Universidade de Aveiro Departamento de Química
Ano 2018/2019

HENRIQUE XAVIER SANTOS BASTOS

ESTUDO DE PARTIÇÃO DE BIOMOLÉCULAS EM SISTEMAS ANFÍFILICOS: DESENVOLVIMENTO DE UMA ESTRATÉGIA RECORRENDO A DINÂMICA MOLECULAR

TOWARDS AN UNDERSTANDING OF BIOMOLECULE PARTITION IN AMPHIPHILIC SYSTEMS: DEVELOPMENT OF A MOLECULAR DYNAMICS FRAMEWORK

2019

Universidade de Aveiro
Ano 2018/2019

Departamento de Química

HENRIQUE XAVIER SANTOS BASTOS

ESTUDO DE PARTIÇÃO DE BIOMOLÉCULAS EM SISTEMAS ANFÍFILICOS: DESENVOLVIMENTO DE UMA ESTRATÉGIA RECORRENDO A DINÂMICA MOLECULAR

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biotecnologia, ramo de Biotecnologia Molecular, realizada sob a orientação científica do Doutor Germán Pérez-Sánchez, Investigador do Departamento de Química da Universidade de Aveiro, e do Doutor Nicolas Schaeffer, Investigador do Departamento de Química da Universidade de Aveiro.

This work was developed within the scope of the project CICECO-Aveiro Institute of Materials, FCT Ref. UID/CTM/50011/2019, financed by national funds through the FCT/MCTES, and when appropriate co-financed by FEDER under the PT2020 Partnership Agreement.

o júri

presidente	Prof. Doutor João Manuel Costa Araújo Pereira Coutinho professor catedrático do Departamento de Química da Universidade de Aveiro
arguente	Prof. Doutor Eduardo Jorge Morilla Filipe professor auxiliar do Departamento de Engenharia Química e Biológica da Universidade Técnica de Lisboa
orientador	Doutor Germán Pérez-Sánchez investigador doutorado Nível 1 do Departamento de Química da Universidade de Aveiro

agradecimentos

Encerra-se um ciclo com o ano mais desafiante que tive até hoje, por vários motivos, mas também o melhor em termos de aprendizagem e crescimento pessoal. Deve-se em parte a este projeto, que exigiu de mim sair da zona de conforto, explorar algo quase totalmente desconhecido, mas extremamente aliciante. Um agradecimento aos meus orientadores, Germán e Nicolas, por terem tido a paciência e disponibilidade de acolher um novato que muitas vezes falhou, mas que foi superando os obstáculos com a sua ajuda. Não só tornaram a curva de aprendizagem mais amigável, como ainda me abriram a porta para boas oportunidades no futuro próximo.

Ainda dentro do Departamento de Química, agradecer a esta casa que também foi minha durante cinco anos, que me alargou os horizontes e para a qual espero ter contribuído de forma positiva.

A todos os colegas de curso, departamento e academia em geral que de uma forma ou outra tiveram a sua marca durante este percurso. Qualquer que tenha sido, teve a sua importância para a pessoa que sou hoje e a que serei. Em particular àqueles que me ajudaram diretamente com esta dissertação, um grande obrigado.

Aos meus amigos de sempre, seja os que já levava quando entrei para esta Universidade, que me continuaram a acompanhar sempre que possível, seja aqueles que levo agora daqui, com o mesmo valor. Sei que posso contar convosco e sabem que podem contar comigo para o que for necessário.

Às duas pessoas que mais me ajudam diariamente e que mais agradeço, Marta e a minha mãe. São as pessoas que mais me inspiram e me motivam a atingir tudo o que tenho como objetivo. À minha mãe em particular, por ser incondicionalmente a melhor amiga que alguém podia ter e por me ter ajudado a chegar até aqui. É altura de devolver o esforço.

