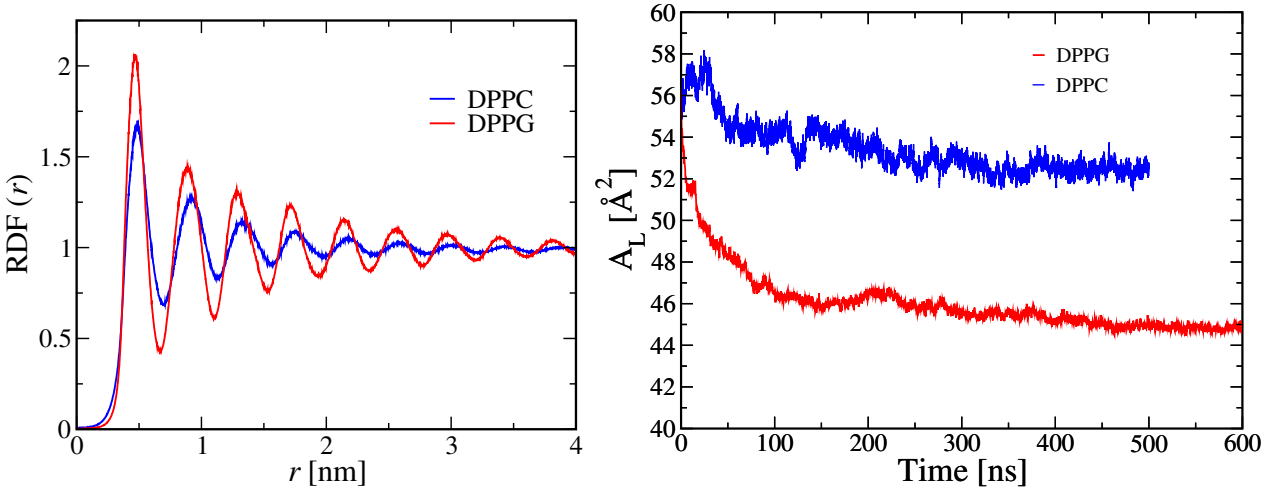


Figure 7.13: REMD replica at $T = 295\,\mathrm{K}$: Left) Ramachandran plot. Right) Probability of secondary structure for the different G4 residues. The probability is represented by coloured squares. See the caption of Figure 7.8 for the colours legend.

7.4 G4 interacting with lipid membranes - Lipid charge effect

In this section, the interactions between the G4 peptide and a lipid monolayer with different head group charges are discussed. The selectivity of the G4 peptide in forming an α-helix structure when interacting with negatively charged vesicles is considered the reason why the G4 peptide can kill both Gram-positive and Gram-negative bacteria with high potency whilst maintaining low cytotoxicity against human erythrocyte cells. This has been confirmed by neutron experiments on DPPC and DPPG monolayers[8]. The peptide binding experiments by Ciumac *et al.*[8] were carried out at an initial surface pressure of $15\,\mathrm{mN\,m^{-1}}$. To reproduce a similar surface pressure, the monolayer simulations were run at a constant surface tension of $\gamma_t = 100\,\mathrm{mN\,m^{-1}}$. Given the surface tension of the SPC/E water model with a cutoff equal

to 1.7 nm, $\gamma_W \approx 58.5\,\mathrm{mN\,m^{-1}}$ [27], this will result in a surface pressure of $\sim (2\gamma_W - \gamma_t) = 17\,\mathrm{mN\,m^{-1}}$. From the simulations of the monolayers in water, the area per lipid, A_L, computed from the fluctuations of the simulation box (see equation 3.5), were 55 $\mathrm{Å}^2$ and 52 $\mathrm{Å}^2$ for the DPPC and DPPG systems, respectively. These values compare well with the area per lipid of those monolayers presented in ref 8 ($53 \pm 2\,\mathrm{Å}^2$ for the DPPC monolayer and $54 \pm 2\,\mathrm{Å}^2$ for the DPPG monolayer). At surface pressure of $15\,\mathrm{mN\,m^{-1}}$, the lipids were found in the liquid tilted condensed phase[8]. Insights into the lipid phase of the simulated monolayer were obtained by computing the lateral radial distribution function (RDF) of three carbon atoms in the middle of each of the DPPC and DPPG lipid chains, as previously proposed in ref. 28. As shown in Figure 7.14a, both RDFs show the long-range oscillatory behaviour, characteristic of the liquid condensed phase.



(a) RDF **(b)** A_L *vs.* time

Figure 7.14: (a) 2D radial distribution function (RDF) of the carbon atoms in the phospholipid chains. The RDFs were obtained using three carbon atoms in the middle of each phospholipid chain (as explained in ref. 28). The blue curve corresponds to DPPC and the red one to DPPG. (b) A_L *vs.* time for the DPPC (blue curve) and DPPG (red curve) monolayers. The G4 peptide is added at Time= 0.

When adding the peptide, the area per lipid of the two monolayers dropped to 52 $\mathrm{Å}^2$ and 45 $\mathrm{Å}^2$, respectively (see Figure 7.14b). In the binding experiments by Ciumac *et al.*[8] after injecting the peptide at constant area in the Langmuir trough, the surface pressure rose sharply due to interactions of the peptide with the lipid monolayers. In our simulations, performed at constant γ_t, this results in a decrease in the area per lipid.

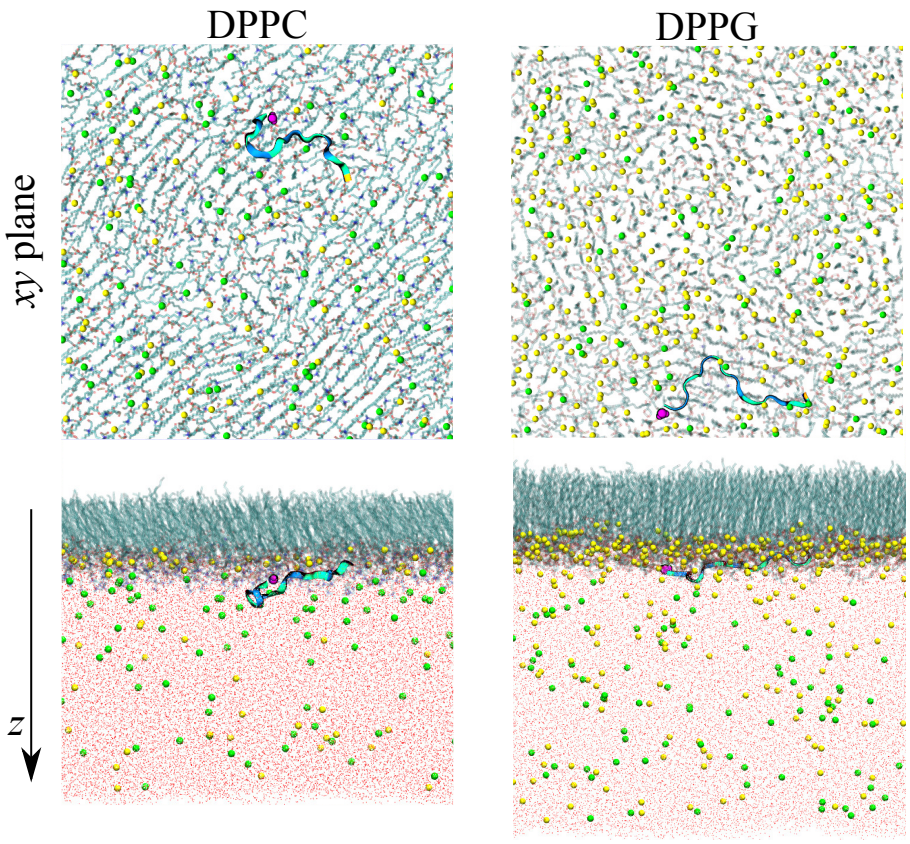


Figure 7.15: Representative snapshots from the simulations showing the G4 peptide interacting with the DPPC (left) and DPPG (right) monolayers. Water O and H are represented as red and white dots, respectively. Na^+ ions as yellow beads and Cl^- ions as green beads. DPPC and DPPG lipid molecules are represented as dynamic bond, C atoms are coloured in cyan, N in blue, O in red , H in white and P in brown. The peptide representation is described in Figure 7.9.

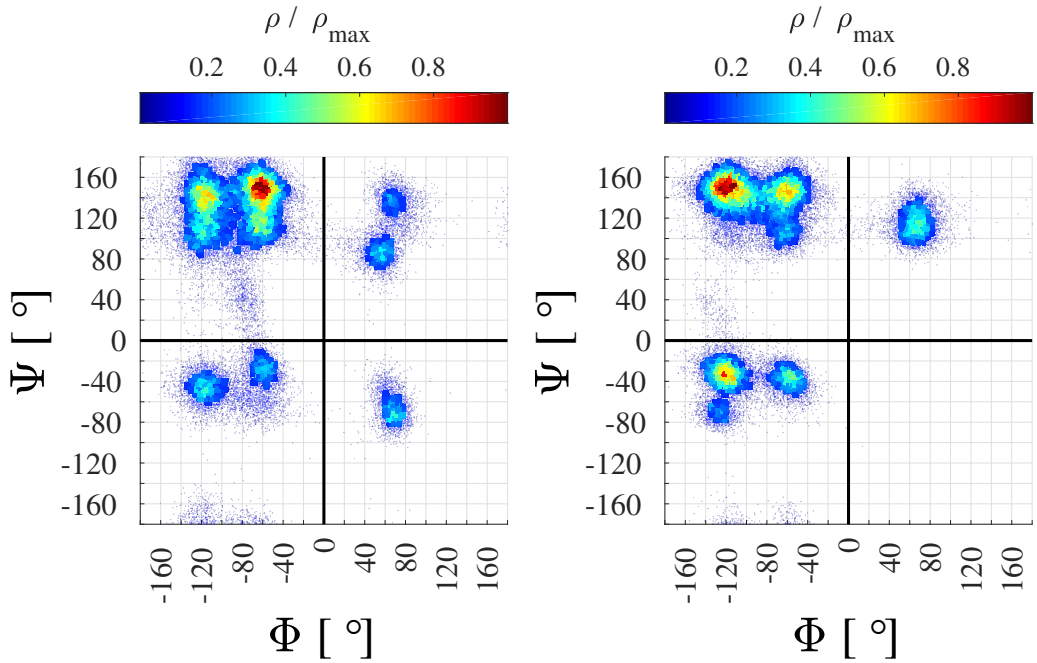
Two representative snapshots from the simulations are shown in Figure 7.15 showing how the peptide interacts with the two monolayers. In both cases the peptide goes to the water-monolayer interface, however, Figure 7.15 suggests that the interaction with DPPG is stronger, with the peptide completely absorbed at the interface. The model which best fits the neutron experiments of Ciumac *et al.*[8], suggests that a few G4 peptides interact with the DPPC monolayer head group region, where the peptide molecules lie flat or tilted. In contrast, more peptide molecules are bound to the DPPG monolayer, with penetration of the peptide into the acyl chain region. Further simulations with a concentration of peptide closer to the experimental one are required to compare our results to the proposed model.