palavras-chave

Moléculas anfifílicas, líquidos iónicos, dinâmica molecular, modelos atómicos, modelos *grão-grosso*, biomoléculas, ácido gálico, ibuprofeno, entrega controlada de fármacos

resumo

As moléculas anfifílicas são elementos de elevado potencial de estruturas auto-organizadas para vários fins biotecnológicos, devido às suas componentes hidrofóbica e hidrofílica. Parte destas são biocompatíveis, capazes de transportar biomoléculas e altamente ajustáveis e controláveis por fatores externos. Estas propriedades são particularmente relevantes em aplicações de libertação controlada de fármacos. Os líquidos iónicos são cada vez mais utilizados desde a descoberta da sua sensibilidade a estímulos, ajustabilidade e possível uso como alternativas sustentáveis a solventes convencionais. Este trabalho teve como objetivo utilizar dinâmica molecular para estudar líquidos iónicos à base de iões amónio para extração e libertação de biomoléculas, particularmente ácido gálico ou ibuprofeno. Foi utilizada uma estratégia de simulação em várias escalas com o pacote de simulação em dinâmica molecular clássica GROMACS, onde modelos com alta resolução foram usados para criar modelos *de grão-grosso* novos, mais eficientes em estudos de partição e comportamento de fases. Foi averiguada a partição de ácido gálico e ibuprofeno nas soluções de líquido iónico em questão, bem como a orientação da biomolécula na estrutura supramolecular do líquido iónico e as interações que levaram à mesma. Foi verificado um efeito à base do pH como o principal fator a afetar os sistemas estudados. Este trabalho tem o potencial de dar origem a uma plataforma transversal e transferível para analisar e testar várias combinações de biomoléculas e líquidos iónicos em soluções aquosas de forma a poupar tempo e recursos experimentais em diversas aplicações.

keywords

Amphiphilic molecules, ionic liquids, molecular dynamics, all-atom, coarse-grained, biomolecules, gallic acid, ibuprofen, drug delivery

abstract

Amphiphilic molecules are interesting building blocks of self-assembled structures for a variety of biotechnological purposes, due to their hydrophobic and hydrophilic moieties. Some of them are deemed as biocompatible, capable of carrying biomolecules, while being highly tuneable and controlled with external cues. Such properties are advantageous in drug delivery applications. Ionic liquids have gained relevance since their discovery as not only responsive and adjustable, but also as promising alternatives to conventionally used solvents. This project aims to use molecular dynamics to the study of ammonium-based ionic liquids in the extraction and delivery of biomolecules, specifically gallic acid and ibuprofen. A multiscale strategy was followed to simulate systems using the GROMACS package for classical molecular dynamics simulations. High-resolution descriptions were used to create a novel coarse-grained model to reproduce the phase behaviour and partition studies. The partition of gallic acid and ibuprofen in the studied ionic liquid solutions was assessed, as well as the particular orientation of the biomolecule in the supramolecular structure of the ionic liquids, as well as the interactions generating each outcome. A pH-driven effect was verified as the main parameter affecting the studied systems. This work has the potential to pave the way for a transferable, transversal platform to analyse and test different biomolecule-IL combinations in aqueous solutions in order to save time and experimental resources in diverse applications.

Abbreviations

$[C_{10}mim]^+$	1-*n*-decyl-3-methylimidazolium
AA	All-Atom
ABS	Aqueous Biphasic System
AN	Aggregation Number
API	Active Pharmaceutical Ingredient
B3LYP	Becke, 3-parameter, Lee-Yang-Parr exchange-correlation functional
CG	Coarse-Grained
CHARMM	Chemistry at HARvard Macromolecular Mechanics
CHELPG	Charges from Electrostatic Potentials using a Grid-based method
CMC	Critical Micellar Concentration
CN	Coordination Number
COM	Center Of Mass
CP	Cloud Point
CPP	Critical Packing Parameter
CTAB	Cetyltrimethylammonium bromide
DFT	Density Functional Theory
DPD	Dissipative Particle Dynamics
EM	Energy Minimization
FF	Force Field
GA	Gallic Acid
GROMACS	GROningen MAchine for Chemical Simulations
[Hbet][Tf2N]	Bis(trifluoromethylsulfonyl)imide
IBP	Ibuprofen
IL	Ionic Liquid
LCST	Lower Critical Solution Temperature
LINCS	LINear Constraint Solver
LJ	Lennard Jones
MD	Molecular Dynamics
MHC	Minimum Hydrotrope Concentration

$[N_{11114}]Cl$	Trimethyltetradecylammonium chloride
$[N_{44414}]Cl$	Tributyltetradecylammonium chloride
$[N_{4444}]Cl$	Tetrabutylammonium chloride
OLW	Oil-Water
OPLS	Optimized Potential for Liquid Simulations
OW	Octanol-Water (used in the partition coefficient)
$[P_{44414}]Cl$	Tributyltetradecylphosphonium chloride
$[P_{4444}]CF_3COO$	Tetrabutylphosphonium trifluoroacetate
PBC	Periodic Boundary Conditions
PDB	Protein Data Bank
PEG	PolyEthylene Glycol
PME	Particle Mesh-Ewald
PW	Polarizable Water
RDF	Radial Distribution Function
RM	Reverse Micelles
RT	Room Temperature
SAIL	Surface-Active Ionic Liquid
SASA	Solvent-Accessible Surface Area
SDF	Spatial Distribution Function
TRAVIS	TRajectory Analyzer and VISualizer
UCST	Upper Critical Solution Temperature
vdW	van der Waals
VMD	Visual Molecular Dynamics

Index

Chapter I: Introduction 3

- **1. Introduction 3**
  - 1.1 Solubilisation of poorly soluble compounds 3
    - a) Possible solubilisation processes 3
    - b) Amphiphilic molecules 4
    - c) Hydrotropes 10
  - 1.2 Aggregation in amphiphilic systems 13
    - a) Forces that drive the aggregation process 13
    - b) Phase behaviour and its relation to the surfactant structure 15
    - c) Phase separation in amphiphilic systems 16
  - 1.3 Ionic liquids 18
    - a) Tuneable solvent 18
    - b) Applications 19
- **2. Computational Chemistry 20**
  - 2.1 Multiscale approach 22
  - 2.2 Molecular Dynamics 24
    - 2.2.1 Simulations in the GROMACS package 25
    - 2.2.2 Coarse-Grained models 27
- **3. Objectives 28**

Chapter II: Methodology 31

- **1. Simulations details 33**
- **2. All-atom molecular dynamics (AA-MD) systems 34**
  - 2.1 Ammonium-based ILs 34
  - 2.2 GA and IBP 35
- **3. AA mapping to CG and validation 37**
- **4. Coarse-grained molecular dynamics (CG-MD) systems 41**
- **5. Analysis 41**

Chapter III: Results and Discussion 47

- **1. Biomolecule model validation 49**
- **2. ILs phase behavior 56**

3. GA in aqueous IL solutions ... 63
   - a) Gallic acid partition ... 64
   - b) Density profile analysis ... 71
   - c) Gallic acid hydration ... 78
   - d) Influence of GA on phase transition – micelle swelling ... 80
   - e) Influence of Gallic acid on phase transition –ion-exchange ... 85
   - f) Summary of findings ... 88

Chapter IV: Conclusions and future work ... 93

Chapter V: Bibliography ... 99

Supporting Information ... 113

Figure Index

Figure 1 - General structure of an amphiphilic molecule. In green, the hydrophobic alkyl chain, with *n* indicating the number of carbon atoms. In purple, the polar hydrophilic head. 4

Figure 2 - Examples of amphiphilic molecules shapes, from simpler, rod-like micelles to more complex spheres and bilayered structures. 5

Figure 3 – Schematic representation of some of the main proposed hydrotrope mechanisms of solubilisation. Hydrotrope molecules are represented in green, water in blue and the target molecule for solubilisation in red. a) Complexation of the hydrotrope with the target molecule, forming aggregates that increase the solubility of the latter; b) Solvent structure disruption by the hydrotrope. Holes in the solvent structure are filled by the target molecule, while the hydrotrope then continues the same process. 11

Figure 4 - Examples of non-bonded interactions involved in self-assembly of amphiphilic molecules. 13

Figure 5 - Multiple non-bonded interactions within one amphiphilic molecule. In this example an imidazolium-based IL cation, 1-*n*-decyl-3-methylimidazolium ($[C_{10}mim]^+$) is used. 14