The key factor for G4 penetration in the anionic lipid monolayer is the transition from non-ordered to α-helix structure. The analysis of the secondary structure of the peptide from the backbone dihedral angles was used to detect the development of helix structure due to

the interaction with the negatively charged lipids. Interestingly, the probabilities of α-helix conformations start to increase when the peptide is in the proximity of the DPPG monolayer (as shown in the % presented in Table 7.4 and the Ramachandran plots in Figure 7.16), and it is higher than in the proximity of the DPPC monolayer (47.8% *vs.* 28.2% for DPPG and DPPC, respectively).

Initial configuration	p_β [%]	$p_{RH\alpha h}$ [%]	$p_{LH\alpha h}$ [%]	p_{none} [%]
G4 + DPPC monolayer	65.8	16.7	11.5	6.0
G4 + DPPG monolayer	52.1	36.6	11.2	0.1

Table 7.4: Probabilities of occurrence of G4 secondary structure. See section 7.2.3 for a description of how the probabilities were defined and computed.



(a) G4-DPPC **(b)** G4-DPPG

Figure 7.16: Ramachandran plot of a) G4 interacting with a DPPC monolayer. b) G4 interacting with a DPPG monolayer.

Looking at the contribution of the single residues to the secondary structure, it is possible to observe that the folding process for G4 interacting with DPPG starts from the amidated C-terminus. This suggests a possible key role for the peptide amidation in promoting the selective antimicrobial properties of the peptide.

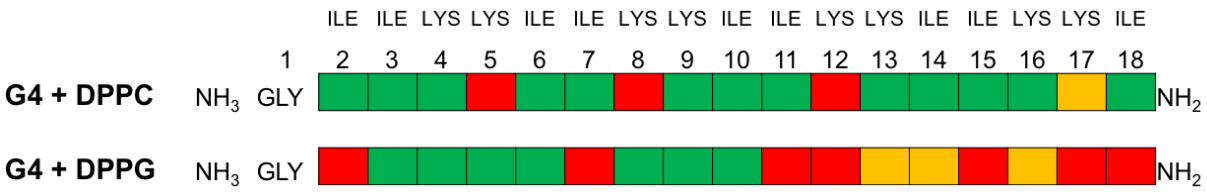


Figure 7.17: Probability of secondary structure for the different G4 residues. The probability is represented by coloured squares. See the caption of Figure 7.8 for the colours legend.

7.5 Conclusion and future work

The study presented in this chapter introduces a computational model for an antimicrobial peptide with the general sequence $G(IIKK)_4I{-}NH_2$ (G4). This study complements and discusses the experimental results presented in ref. 8.

The computation of the peptide secondary structure from its backbone torsional angles confirms the hypothesis of a non-ordered coil structure in water. Initial configurations with the G4 peptide in an α-helical structure unfold from this configuration, sampling the coil region of the Ramachandran plot. The unfolding process preferentially happens at the peptide N and C termini, while the central IIKK residues are more stable in a helical conformation. Adding 137 mM NaCl solution seems to favour the formation of the extended configuration and stabilise the peptide secondary structure.

The mass rescaling and replica exchange molecular dynamic techniques were also used to enhance the conformational sampling of the peptide in water. By rescaling the solvent masses by a factor of 0.1, the peptide structure converged faster to its extended state. REMD helped the peptide to escape from the local energy minimum of helical conformations.

A preliminary study of the G4 peptide interacting with DPPC and DPPG lipid monolayers shows that the peptide exhibits a preferential interaction with the DPPG monolayer. In addition, when interacting with the anionic lipids the probability of helical conformations is higher. In particular the residues near the amidated C terminus tend to assume helical conformations, suggesting a microscopic explanation for the amidations of wild-type peptides.

Further analysis is necessary to fully unravel these questions. Specifically:

- Despite giving a first idea of the peptide secondary structure, the study of the Ramachandran diagram presents several limitations. The link between secondary structure and the regions of the graph is not unique in literature and several classifications exist (*e.g.* STRIDE[40] and DSSP[41]). Further work could focus on a more detailed analysis than the one presented in this chapter.

- The mass rescaling technique has previously shown to further improve the sampling when the amino acid side-chain masses are also reduced[24]. A future plan is to test this on the G4 peptide, as well.

- The REMD simulations presented here have two limitations. The first is the narrow temperature range studied here. Future calculations will increase the maximum value of the temperature range to 400 K. Moreover, the 50 ns simulation time scale might not be long enough to ensure the convergence of the peptide folding behaviour and allow the system to climb out of local energy minima. Longer simulations should also be performed.

- The simulations of the interaction between monolayers and G4 are still at a preliminary step. The effect of applying different surface tensions, γ_t, remains to be tested. Moreover, further simulations of the lipid monolayers with a number of peptides reproducing the experimental peptide concentration will allow the neutron experimental model to be tested and the effect of mutual peptide-peptide interactions in promoting α-helical conformations and penetrating the lipid monolayer to be investigated.

- A recent study by Ciumac *et al.*[42] showed that in addition to the effect of the lipid head group charge, the saturation of the lipid chains has also an impact on the selective interaction of lipid monolayers and G4 peptide. This remains to be tested and explained with computer simulations.

Bibliography to chapter 7

[1] J. M. Blair, M. A. Webber, A. J. Baylay, D. O. Ogbolu and L. J. Piddock, *Nature Reviews Microbiology*, 2011, **47**, 4055–4061.

[2] S. B. Levy, *Scientific American*, 1998, **278**, 46–53.

[3] C. Ventola, *Pharmacy and Therapeutics*, 2015, **40**, 277–283.

[4] ECDF, *European Centre for Disease Prevention and Control. Antimicrobial resistance surveillance in Europe 2016. Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)*, Ecdf technical report, 2017.

[5] L. Zhang and R. L. Gallo, *Current Biology*, 2016, **26**, R14–R19.

[6] K. A. Brogden, *Nature Reviews Microbiology*, 2005, **3**, 238–250.

[7] J. Hu, C. Chen, S. Zhang, X. Zhao, H. Xu, X. Zhao and J. R. Lu, *Biomacromolecules*, 2011, **12**, 3839–3843.

[8] D. Ciumac, R. A. Campbell, H. Xu, L. A. Clifton, A. V. Hughes, J. R. Webster and J. R. Lu, *Colloids and Surfaces B: Biointerfaces*, 2017, **150**, 308–316.

[9] J. W. Carter, A. S. Tascini, J. M. Seddon and F. Bresme, *Computational Tools for Chemical Biology*, The Royal Society of Chemistry, 2018, ch. 2, pp. 39–68.

[10] M. B. Ulmschneider, J. P. F. Doux, J. A. Killian, J. C. Smith and J. P. Ulmschneider, *Journal of the American Chemical Society*, 2010, **132**, 3452–3460.

[11] J. P. Ulmschneider, J. C. Smith, S. H. White and M. B. Ulmschneider, *Journal of the American Chemical Society*, 2011, **133**, 15487–15495.

[12] Y. Wang, C. H. Chen, D. Hu, M. B. Ulmschneider and J. P. Ulmschneider, *Nature Communications*, 2016, **7**, 1–9.

[13] L. Thøgersen, B. Schiøtt, T. Vosegaard, N. C. Nielsen and E. Tajkhorshid, *Biophysical Journal*, 2008, **95**, 4337–4347.

[14] C. Kandt, W. L. Ash and D. Peter Tieleman, *Methods*, 2007, **41**, 475–488.

[15] X. He, M. Lin, B. Sha, S. Feng, X. Shi, Z. Qu and F. Xu, *Scientific Reports*, 2015, **5**, 12808.

[16] M. Shahmiri, M. Enciso and A. Mechler, *Scientific Reports*, 2015, **5**, 1–14.

[17] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeerschd, E. Zurek and G. R. Hutchison, *Journal of Cheminformatics*, 2012, **4**, 1–17.

[18] F. W. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott and I. G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual*

and User Guide, 1996, pp. 1–1042.

[19] R. B. Best, N. V. Buchete and G. Hummer, *Biophysical Journal*, 2008, **95**, L7–L9.

[20] E. A. Cino, W. Y. Choy and M. Karttunen, *Journal of Chemical Theory and Computation*, 2012, **8**, 2725–2740.

[21] D. van der Spoel, E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, K. A. Feenstra, R. van Drunen and H. J. C. Berendsen, *GROMACS user manual*, 2010, p. 7778.

[22] G. Bussi, D. Donadio and M. Parrinello, *Journal of Chemical Physics*, 2007, **126**, 014101.

[23] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *The Journal of Chemical Physics*, 1995, **103**, 8577–8593.

[24] I. C. Lin and M. E. Tuckerman, *Journal of Physical Chemistry B*, 2010, **114**, 15935–15940.

[25] A. Patriksson and D. van der Spoel, *Physical Chemistry Chemical Physics*, 2008, **10**, 2073.