Figure 6 – Example of a phase diagram relating the concentration of a polymer aqueous solution and the temperature. The phase behaviour is depicted by either monophasic (grey) or biphasic (blue). The critical boundaries of the solubility curves, UCST and LCST points, are represented by red dots. 17

Figure 7 – Archetypical cations and anions commonly used as constituents of IL systems. 18

Figure 8 - Models for bulk structures of ionic liquids, from single ionic interactions (ion pairs), to small IL networks (H-bond networks) and clusters of self-assembled structures.[78] 19

Figure 9 – Multiple scales involved in this work, using a GA molecule as an example. The multiscale strategy description applies to the development of each subsequent level in this scale. 23

Figure 10 - Lennard Jones potential, relating the energy of interaction (V) and the distance between particles (r). The minimum value indicates the optimal balance between the repulsion and the attraction terms.[110] 26

Figure 11 - MARTINI energy matrix of interaction.[111] Each level of interaction is used as the energy of interaction (ε) in the LJ calculations and the bead size (σ) is 0.47 nm by default. Beads are categorized as charged (Q), polar (P), non-polar (N) and apolar (C). In charged Q beads, d, a, da or 0 indicate the hydrogen bond capabilities; donor, acceptor, donor and acceptor and none, respectively. 1 to 5 indexes indicate increasing polar or apolar affinity being 1 the lowest polar/apolar character. From O to IX, the interaction strength is: 5.6 kJ/mol, 5.0 kJ/mol, 4.5 kJ/mol, 4.0 kJ/mol, 3.5 kJ/mol, 3.1 kJ/mol, 2.7 kJ/mol, 2.3 kJ/mol, 2.0 kJ/mol, 2.0 kJ/mol ($\sigma = 0.62$ nm at the IX selection). 28

Figure 12 - Flowchart of the general line of thought used in this project and the main objectives. 29

Figure 13 - Charge distribution comparison between the deprotonated (left) (system 2) and protonated (right) (system 1) version of GA, related to the carboxylic acid group. These were calculated with the B3LYP/6-311+G(d,p) level of theory. Red colour indicates negatively charged sections while blue colour illustrates positively charged regions. 37

Figure 14 – Charge distribution comparison between the deprotonated (left) (system 6) and protonated (right) (system 5) version of IBP. These were calculated with the B3LYP/6-311+G(d,p) level of theory. Red colour indicates negatively charged sections while blue colour illustrates positively charged regions. 37

Figure 15 – AA to CG mapping for $[N_{11114}]^+$, with a charged polarity being selected for the surfactant head (Q0), and an apolar one for the alkyl tail (C1). 38

Figure 16 – Mapping from AA to CG of $[N_{44414}]^+$. The center bead of the polar head is set as charged (Q0), while the butyl chains of head are mildly apolar (C3). The alkyl tail is described as apolar (C1), exactly as in $[N_{11114}]^+$. 38

Figure 17 – AA mapping to CG of $[N_{4444}]^+$. Similarly to $[N_{44414}]^+$, the central ammonium atom is set as charged (Q0), along with the initial carbon atoms surrounding it. The butyl chains are set as mildly apolar (C3). 38

Figure 18 – CG mapping from AA for GA. A special mapping was employed, labelled by an “S” preceding the actual bead polarity. The aromatic ring was set as mildly apolar (SC3), while the carboxylic acid and hydroxylic groups had different polarities depending on the simulated pH, and consequently their protonated states. The former ranged from nonpolar (SNda) to charged (SQa) while the latter ranged from weakly apolar (SC5) to charged (SQa). 39

Figure 19 – IBP mapping from AA to CG. While the most apolar region is defined by a default mapping, apolar (C1), the rest of the procedure uses the special mapping, also applied to GA. The aromatic ring is again mildly apolar (SC3), while the intermediate alkyl region between it and the carboxylic acid group is more polar (SC4). The latter group is either set as nonpolar (SNda) if protonated or charged (SQa) if deprotonated. 39