[26] O. Berger, O. Edholm and F. Jähnig, *Biophysical Journal*, 1997, **72**, 2002–2013.

[27] D. Mohammad-Aghaie and F. Bresme, *Molecular Simulation*, 2016, **42**, 391–397.

[28] D. Mohammad-Aghaie, E. Mace', C. A. Sennoga, J. M. Seddon and F. Bresme, *Journal of Physical Chemistry B*, 2010, **114**, 1325–1335.

[29] I. C. Yeh and M. L. Berkowitz, *Journal of Chemical Physics*, 1999, **111**, 3155–3162.

[30] J. L. Abascal and C. Vega, *The Journal of chemical physics*, 2005, **123**, 234505.

[31] S. Izadi, R. Anandakrishnan and A. V. Onufriev, *The Journal of Physical Chemistry Letters*, 2014, **5**, 3863–3871.

[32] M. Javanainen, A. Lamberg, L. Cwiklik, I. Vattulainen and O. H. Ollila, *Langmuir*, 2018, **34**, 2565–2572.

[33] I. Kufareva and R. Abagyan, *Methods in molecular biology (Clifton, N.J.)*, 2012, **857**, 231–57.

[34] F. Ioannou, E. Leontidis and G. Archontis, *The Journal of Physical Chemistry B*, 2013, **117**, 9866–9876.

[35] M. V. Fedorov, J. M. Goodman and S. Schumm, *Physical Chemistry Chemical Physics*, 2007, **9**, 5423–5435.

[36] M. V. Fedorov, J. M. Goodman, V. V. Kolombet, S. Schumm and I. M. Socorro, *Journal*

of Molecular Liquids, 2009, **147**, 117–123.

[37] K. B. Rembert, J. Paterová, J. Heyda, C. Hilty, P. Jungwirth and P. S. Cremer, *Journal of the American Chemical Society*, 2012, **134**, 10039–10046.

[38] C. E. Dempsey, P. E. Mason and P. Jungwirth, *Journal of the American Chemical Society*, 2011, **133**, 7300–7303.

[39] X. Daura, R. Suter and W. F. van Gunsteren, *The Journal of Chemical Physics*, 1999, **110**, 3049.

[40] M. Heinig and D. Frishman, *Nucleic Acids Research*, 2004, **32**, W500–W502.

[41] W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577–2637.

[42] D. Ciumac, R. A. Campbell, L. A. Clifton, H. Xu, G. Fragneto and J. R. Lu, *ACS Omega*, 2017, **2**, 7482–7492.

Chapter 8

Fats' split personality

An outreach and classroom activity to introduce the amphiphilicity of lipids

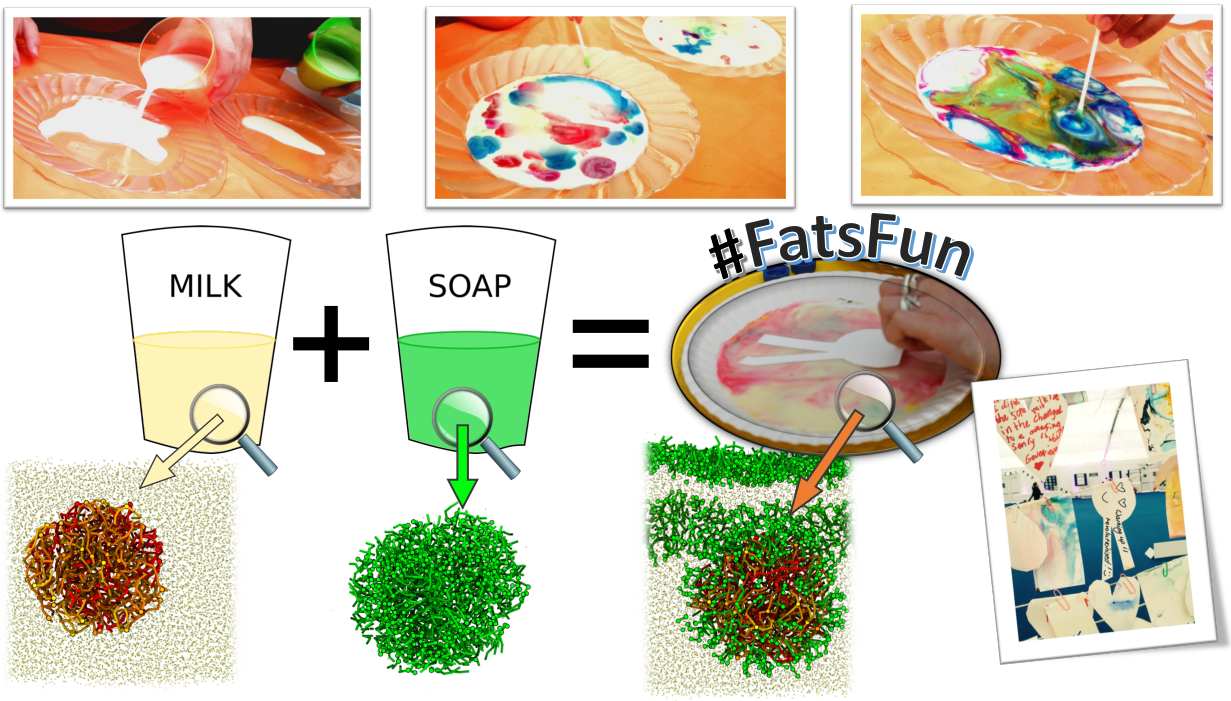


Figure 8.1: Title page figure for chapter 8. This image was used as the Table of contents for Tascini *et al*, *J. Chem. Educ.*, 2018, submitted

8.1 Introduction

Lipidomics is the science of lipids, of which fats are one common type[1]. Despite being one of the four major classes of biomolecules, the importance of lipids is often overlooked and, consequently, there is not much information about their functions in school curricula. Furthermore, the general public tends to have a bad impression of fats, due to the emphasis placed on dangers to health by the public press and social media. Few people know that fats are essential building blocks of cells and also vital components in the cosmetics industry. Therefore, it is necessary to introduce the fundamental importance of lipids to the public and engage school students on this topic. Some interesting activities have been previously reported which provide students with both theoretical knowledge and practical experience with lipids[2, 3, 4, 5, 6]. Most of these works have been designed for undergraduates or GCSE and A–level students students studying biochemistry. They involve lipid extraction from food and subsequent analysis via chromatography, spectroscopy or microscopy. However, since they require professional laboratory equipment and basic biochemistry knowledge, these experiments are less suitable for the general public or younger students.

This chapter describes a simple, safe, inexpensive and easy to perform outreach activity to communicate to the general public, including children over 6 years old, why and how lipid molecules play a fundamental role in all living organisms and in our bodies. The activity is called "Fats' split personality" to describe in colloquial language the amphiphilic nature of lipid molecules which is vital to their biological functions. It investigates the interactions between amphiphilic milk lipids and soap surfactants, which are visualised with food colouring (or watercolour paint). Beautiful complex colourful patterns are produced and can be transferred to watercolour paper, which stimulates creativity and inspires interest in science. The pieces of paper, where the artwork was transferred, were also used to collect the comments and feedback of the participants and allowed us to learn how to improve the activity.

The practical activity is accompanied by a discussion and a series of posters highlighting the role of lipids and how their "split personality" gives rise to the observed behaviour. The main novelty of this activity is that the experiment is interpreted at a molecular level via a computational

molecular dynamics simulation. In this simulation, milk is modelled as a mixture of water and triglycerides (the most abundant lipid component in milk[7]), and the soap is modelled as a dodecylphosphocholine (DPC) micelle. The MARTINI coarse-grained forcefield [8] is used, providing a qualitative visualisation of the lipid-lipid interactions.

When presenting the activity in science festivals and schools I have been helped by many volunteers (reported in alphabetic order): Zainab Ahmed, Marie Bachelet, James Carter, Yu Han Chang, Siyuan Chen, Silvia Di Lecce, Gøran Svaland, Christian Jürgensen, Michal Kopytynski, Berardo Mario Manzi, Enrica Mazzon, Suhuur Osman, Oliver Robotham and Shiqi Wang.

8.2 Activity preparation

8.2.1 Materials

The activity requires the materials listed below. Everything is inexpensive and very easy to find.

- Experimental consumables: different types of milk (whole milk, soya milk and skimmed milk), washing-up liquid, cotton buds, food colouring solutions (or watercolour paint solutions), disposable dishes (to hold the milk), plastic pipettes (for dropping the food colouring onto the milk surface), watercolour paper (for printing the artwork from the milk surface), and permanent markers (for leaving notes and feedback).
- Computer simulation: a screen and a device (*e.g.* a laptop, a tablet, a smartphone or a RaspberryPi) to play the video.
- A "Lipid-box": a box containing lipid-rich products for use in the introduction of the activity, including cooking oil, chocolates, nuts, avocados, hand creams, soaps, shampoo, milk and lipsticks.
- Posters and flyers: posters to aid in the introduction and discussion and a take home flyer to remember the activity and encourage people to repeat it and share it with others.

- Other materials: bins for waste dishes and milk, kitchen rolls for wiping up spilt liquids.

Figure 8.2: The stand for the activity in different locations, showing different set-ups. From the top left-hand side in anti-clock-wise order: 1) "Fats' split personality" at Imperial Festival 2016: the picture shows the "Fats' box", the bin for milk waste, the informative material and the artworks hanging to dry). 2) "Fats' split personality" at Imperial Festival 2017: the picture shows the SNAL poster milk, soap and colours ready on the table, the T-shirts designed for the activity are also worn by the presenters. 3) "Fat is fun" in Altafulla (Spain), here the stand was placed outside in a public space (by the beach). 4) "Fats' split personality" at the Fun Family Day organised by the Royal Institution: the picture shows the experiment, the informative material, the artworks hanging to dry, the SNAL poster, the simulation video running on the left.