Figure 20 – Counter ion mapping procedure from AA to CG. Chloride was used as anion, for the ionic liquids used, being described as charged and hydrogen bonding acceptor (Qa), while sodium as used as cation, to charge balance deprotonated GA and IBP, being set as charged and donor in hydrogen bonds (Qd). Both CG mappings have six implicit water molecules. 39

Figure 21 – Polarizable CG water PW, containing three particles: a central center of mass (W) and two charged dummy interaction centers, a positive (WP) and a negative (WM) to mimic the orientational polarizability of water. Both are connected to W at a distance of *l* and can harmonically vibrate with a spring constant of K_θ (harmonic function).[143] 40

Figure 22 – Schematic representation of the RDF calculation using GA (in dark blue and cyan) and water (red and white) as example. In this case, the carboxylic acid group of the GA molecule is used (dark blue), from which center of mass (COM) of the selected water molecules (red and white) is performed at a given distance *r*. At each value of *r*, the probability of finding water around the carboxylic acid group of GA is calculated and described by $g(r)$. 42

Figure 23 – Two-dimensional view of the double cubic lattice method to calculate SASA. A three-dimensional grid is displayed in the system, with *r* indicating the maximum atomic radius of the reference atoms (in blue). This radius is given by the sum of the vdW radius and the accessible surface to a probe molecule. The SASA is calculated using the disposition of the selection atoms (in red) towards the reference ones, considering the atomic radius. 43

Figure 24 – RDF profile of AA (left) and CG (right) GAwith the carboxylic acid group selected as a reference, both in its protonated (system 1,3) and deprotonated (system 2,4) states and water as the selection. 51

Figure 25 - SDF of water (blue) and other GA molecules (green) surrounding a Prot. GA molecule (top) and a Deprot. GA (-1) molecule (bottom), as references. A side view (left) and a top view (right) of the SDF are provided, the GA orientation is set with the carboxylic acid group on the left and the hydroxyls on the right. 53

Figure 26 – RDF profile of AA (left) and CG (right) IBP with the carboxylic acid group selected as a reference, both in its protonated (system 5,7) and deprotonated (system 6,8) state and water as the selection. 54

Figure 27 – SDF images of water (blue) and other IBP (green) structures surrounding a reference IBP molecule, both in the protonated (top) and deprotonated states (bottom). The reference structure is placed with the carboxylic acid group on the left part of the image. 55

Figure 28 – Visualization of the resulting mesophases of [N_{11114}]Cl and [N_{44414}] and after the addition of NaCl at lower and higher concentrations (systems 9-14). Cation heads are depicted in purple, with the alkyl tail in green and chloride ions (both the IL anion and salt component) in black. Water and sodium molecules were removed for clarity...... 57

Figure 29 – RDF profile of CG [N_{11114}]Cl (systems 9-11) and [N_{44414}]Cl (systems 12-14), using the polar head of the cation as a reference and the PW as the selection. 59

Figure 30 – SASA profiles for the polar head group of the CG [N_{11114}]Cl (left) and [N_{44414}]Cl (right) cation (systems 9-14). A moving average was applied to smooth the function. 59

Figure 31 – Density profiles of [N_{11114}]Cl before the addition of salt (top), with NaCl at 0.84 wt.% (center) and at 4.75 wt.% (bottom) (systems 9-11). The SAIL has its cation heads represented in purple and tails in green. Chloride, both from the SAIL anion and NaCl is depicted in black. Sodium is presented in dark yellow and water in cyan. In each block, water density is represented in the top segment while the rest of the components is depicted in the bottom one, for increased clarity. The profile was generated along the Y axis of the simulation box. Snapshots of a cross-section of the simulation box are presented on the left side of the figure, following the same order as the density profiles. 62

Figure 32 - Final snapshots of the production runs comprising [N_{11114}]Cl, [N_{44414}]Cl and [N_{4444}]Cl with GA at low and high concentrations and various protonation states (systems 16-21, 23-28 and 30-35). IL cation head groups are depicted in purple and alkyl chains in green, while the carboxylic acid group of GA is presented in blue, its aromatic ring in yellow and the hydroxyls in red. Both water, chloride and sodium counter ions were removed for increased clarity. 69