8.2.2 Preparation before the activity

Watercolour paper was cut into small pieces with different shapes (e.g. heart, T-shirt, lipid molecule, and water droplet) to engage participants. For open day/science festivals, the stand for this activity was decorated with visual materials (see some examples in Figure 8.2 and 8.3), a computer simulation video was played in repeat mode, and a washing line was hung up to display the artworks and allow them to dry. Figure 8.2 shows the stand and the activity materials set-up in several festivals. Different set-ups are required to adapt to different exhibition spaces and

to different audiences. Regardless of location, the use of bright colours was found to be a key factor in catching people's attention and generating interest towards the stand.

8.3 The activity

When possible, the activity began with a brief introduction to lipids and several motivating questions for the audience:

- Where can we find fats in everyday life?
- What are the functions of fats?
- Why do fats function like that?

According to the experience acquired by presenting the activity over the past three years, most of the audience replied to the first question naming a few fats contained in food including cooking oils, meats and dairy products. A few students mentioned the body or even cells, but hardly anyone could link fats with surfactants in household (*e.g.* washing-up liquids) or cosmetic products (*e.g.* hand creams). Therefore, the "lipid-box" was added to the activity and used to reveal unexpected fat-containing products, such as soaps and cosmetics.

Regarding fats' functionality, most answers from the audience were related to nutrition and energy storage for the body. Many students would say fats "keep us warm" or "give us energy". This was usually the moment used to discuss the structural and protective roles of lipids as a key component of cell membranes, and the cleaning function of surfactants in washing-up liquids. The audience was interested by those extremely different roles of fats, and began to think and speculate about the reasons behind it, creating a climax to discuss the amphiphilic molecular structure, so-called "fats' split personality". Figure 8.3 shows the visual material I designed to explain lipid amphiphilicity. The fat molecule is represented as a person (see main figure on the left-hand side of Figure 8.3). The molecule's "head" is in love with water and would like to stay as close as possible to it. However, the molecule's "legs" (*i.e.* the lipid acyl

chains) would like to run away from water. The "split personality" is crucial in forming the structures of surfactant micelles, milk globules and cell membrane bilayers (as shown in the insets on the right-hand side of Figure 8.3). In addition, understanding this important feature of lipid molecules reveals an opportunity to design drug delivery strategies (see inset on the bottom-right of Figure 8.3.)

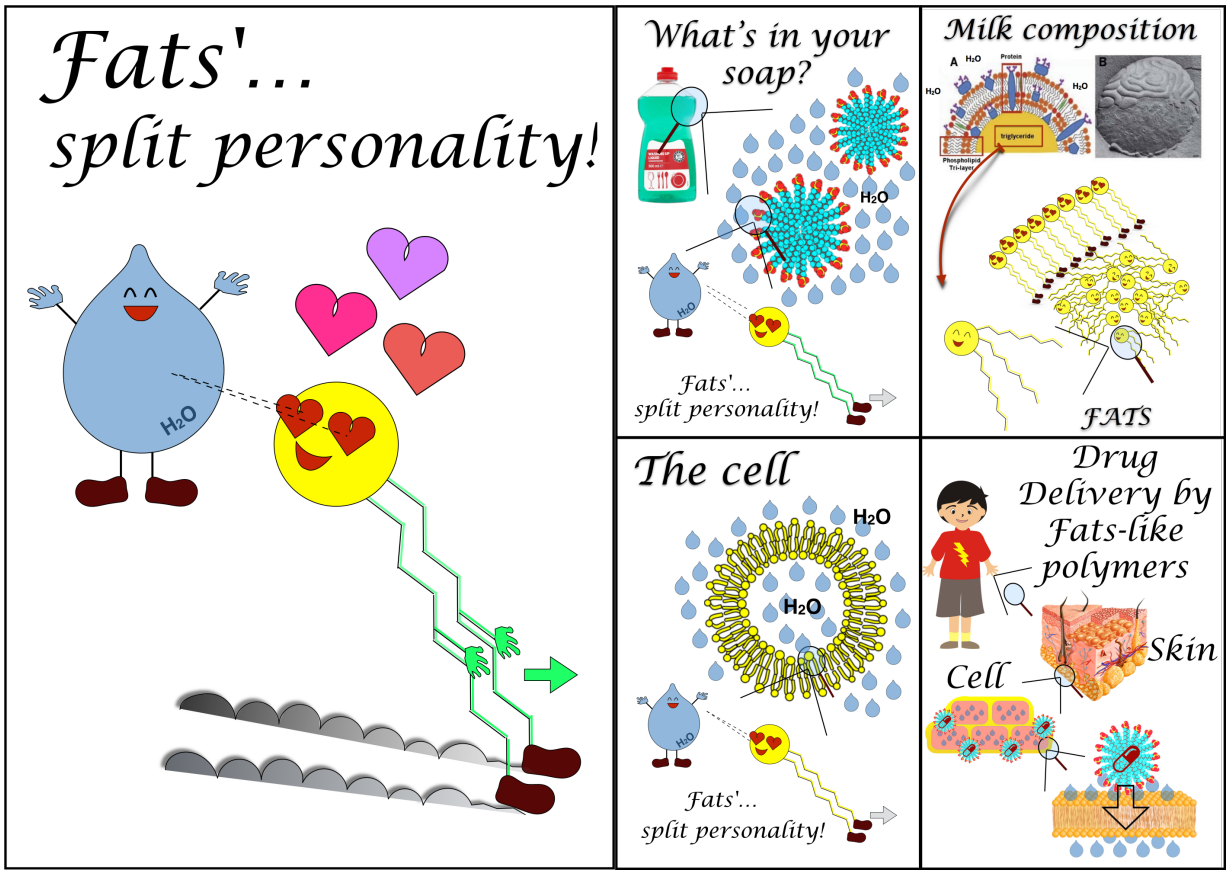


Figure 8.3: A schematic (cartoon-style) representation of the "Fats' split personality" (left-hand side). It describes the fat molecule as a person with two different parts, a "head" which loves water and the "legs" which hate it. This can be directly related to the head and tail groups of lipid molecules and this metaphor makes the amphiphilic properties of lipid molecules easy to imagine. The role of this "split personality" in forming the structures of surfactant micelles and cell membrane bilayers is then shown (right-hand side).

After explaining this concept, the hands-on experiment was carried out to show how different fats with "split personality" interact with each other. Figure 8.4 shows the four steps of the experiment:

1. Milk was poured into a shallow dish. Participants were encouraged to choose different types of milks (soya, skimmed or whole milk) for comparison. This would also avoid any

potential problems for people with allergies to lactose or soya.

2. Several drops of food colouring or watercolour solutions were pipetted onto the milk surface to produce small, concentrated patches of different colours. For young children, it was necessary to explain how to use the plastic pipettes to suck up and release the solutions before this step.

3. The tip of a cotton bud was loaded with the washing-up liquid, and the cotton bud tip was placed onto the surface of the milk, near or within the colour drops.

4. The last step is the experimental result: the colour rapidly spread out in all directions, away from the cotton bud as the washing-up liquid interacts with the milk.

Figure 8.4: The 4-step experiment, visualised by pictures taken during a science festival (Imperial Festival 2017). From left to right, milk poured into a dish, the addition of colour drops, the first touch of a cotton bud with the washing-up liquid and the final colour patterns.

It should be noted that different types of milk result in different effects. In cow's milk, fats are assembled in globular units and the milk is less dense[7]. Therefore, the interaction between the milk fats and the soap surfactant is faster and the colours dissolve more compared to soya milk. Skimmed cow's milk has a lower fat concentration, so the interaction is even faster and stops earlier. We found that the most successful milk for this experiment was the soya milk.

At the end of the experiment, the colourful patterns on the milk surface were transferred to the watercolour paper, which served as a take-home souvenir to remind the participants what they had learned. They were also asked to write on the piece of paper their names and any feedback about the activity. In this way, students had a chance to stop and reflect on what they had learnt from the explanation and experiment. In addition, this allowed us to collect direct feedback of the impact of the activity. Along with the artwork, the participants took home flyers containing the experimental instructions (see Figure 8.5) for repeating the experiment at home and hopefully sharing and explaining it to friends and family.

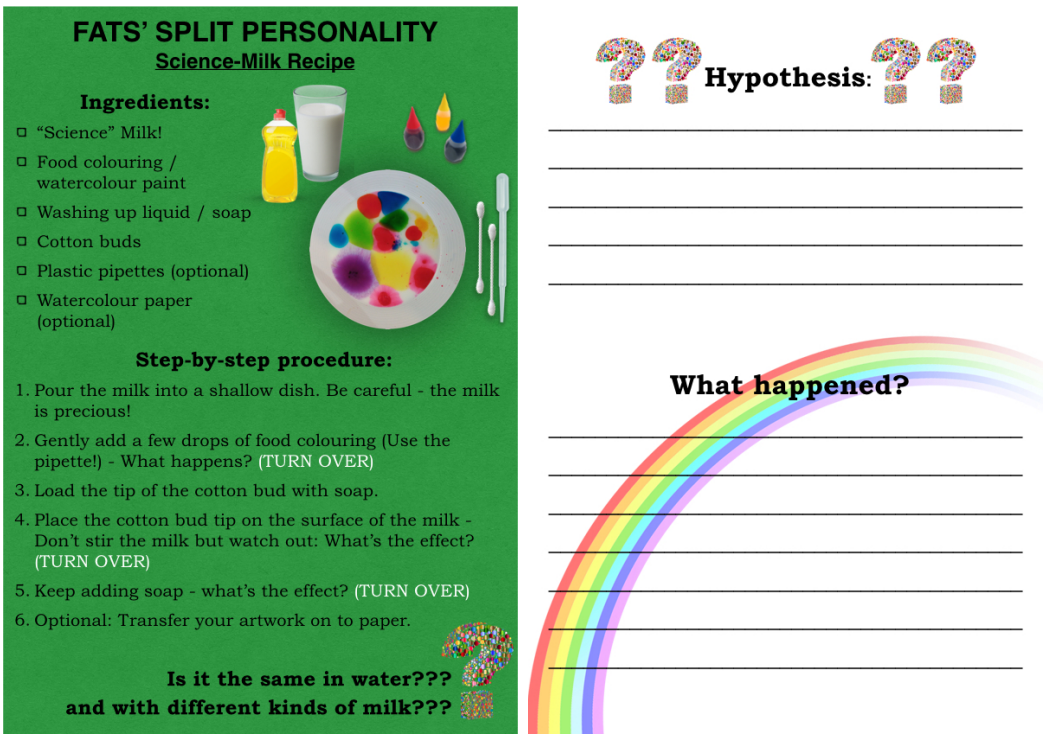


Figure 8.5: Flyer (2 sides) with experimental instructions on the front and a space for participants to write their hypotheses and observations on the back.

8.4 Computer simulations for an insight at the molecular level

Alongside this simple experiment, a video explaining the main features of fats and a computer simulation explaining the molecular dynamics (MD) of milk-soap interactions in water were played to illustrate what was going on at a molecular level, to produce the striking macroscopic effects.

8.4.1 Method

As described in chapter 2, MD is a simulation technique which employs Newtonian mechanics to compute the evolution of molecular systems, based on the interactions between atoms[9]. The milk was represented as a mixture of water and triglycerides (the most abundant lipids in milk[7]) with different degrees of saturation, namely Tripalmitin (TPG), Triolein (TOG) and Trilaurin (TLG). Milk proteins and phospholipids were ignored in these simulations to simplify

the general message for non-specialist members of the public. Soap was represented by a common surfactant: dodecylphosphocholine (DPC). In these simulations, instead of defining the interactions of every single atom, the coarse-grained (CG) MARTINI forcefield[8] was used. In a CG representation, groups of atoms are combined into a smaller number of CG beads. The advantage of this approach is a speed-up in simulation time and the ability to simulate larger systems. The MARTINI forcefield, one of the most widely used biomolecular CG models, is based on a 4:1 mapping of heavy (non-hydrogen) atoms to CG beads and is parametrised to reproduce experimental thermodynamic data. It works especially well for lipids[10], reproducing many properties with semi-quantitative agreement with experiment. According to this forcefield, all particles interact through pairwise short-range Lennard–Jones (LJ) potentials, which are defined as:

$$U_{LJ}(r) = 4\epsilon_{ij}\left[\frac{\sigma_{ij}}{r}^{12} - \frac{\sigma_{ij}}{r}^{6}\right], \tag{8.1}$$

where i and j label two interacting particles, σ_{ij} is the distance at which the potential switches from attractive to repulsive and ϵ_{ij} is the interaction strength. To reproduce the lipid "split personality", with a structure consisting of a water-loving (hydrophilic) head group and water-hating (hydrophobic) tails, in the simulations, the beads composing the lipids differ in their degree of hydrophilicity, which is reflected by different interaction strengths (ϵ_{ij}). A hydrophilic particle will be attracted more strongly to other hydrophilic particles (larger ϵ_{ij}) than to hydrophobic particles and vice versa.

Specifically the molecules used in this work:

- The triglyceride molecules were composed of a site of intermediate hydrophilicity (pink beads in Figure 8.6), used to represent the glycerol ester moiety, and three hydrophobic tails (cyan and purple beads in Figure 8.6), where the number of beads in the tails reflects the different chemical structures of TPG, TOG and TLG. The double bond in the tails is represented by a slightly different bead type (purple) to the saturated parts of the tails (cyan).

- The DPC molecule is composed of two strongly hydrophilic beads (blue and gold beads in Figure 8.6) and a short hydrophobic tail (cyan beads in Figure 8.6).

- The water is modelled by an individual hydrophilic particle representing four 'real' water molecules (green bead in Figure 8.6).

The molecule representations are shown in Figure 8.6.

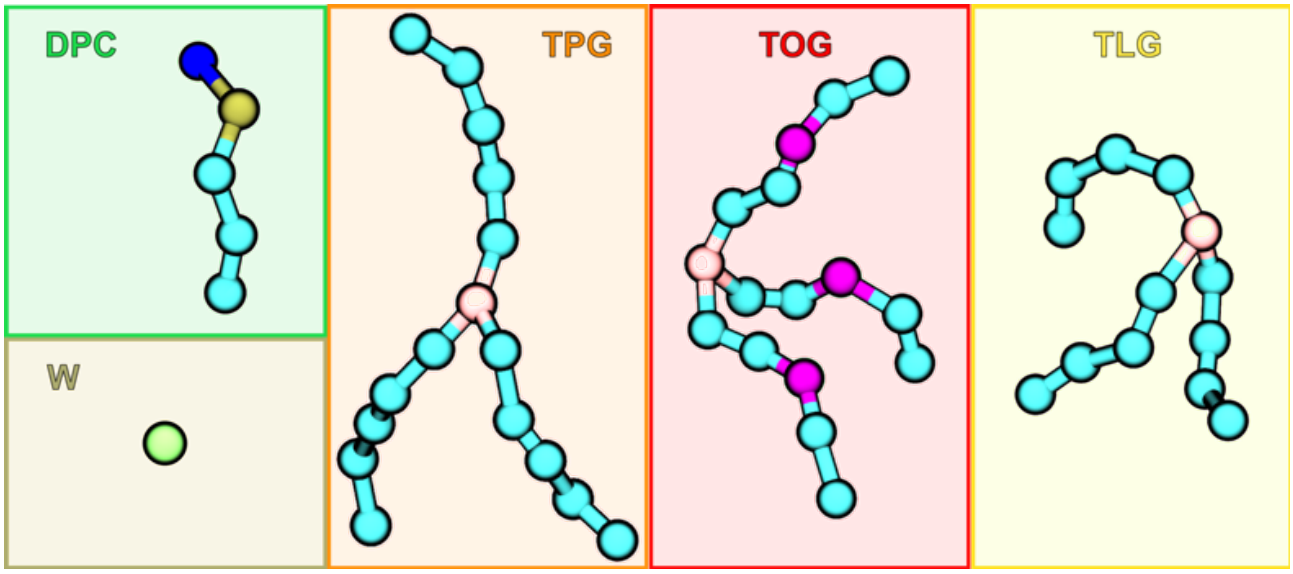


Figure 8.6: Representations of the system components in the MARTINI forcefield: DPC (soap), W (water), TPG, TOG and TLG (milk triglycerides). The coloured box behind the molecules corresponds to the colours of the molecules in Figure 8.7.

Bonded interactions (within a lipid molecule) are modelled by a weak harmonic potential. Harmonic angle potentials provide the appropriate stiffness for the molecules. The topology of the lipids used in this work and the forcefield parameters can be downloaded from the MARTINI website[11, 12].

The simulations were performed using the GROMACS/4.5.5 package[13]. The milk triglycerides (50 TPG, 30 TOG and 20 TLG) were firstly randomly distributed in a box with 10 306 water beads and then equilibrated in the *NPT* ensemble for 3 µs. The temperature was equilibrated to a constant value of 300 K using the velocity rescaling algorithm[14], with a time coupling constant, $\tau = 1$ ps. The pressure was set to 1 bar, with the Berendsen barostat[15] applied isotropically, with $\tau = 5$ ps and compressibility 4×10^{-5} bar^{-1}. Periodic boundary conditions were applied in all three spatial directions. The equations of motion were integrated using a time step of 40 fs. At the end of this equilibration the milk triglycerides clustered together, forming a spherical structure, which resembled the structure of a milk globule[16]. The simulation box

was then elongated along the z direction to create two water/vapour interfaces. An equilibrated structure of the DPC surfactants, self-assembled in a micellar shape, was taken from the open source MARTINI repository[11, 12]. It was added above one of the two interfaces and the system was run in the NVT ensemble for 1 µs. The temperature was again kept constant at 300 K with the same thermostat and coupling constant as before.