Figure 33 - Density profiles for systems 19-21, containing GA and [N_{11114}]Cl, using the COM of the IL micelle as the reference and presenting the density of each component in relation to the distance to it. A snapshot on the left side corresponding to a representative micelle in each state (Prot. GA in the top, Deprot. GA (-1) in the middle and Deprot. GA (-3) in the bottom) is presented on the left side of the figure. The micelle surface is highlighted by a red circle. 73

Figure 34 - Density profiles for systems 26-27, containing GA and [N_{44414}]Cl, using either the COM of the IL micelle as the reference and presenting the density of each component in relation to the distance to it. For Deprot. GA (-3) the low concentration system (25) was used as it is still micellar and allows the use of the cluster counting algorithm A snapshot on the left side corresponding to a representative micelle in each state or the full simulation box (Prot. GA in the top, Deprot. GA (-1) in the middle and Deprot. GA (-3) in the bottom) is presented on the left side of the figure. The micelle surface is highlighted by a red circle........ 74

Figure 35 - Density profiles for systems 33-35 containing GA and [N_{4444}]Cl, using the full simulation box in the *z* axis direction and measuring the density of each component throughout it. A snapshot on the left side corresponding to the simulation box (Prot. GA in the top, Deprot. GA (-1) in the middle and Deprot. GA (-3) in the bottom) is presented on the left side of the figure. Water was removed in these for increased clarity. 76

Figure 36 – RDF profile using the Ct3 bead of [N_{44414}]Cl as the reference and the hydroxyl beads of GA as the selection (systems 23-25). 77

Figure 37 – RDF profiles using the carboxylic (left) and hydroxylic (right) moieties of GA as the reference and PW as the selection (systems 19-21, 26-28, 33-35). 79

Figure 38 – RDF profiles using either the carboxylic or hydroxylic moieties of GA as the reference and the polar heads of the IL cation as a selection (systems 19-21, 26-28, 33-35). 82

Figure 39 – RDF profiles using either the polar heads or the alkyl chains group of each IL cation as the reference and PW as the selection (systems 15, 19-21, 22, 26-28, 29, 33-35). 85

Figure 40 – RDF profiles using the polar heads of each IL cation as the reference and the respective IL anion (chloride) as the selection (systems 15-35). 87

Table Index

Table 1 - Recent studies on applications combining surfactants and molecules of interest. 6

Table 2 - Recent studies on applications of hydrotropes in drug delivery. 12

Table 3 – An example of recent MD simulations involving biomolecules. 21

Table 4 - Advantages and disadvantages of using AA and CG models. 24

Table 5 - Input and output files used in the GROMACS MD simulations. 34

Table 6 – Reference values related to pH, solubility and partition used for GA and IBP. . 36

Table 7 – Simulation parameters of each system, regarding the production run, with the respective notation. Box sizes are set as “3” for cubic boxes, “2” for rectangular ones and “dode” for dodecahedric boxes. 44

Table 8 – CN of the RDF profiles used for validation of CG models of GA and IBP. Number in brackets represent the coordinates of the peak (*r* in the RDF profile) where the CN is estimated, relatively to the reference group, in nanometres. 50

Table 9 – Summary of the phase behaviour, CNs and SASA (area) values obtained for systems 9-14. The coordinates at which the CNs were obtained relatively to the reference group are presented in brackets. The SASA is the result of a running average, with the SASA per bead of the group being presented in brackets when applicable. 57

Table 10 – Summary of the main observations obtained with the ILs + GA systems (systems 15-35). The coordinates at which the CNs were obtained relatively to the reference group are presented in brackets. The SASA presented is the results of a running average, with the SASA per bead of the group being presented in brackets when applicable. 66

Table 11 – Aggregation number (AN) of the SAIL micelles and CN of each group of GA when using the SAIL micelle core as a reference (Ct3 bead). 80