8.4.2 Results

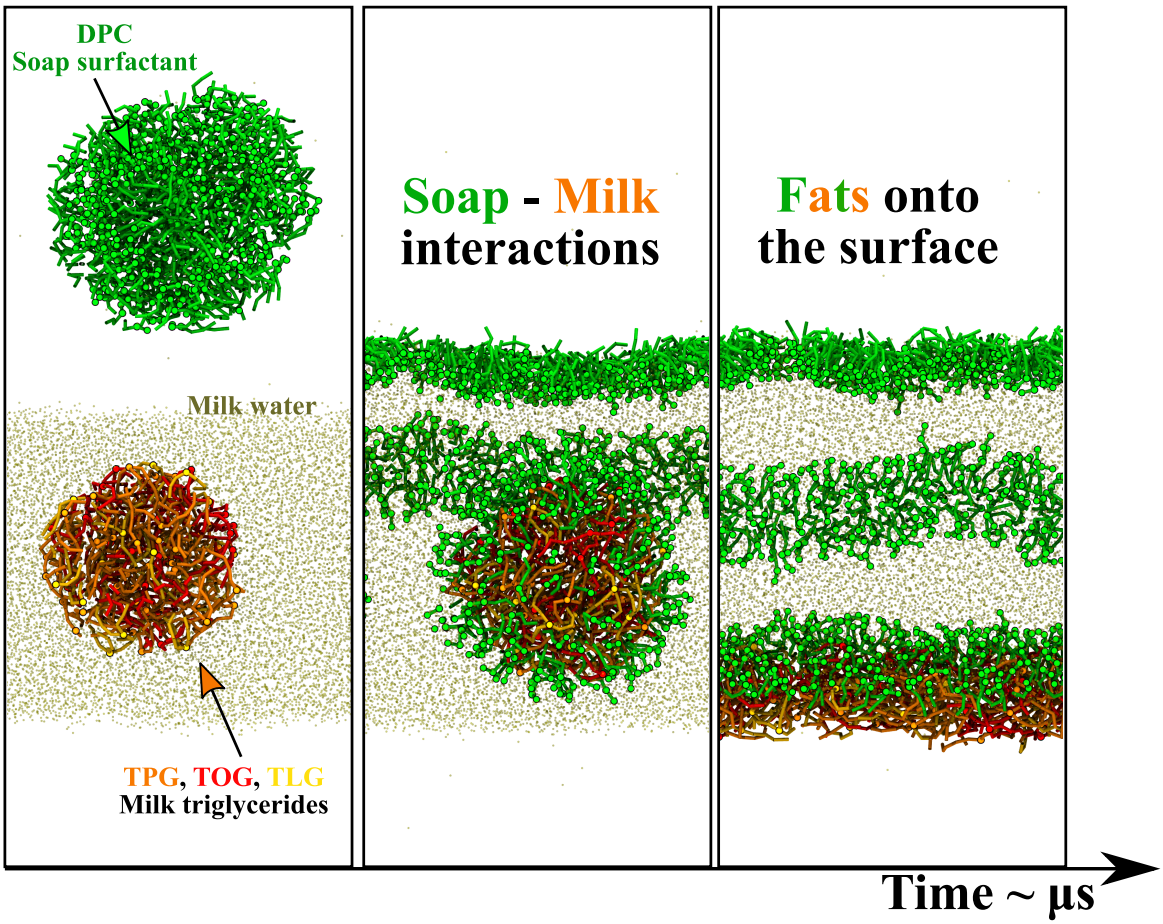


Figure 8.7: Three representative snapshots from the MD trajectory. (Left) Soap surfactant droplet approaching the milk surface. (Middle) Interactions between soap and milk fats, forming a soap monolayer on the milk surface and a soap-milk micelle around the milk triglycerides. (Right) The milk globule is destroyed, and the milk fats are carried to the interface lowering the interfacial surface tension. The DPC soap surfactant is represented in green and different milk triglycerides in orange (TPG), red (TOG) and yellow (TLG). The head group beads are represented as a ball and the hydrophobic tails as sticks. Water in the milk is represented with brown dots.

Figure 8.7 shows the crucial steps of the simulation video. Firstly, the surfactant micelle quickly approaches the milk interface. Due to the strong interaction of DPC molecules with

water, some of the soap lipids reorient, with the hydrophilic beads facing the water interface and the hydrophobic tails pointing away from the water, towards the air. This creates a surfactant monolayer at the interface, and would be the equilibrium situation if, instead of milk, there was only water. However, in this specific case there are also the milk lipids, which the soap starts interacting with, leading to the destruction of the milk globule and transport of the milk lipids to the interface. The combination of these interactions drives the macroscopic effects observed in the experiment and ultimately produces the colourful patterns which form on the milk surface. The formation of a lipid layer on the milk surface can also be observed at the end of the experiment by dropping a small droplet of water from the cotton bud onto the milk surface after adding the washing up liquid. This droplet will run on the milk surface.

8.5 Discussion

The activity was presented in a series of events, listed in Table 8.1.

Event name	Location	Event Type	Target	Dates	N_P
Imperial Festival 2015[17]	South Ken., London (UK)	Science fest.	General public	9 – 10/05/2015	12 000
School day - Imperial Festival 2016[17]	South Ken., London (UK)	Science fest.	Students, 7–15yo	6/05/2016	200
Imperial Festival 2016[17]	South Ken., London (UK)	Science fest.	General public	7 – 8/05/2016	16 000
Imperial Festival 2017[17]	South Ken., London (UK)	Science fest.	General public	6 – 7/05/2017	20 000
Fun Family Day[18]	The RI, London (UK)	Science fest.	General public	13/05/2017	450
The invention room open day[19]	White City, London (UK)	Open day	General public	28/10/2017	700
Fat is fun[20, 21]	Altafulla (Spain)	Public event	General public	9/09/2017	80
Creativity Quarter[22]	South Ken., London (UK)	Open Day	Students, 17 yo	17/11/2017	100
The Avenue Primary Academy	London (UK)	Class demo.	Students, 7 yo	22/02/2018	25
Platanos College	London (UK)	Class demo.	Students, 12 yo	24/04/2017	22
City of London Academy Southwark	Surrey (UK)	Class demo.	Students, 16 yo	27/06/2017	12

Table 8.1: List of events where the activity was presented. N_P is the number of participants.

It was presented for the first time at the Imperial Festival[17] in 2015. The festival has received more than 10 000 visitors every year from 2015 to 2017, with a record of more than 20,000 visitors in 2017. Imperial festival provided the opportunity to engage with many people of different backgrounds and different ages, from primary school students to pensioners. With hundreds of visitors, our stand was continuously full of participants, listening to the explanation, performing the experiment and producing hundreds of artworks every day. Whilst presenting the activity in the last three years, in different environments and for a wide range of participants,

I have improved it by developing more informative materials and enhancing the efficacy of the communication.

In 2017, after the success of the 2017 edition of the Imperial festival, the activity was also presented in the Fun Family Day edition, "The chemistry of life"[18], organised by the Royal Institution (RI) of Great Britain. The event registered 450 attendees, with over 250 children, all engaging with the activity. In addition, in 2017 we joined the Imperial Fringe event "The Invention Room Open Day"[19] and the "Creativity quarter"[22] organised by the South Kensington community. For the former event, the activity was adapted to fit in the general topic of "Free creative fun to bring out the maker in you!". The stand was titled "Make art from fats" and in addition to the setup described earlier I also added a video showing Zainab Ahmed, from the membrane biophysics group of Imperial College, creating and analysing microfluidic devices in the laboratory. At the "Creativity quarter" event our activity got very positive feedback that was reported in the Imperial College newsletter[22]. From the words of Chris Markou, a visiting student from Caterham High School: "*So far my favourite activity has been seeing how the fat in soap reacts with milk. The colour and the way that it moves is so unusual, and it's fascinating*".

The demonstration was also adapted for the Spanish public and presented in Altafulla (Spain) during the 6^{th} Workshop on Biomaterials and Membranes in 2017[21, 23], with the aid of a Spanish-speaking volunteer. The activity, titled "La personalidad dividida de la grasa", was presented along the beach, transforming a public area into a space for public learning and scientific discussion. It allowed us to test the activity with a different type of audience to scientific festivals, where people deliberately choose to go to because they already have a strong interest in science. In this setting it was crucial to attract the audience's interest with questions and to explain it was a free event with a scientific demonstration. Again in this case, the activity was well received, and of interest for most of the people passing by ($\sim$ 80 people).

Figure 8.8 shows the engagement of the general public with the stand.

To collect feedback from the participants, they were encouraged to write their comments on the final artwork. All comments were extremely positive, including "I think it was really cool to

Figure 8.8: Three pictures showing the engagement of the general public with the activity in different settings. From left to right: Imperial Festival 2017, The invention room open day, Altafulla (Spain).

see how different fats react to a specific fat"; "This was very fun and interesting. I never knew fats could be so cool"; "Washing-up is revolutionized"; "I found this experiment highly engaging! Thank you."; "This was totally awesome, chemistry Rocks!"; "Fun and informative"; "Thank you for the explanation, I feel like I've learnt a lot"; "Science is great!". This has highlighted three main outcomes of the activity: effectively communicating the basic principles of how lipids work and their fundamental roles in biology; engaging for people of all ages and backgrounds; inspiring the interests of the general public towards science. Photographs of some of the artwork and comments are shown in Figure 8.9.

Figure 8.9: Artworks produced at the end of the milk experiment, showing some of the feedbacks from the general public.

Pictures of the artwork and comments were shared on Twitter, with the hashtag #FatsFun, highlighting the engagement of the audience and further increasing the visibility of the events on social media. During Imperial Festival 2017, an image contest of the artworks was run to promote the outreach activity on social media. Participants were encouraged to share their experiences on Twitter or Facebook with the hashtag #Fatsfun to win our specially designed T-shirt (with the slogan "I'm made of fats"). The flyer to promote the contest, showing the designed T-shirt, is presented in Figure 8.10. Other media coverage included the newsletters of the Departments of Chemistry and Chemical Engineering at Imperial College London[24] and the ITN-SNAL website[25].

Figure 8.10: Flyer to advertise the picture competition, produced to promote the outreach activity and increase visibility on social media. The T-shirt was awarded as the competition prize and the same T-shirts were also worn by volunteers working on the stand (see yellow T-shirts in Figures 8.2 and 8.8).

The activity has also been adapted to target students of local primary and secondary schools (see list in Table 8.1). Working with students of different ages required us to adjust the contents to match their knowledge and vocabulary. In general, regardless of age, the students did not have much previous knowledge of the importance of lipids. Therefore, it was very effective to introduce the amphiphilicity property of lipids as a "split personality". The hydrophilic head was interpreted as "loving water", while the hydrophobic tail as "hating water". This interpretation made it easy for audiences of all ages to understand the scientific concept of

amphiphilicity. Compared to scientific festivals, the advantage of working in a classroom is the possibility of dividing the students into small groups and doing the activity at the same time in a controlled environment, using a presentation showing the steps of the experiment and the MD video played at the end for discussions. An example is shown in Figure 8.11.

Figure 8.11: Outreach activity at Platanos College.

A limitation of the activity, which could be improved in the future, is that the transfer of colourful patterns to paper to generate the artwork does not always work. First of all, the type of paper and colour is crucial. The best results were produced by using watercolour paper, due to its thickness and the rough surface, and quick-setting watercolour paints. Watercolour solutions need to be prepared in advance by properly mixing the paint with water to avoid precipitation. Secondly, it was noticed that participants often stirred the milk once the reaction slowed down, which destroys the colour patterns produced during the experiment. So, it is important to remind the audience not to stir their milk once the fun is over. Finally, compared to soya and almond milks, the use of cow's milks led to a faster reaction and more dissolved colours which caused the colour to fade, making it more difficult to transfer to paper.

8.6 Conclusion

An outreach activity has been designed to communicate the basics of lipodomics to the general public and school students. It introduces the amphiphilicity of lipid molecules in colloquial language, as a "split personality". It consists of a discussion about lipids, followed by a simple experiment, along with a computer simulation to demonstrate the principles and behaviour at a molecular level. The activity has been tuned and developed for different settings (*e.g.*

science festivals of different sizes, public spaces, schools), audiences (from young students to the elderly) and countries (UK and Spain). It received excellent feedback in terms of both the number of attendees and the collected comments, which have highlighted three main outcomes of the activity: effectively communicating the basic principle of how lipids work and their fundamental role in biology; engaging for people of all ages and backgrounds; inspiring the interests of the general public towards science.

Bibliography to chapter 8

[1] P. Laggner and J. P. Slotte, *European Journal of Lipid Science and Technology*, 2007, **109**, 1237–1240.

[2] V. G. Almendro Vedia, P. Natale, S. Chen, F. Monroy, V. Rosilio and I. López-Montero, *Journal of Chemical Education*, 2017, **94**, 644–649.

[3] A. M. Hartel and A. C. Moore, *Journal of Chemical Education*, 2014, **91**, 1702–1705.

[4] A. C. Rowat, K. A. Hollar, H. A. Stone and D. Rosenberg, *Journal of Chemical Education*, 2011, **88**, 29–33.

[5] R. Abu-Much, S. Basheer, A. Basheer and M. Hugerat, *Journal of Chemical Education*, 2013, **90**, 1207–1211.

[6] K. Bendinskas, B. Weber, T. Nsouli, H. V. Nguyen, C. Joyce, V. Niri and T. W. Jaskolla, *Journal of Chemical Education*, 2014, **91**, 1697–1701.

[7] T. Huppertz, A. L. Kelly and P. F. Fox, *Dairy Fats and Related Products*, 2009, pp. 1–27.

[8] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *The Journal of Physical Chemistry B*, 2007, **111**, 7812–7824.

[9] J. W. Carter, A. S. Tascini, J. M. Seddon and F. Bresme, *Computational Tools for Chemical Biology*, The Royal Society of Chemistry, 2018, ch. 2, pp. 39–68.

[10] S. J. Marrink and D. P. Tieleman, *Chemical Society Reviews*, 2013, **42**, 6801–6822.

[11] *Martini lipodome*, `http://cgmartini.nl/index.php/force-field-parameters/lipids`.

[12] T. A. Wassenaar, H. I. Ingólfsson, R. A. Böckmann, D. P. Tieleman and S. J. Marrink,

Journal of Chemical Theory and Computation, 2015, **11**, 2144–2155.

[13] D. van der Spoel, E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, K. A. Feenstra, R. van Drunen and H. J. C. Berendsen, *GROMACS user manual*, 2010, p. 7778.

[14] G. Bussi, D. Donadio and M. Parrinello, *Journal of Chemical Physics*, 2007, **126**, 014101.

[15] H. J. Berendsen, J. P. Postma, W. F. van Gunsteren, A. Dinola and J. R. Haak, *The Journal of Chemical Physics*, 1984, **81**, 3684–3690.

[16] P. Walstra and R. Jenness, *Dairy chemistry and physics*, 1984, p. 467.

[17] *Imperial Festival | Imperial College London*, http://www.imperial.ac.uk/festival/about/.

[18] *Family Fun Day: The Chemistry of Life | The Royal Institution: Science Lives Here*, http://www.rigb.org/whats-on/events-2017/may/public-family-fun-day--the-chemistry-of-life.

[19] *The Invention Rooms Open Day*, http://www3.imperial.ac.uk/newsandeventspggrp/imperialcollege/eventssummary/event_18-8-2017-10-31-46.

[20] V. Baulin, *ITN SNAL - Marie Curie Initial Training Network - Smart Nano-objects for Alteration of Lipid bilayers*, https://itn-snal.net/.

[21] B. M. Manzi and A. S. Tascini, *Altafulla: Fat is fun*, https://itn-snal.net/2018/03/altafulla-fat-fun/.

[22] J. Rawling, *School students explore science and creativity at Creative Quarter 2017 | Imperial News | Imperial College London*, http://www.imperial.ac.uk/news/183225/school-students-explore-science-creativity-creative/.

[23] V. Baulin, *SNAL workshop web page*, http://meeting.softmat.net/program.

[24] M. Tortelli, *Chemistry research showcased at the Imperial Festival | Imperial News | Imperial College London*, http://www.imperial.ac.uk/news/172332/chemistry-research-showcased-/.

[25] S. Wang and A. S. Tascini, *SNAL @ Imperial Festival - ITN SNAL - Marie Curie Initial Training Network*, https://itn-snal.net/2015/09/snal-imperial-festival.

Chapter 9

Conclusion

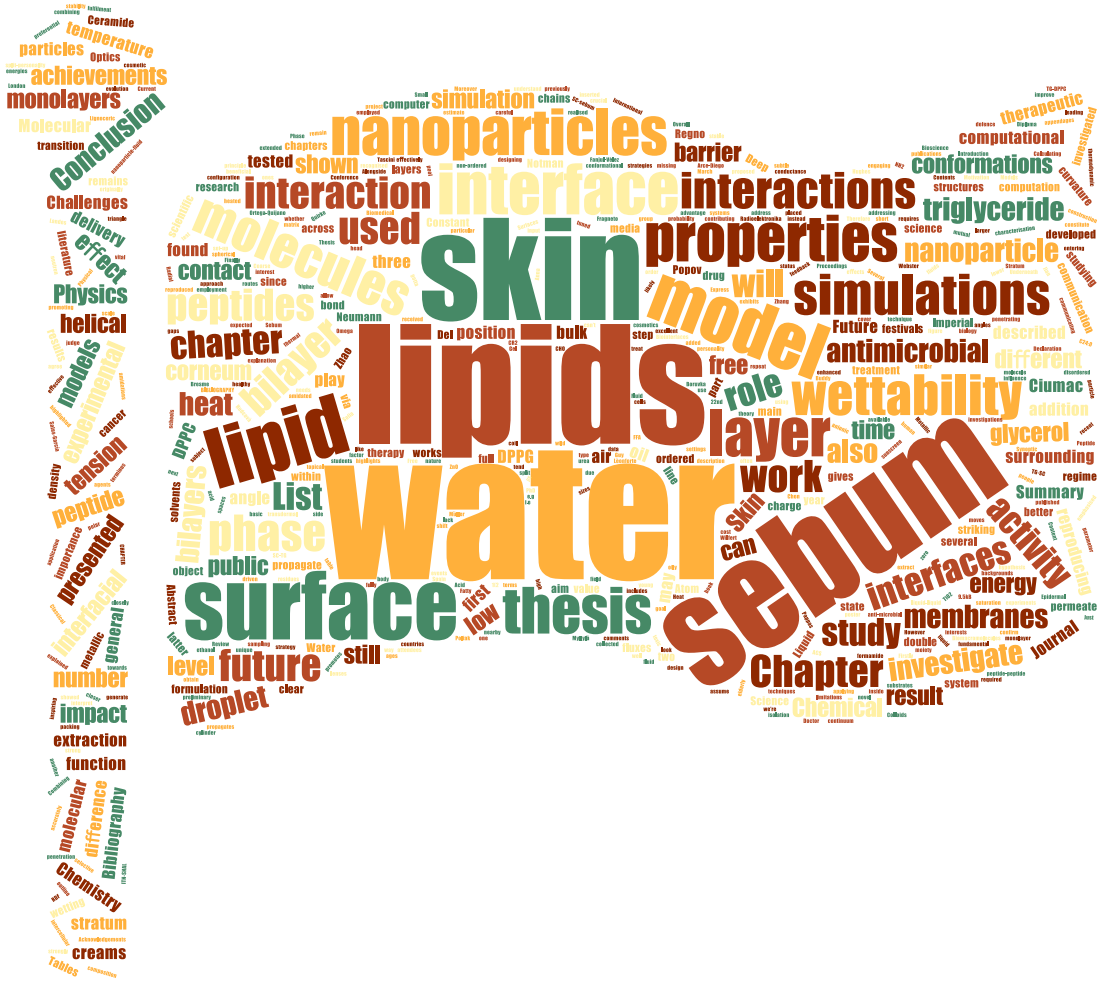


Figure 9.1: Title page figure for chapter 9. Word cloud realised using the words in this chapter.

9.1 Summary of thesis conclusions

9.1.1 Research outcomes

This work gives insights into the properties of the skin surface at the microscopic level by introducing a model of the sebum layer, studying its interactions with water and SC lipids and analysing the skin surface wettability. The interfacial properties of metallic nanoparticles and the conformations of an antimicrobial peptide are also studied. The future challenge is to combine the nanoparticles and peptides with skin lipids (and also phospholipid bilayers) to explore their interactions at the molecular level. A detailed overview of the main outcomes of this thesis is presented below.

For the first time, a computational model of sebum oil has been introduced. The sebum was represented by its most abundant component, the triglyceride tri-cis-6-hexadecacenoin (TG). The model was tested against the structural and interfacial properties of TG, closely reproducing the density of triglycerides with similar structures, the surface tension of the sebum oil at the air interface and the collapse pressures of TG-monolayers at the air-water interface.

The development of this model has helped to shed light on several important questions:

1. **What impact does the sebum layer have on skin hydration?**

   The sebum model was found to form an ordered structure in bulk with glycerol groups clustered together to form an extended network. Water molecules can permeate into the TG phase by moving through this network of glycerol moieties. Evidence for this result was found by studying TG at the water interface (in chapter 3) and the sebum-SC-water system (chapter 5) where water moves to the opposite side of the bilayer via the sebum layer and, specifically, via this network of clustered glycerol groups. The free energy barrier to water permeation was computed. The magnitude of the activation energy, $\Delta G \approx 9.5 k_B T$, is relatively low and suggests that water molecules can permeate the TG layer in a short time scale ($\approx 100\,\text{ns}$). The estimated difference in free energy between water molecules in bulk water and water molecules in the TG phase is $\approx 7 k_B T$,

and reflects the water solubility in the sebum oil. This result suggests that triglyceride molecules may play a role in regulating the water concentration on the skin surface by enabling a controlled permeation of water through the skin surface and thereby influencing the hydration of the outermost layers of the skin.