Chapter I: Introduction

1. Introduction

Supporting the current green chemistry trend, the United Nations selected 17 sustainable development goals in which the innovation of industry is included.[1]Amphiphilic molecules, which contain both a polar and an apolar region, are proposed as more environmentally sustainable alternatives to organic solvents in many biochemical applications.[2] The possibility of tailoring the molecule properties for specific applications extends their potential. This effort could be reproducible in the biotechnological field, such as in the extraction of biomolecules and in the delivery of certain pharmaceutical compounds. Computational methods, which include molecular dynamic simulations, have emerged as powerful tools to shed light on the molecular interactions and complementing experimental results.[3] This work describes the development and application of molecular dynamics (MD) all-atom (AA) and coarse-grained (CG) models of amphiphilic systems, namely the effect of the molecule structure on the extraction of biomolecules, the interactions of the latter at the interface of the aggregates and finally the conceptual design of a responsive drug-release system. Gallic acid (GA) is the biomolecule assessed in this project as it is well established in the pharmaceutical field, as a probe, and validation data, critical to computational approaches, is plentiful. A new model for ibuprofen (IBP), a common drug with broad potential, is also developed with the aim of extending the obtained results to future works. Systems comprising both ionic liquids (ILs) and these active pharmaceutical ingredients (APIs) are addressed for the drug solubilisation processes involved, *i.e.* use of cosolvents in general, micellar solubilization through surface active molecules and hydrotropy.[4]

1.1 Solubilisation of poorly soluble compounds

a) Possible solubilisation processes

The low solubility of a wide range of biomolecules with biotechnological interest is a major drawback to their application. For APIs and related biomolecules, this affects their extraction and purification, as well as drug delivery efficiency. To increase the solubility of these important molecules, co-solvents, which are compounds able to increase the solubility of another solute, are often added to the target solution.[4] Some of the main alternatives to conventional co-solvents used comprise amphiphilic molecules, with surface activity, or hydrotropes. Although amphiphilic molecules, hydrotropes are known to induce a degree of water structuring[5] but resort to other mechanisms of action than the ones of surface active molecules. In this work these two different categories were assessed for their impact on the solubility of GA, particularly by using ILs which have either of these co-solvent effects and determining the weight of each action. The differences between amphiphilic molecules as surfactants and hydrotropes will be addressed in the next sections.

b) Amphiphilic molecules

Amphiphilic molecules possess two distinct parts in their structures: a polar hydrophilic and an apolar hydrophobic regions as illustrated in **Figure 1**.[6] This functional duality leads to the assembly of amphiphiles in different structures when in contact with a solvent of sufficient polarity, such as aqueous solutions.[7] The self-assembly results in a variety of aggregates, with different shapes and sizes depending on their physical interactions.[8]

Molecular aggregates are formed by a spontaneous self-assembly that coalesce individual particles into an organized structure with a wide range of different properties.[9,10] These can self-assemble in a variety of morphologies depending on the conditions, such as rods, cylinders, discs, sheets, spheres, vesicles, interconnected rods, bilayers, and other 3-D periodic structures.[9] Visual representations of these possibilities are depicted in **Figure 2**.

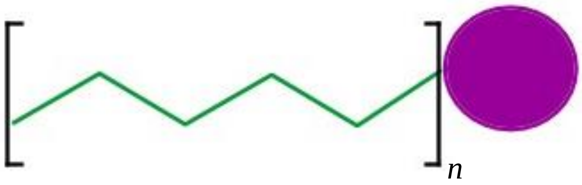


Figure 1 - General structure of an amphiphilic molecule. In green, the hydrophobic alkyl chain, with *n* indicating the number of carbon atoms. In purple, the polar hydrophilic head.