2. **Why are there SC lipids in the sebum - *i.e.* what interaction is there between the sebum and the SC?**

   Experimental studies have shown that stratum corneum lipids are present in the sebum layer[1]. This observation suggests the existence of strong interactions between the sebum molecules and the skin lipids, with the possibility of inducing detachment of SC lipids that then penetrate the sebum. The results included in this thesis (chapters 4 and 5) demonstrate that TG interacts with SC lipids in such a way that the latter can be extracted directly from SC bilayers. The microsecond MD simulations presented here have identified the pathways leading to extraction as well as the energetic cost associated with the dissolution of the lipids in the sebum phase. A different mechanism and behaviour was found for each of the three different SC lipid species. Free fatty acids moved through the same TG glycerol group network as water, whereas cholesterol tended to remain associated to the bilayer and ceramide remained within the membrane, at most inserting one tail into the sebum. Some sebum TG molecules also inserted tails into the SC bilayer. However, despite the SC lipids extraction and TG anchoring to the SC bilayer, the bilayer was found to be stable over time, with most of the lipids remaining inside the bilayer.

3. **How wettable is the SC?**

   Despite the fact that most of the previous computational studies of SC membranes use fully hydrated bilayers, the water concentration in the SC layer is very low[2] and the location of water within the SC is still not clear. The wettability of the SC bilayer at the sebum interface has been investigated in this thesis (chapter 5) with a novel simulation set-up: a water droplet was placed at the sebum-SC interface and the contact angles of the droplet were measured. The stability of the droplet indicates a low affinity of the skin surface for water, with the interface in the partial wetting regime (contact angle

$\approx 100°$). The difference with DPPC bilayers is striking: the water spreads at the TG-DPPC interface (contact angle = 0°), resulting in a full wetting regime, showing the crucial role of SC lipids in determining the skin wettability. This result agrees with the work of Del Regno and Notman[3], in which water forms water pools within SC lipid bilayers, and should be considered in future simulations of SC lipids.

4. **How do heated nanoparticles and antimicrobial peptides behave?**

   The aim of my project was also to look at therapeutic agents and their interactions with lipid membranes. Firstly, since nanoparticles are currently being considered for skin cancer treatment by photothermal therapy[4, 5] and in suncream formulations[6, 7], MD simulations of a nanoparticle immersed in a fluid were used to investigate how heat fluxes propagate from nanoparticles to the surrounding media. Our analysis (chapter 6) reveals a subtle interplay of nanoparticle size and interfacial interactions in determining the efficiency of the heat transfer across nanoparticle interfaces. Small nanoparticles are better heat conductors than larger ones and the strength of the dependence with curvature is regulated by the degree of hydrophilicity of the particle.

   Secondly, in collaboration with our experimental partners (Jian Lu and co-workers[8, 9]) a computational model for their antimicrobial peptide, G4, has been introduced, providing molecular insight into the selectivity of G4 for negatively charged membranes and explaining its antimicrobial activity. Classical MD and enhanced conformational sampling techniques showed the transition from α-helical conformations to a non-ordered coil in water. When interacting with DPPC and DPPG lipid monolayers, G4 exhibits a stronger interaction with DPPG. In addition, when interacting with the anionic lipids the probability of helical conformations is higher. In particular the residues nearby the amidated C terminus tend to assume helical conformations, suggesting a microscopic explanation for the advantage of the peptide amidations of wild type peptides.

   Having investigated these two systems in isolation, the next step would be to study their interactions with our sebum-SC model. Combining these two models with the sebum-SC system remains the subject of future work.

Overall, it is hoped that this work makes a useful contribution to the field since the interaction between sebaceous and skin lipids is a contributing factor for the healthy state of the skin and for designing better strategies of transdermal drug delivery.

9.1.2 Science communication

Alongside scientific research, this thesis also includes the description of an outreach activity, which I have developed and presented at several events and scientific festivals during my PhD. This activity promotes the general public's understanding about lipid molecules, and highlights the importance of their amphiphilic nature (here described as their "split personality") in biological cells, cosmetics and soaps, and in designing strategies for drug delivery.

The activity has been tuned and developed for different settings (*e.g.* science festivals of different sizes, public spaces, schools), audiences (from young students to the elderly) and countries (UK and Spain). It has received excellent feedback both in terms of the number of attendees and the collected comments, which has highlighted three main outcomes of the activity: effectively communicating the basic principle of how lipids work and their fundamental role in biology; engaging for people of all ages and backgrounds; inspiring the interests of the general public towards science.

9.2 Future challenges

The work described in this thesis could be extended to further investigations:

- The research on the sebum triglyceride (TG) has emphasised the importance of addressing the impact of the position and degree of unsaturation of the chains on triglyceride properties. There is still a lack of a clear understanding of the role that the double bond position has in determining the interfacial properties of TG fluids. Here it has been hypothesised that a shift of the double bond to a position closer to the polar part of the molecule will result in a lower surface tension at the air and water interfaces. This effect

would likely be driven by the influence of the double bond position on the packing and the conformation of the TG molecules at the interface. Further experimental and computational studies are required to confirm this hypothesis, which will be the subject of future work.

- A full characterisation of the Neumann triangle construction[10] used to study the wettability of the TG-SC interface requires the computation of the surface tensions of the three interfaces: γ_{TG-W}, γ_{SC-W} and γ_{SC-TG}. However, while the value of γ_{TG-W} is available from the simulation presented in chapter 3, the value of γ_{SC-W} and γ_{SC-TG} needs to be computed to address whether the continuum theory of wettability is effective in reproducing the properties of the SC-TG interface. This will be the aim of future work.

- The same approach, used to investigate the wettability of the SC bilayer to water, could be reproduced to investigate the wettability properties of SC or other lipid membranes to other solvents such as urea, ethanol, glycerol and formamide. These solvents are often used in the formulation of creams for topical application, where the wettability of sebum and SC lipids is of high interest.

- Several limitations have been recognised in the use of spherical water droplet to estimate the wettability properties of substrates, due to line tension effects[11]. Therefore, it would be beneficial to repeat the simulations on SC wettability with a different input configuration, where a cylinder of water is inserted at the interface, instead of the water droplet. This technique has been previously employed in computer simulation studies of liquid lenses at liquid-liquid interfaces[12] in order to generate a three phase line with zero curvature.

- The study of the thermal conductance of nanoparticle-fluid interfaces proposed in this thesis, was performed to understand how heat fluxes propagate from nanoparticles to the surrounding medium. Metallic nanoparticles are currently being considered for skin cancer treatment using photothermal therapy[5]. In addition to metallic nanoparticles, TiO_2 and ZnO nanoparticles[6, 7] are currently added to sunscreen cream formulations. A study of how heat propagates from a nanoparticle to surrounding SC lipids is still missing

and will be the subject of future work.

- The simulation of the interaction between monolayers and the antimicrobial peptide G4 are still at a preliminary stage. The effect of applying different surface tensions, γ_t, remains to be tested. Moreover, further simulations of the lipid monolayers with a larger number of peptides to get closer to reproducing the experimental peptide concentration will allow us to test the model used to interpret the neutron experiments, and to investigate the effect of mutual peptide-peptide interactions in promoting α-helical conformations and penetrating the lipid monolayer. A recent study by Ciumac *et al.*[13] has shown that, in addition to the effect of the lipid head group charge, the saturation of the lipid chains also has an impact on the selective interaction of lipid monolayers and G4 peptides. This remains to be tested and explored with computer simulations.

Bibliography to chapter 9

[1] H. M. Sheu, S. C. Chao, T. W. Wong, J. Y. Y. Lee and J. C. Tsai, *British Journal of Dermatology*, 1999, **140**, 385–391.

[2] R. R. Warner, M. C. Myers and D. A. Taylor, *The Journal of Investigative Dermatology*, 1988, **90**, 218–224.

[3] A. Del Regno and R. Notman, *Physical Chemistry Chemical Physics*, 2018, **20**, 2162–2174.

[4] I. Salas-García, F. Fanjul-Vélez, N. Ortega-Quijano, A. Lavín-Castanedo, P. Mingo-Ortega, M. López-Escobar and J. L. Arce-Diego, Proc. SPIE 8092, Medical Laser Applications and Laser-Tissue Interactions V, 2011, pp. 809204–809204.

[5] I. Salas-García, F. Fanjul-Vélez, N. Ortega-Quijano, J. Poliak, L. Hudcová, O. Wilfert and J. L. Arce-Diego, Proceedings of 22nd International Conference Radioelektronika 2012, 2012, pp. 1–4.

[6] I. Krasnikov, A. Popov, A. Seteikin and R. Myllylä, *Biomedical Optics Express*, 2011, **2**, 3278.

[7] I. Krasnikov, A. Seteikin and A. Popov, *Optics and Spectroscopy*, 2015, **118**, 668–673.

[8] J. Hu, C. Chen, S. Zhang, X. Zhao, H. Xu, X. Zhao and J. R. Lu, *Biomacromolecules*, 2011, **12**, 3839–3843.

[9] D. Ciumac, R. A. Campbell, H. Xu, L. A. Clifton, A. V. Hughes, J. R. Webster and J. R. Lu, *Colloids and Surfaces B: Biointerfaces*, 2017, **150**, 308–316.

[10] L. Boruvka and A. W. Neumann, *The Journal of Chemical Physics*, 1977, **66**, 5464–5476.

[11] F. Léonforte and M. Müller, *Journal of Chemical Physics*, 2011, **135**, 1–13.

[12] F. Bresme and N. Quirke, *The Journal of Chemical Physics*, 2000, **112**, 5985–5990.

[13] D. Ciumac, R. A. Campbell, L. A. Clifton, H. Xu, G. Fragneto and J. R. Lu, *ACS Omega*, 2017, **2**, 7482–7492.