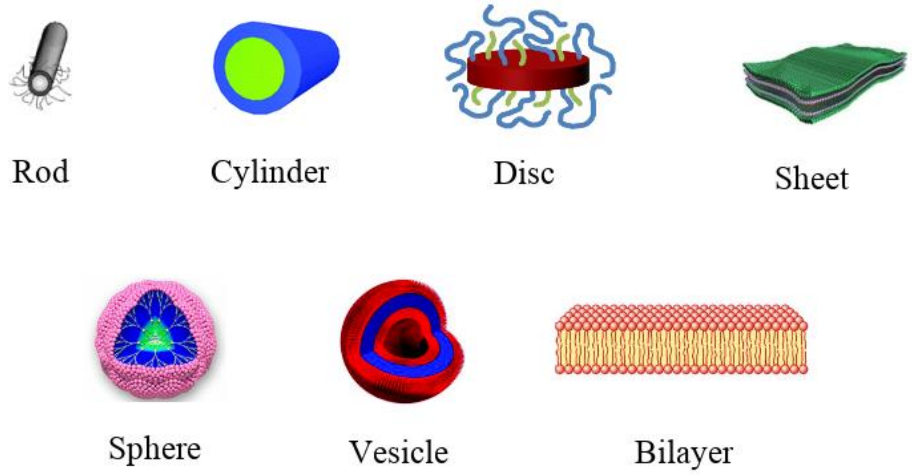


Figure 2 - Examples of amphiphilic molecules shapes, from simpler, rod-like micelles to more complex spheres and bilayered structures.

Examples of amphiphilic molecules currently used in biotechnological applications are ionic, non-ionic, lipids and others. Among these, a segregation is possible between non-ionic- and ionic-based amphiphilic molecules. Both are categorized depending on the hydrophilic moiety, with the former indicating polar uncharged groups, while the latter represents a cationic or anionic hydrophilic part (also called polyelectroytes).[11] This definition is important due to several properties that emerge with ionic amphiphilic molecules, such as certain ILs, which will be used in this work. One of the main properties of amphiphiles is the ability to act as surface active agents (surfactants), that is, to reduce the surface tension between two interfaces,[11] useful for a wide array of applications, such as drug design and delivery[12], emulsion stabilizers[13] and even biofuel production[14]. Important applications involve important molecules, particularly biological amphiphilic molecules, such as pharmaceuticals, as illustrated in **Table 1**.

Table 1 - Recent studies on applications combining surfactants and molecules of interest.

Surfactant(s) (type)	Target molecule(s)	Application	Relevant findings	Ref.
Tween 65 and 80 (non-ionic)	GA and other antioxidants	Stabilization of fatty acid emulsions suffering from oxidation	Surfactant/antioxidant combination and hydrophobicity to create an optimal stabilization	15
N,N-dimethyl-1-adamantanamine (cationic)	IBP	Controlled release of IBP in a cellulose and cyclodextrin-based hydrogel crosslinked by the surfactant	Loading of an insoluble drug into a hydrophilic structure; Triggered release	16
19 variants (non-ionic, anionic, cationic)	Danazol; fenofibrate	Solubilization study of hydrophobic drugs with various surfactants	Surfactant molecular structure affects drug solubilization; Solubilization is directly proportional to chain length; Ion-dipole interactions induce high solubilization in ionic surfactants	17
Cetyltrimethylammonium bromide (CTAB) (cationic)	Sodium valproate	Development of a drug-based IL surfactant after studying interactions between individual components	Strong one-on-one interactions between drug and surfactant; Increased solubility of the API in the IL; Ability to tune delivery and release	18
CTAB (cationic)	Diclofenac; IBP; ketoprofen	Removal of drugs from hospital wastewaters using an electrocoagulation-flotation process	Cationic surfactants reduce gas bubbles size improving flotation processes and neutralizing anionic drugs; Efficient application to real hospital wastewaters	19
Irinotecan hydrochloride (amphiphilic prodrug)	7-ethyl-10-hydroxy camptothecin; paclitaxel; camptothecin	Nano formulations of amphiphilic drugs as excipients for hydrophobic drugs	Readily clinical application of insoluble drugs; Use of amphiphilic drugs as surfactants; Improved therapeutic potency; Possible combinational therapy	20
Cetylpyridinium chloride (cationic)	Diclofenac; IBP; sodium salicylate	Use of drugs to modify the rheology of surfactants	Drugs modified rheology of surfactant solutions as regular salts; Morphological transitions up to a threshold concentration of drugs,	21

			highly dependent on molecular structure of the latter	
Lecithin; monoglycerides (zwitterionic, non-ionic)	Hydroxytyrosol; GA	Formation of reverse micelles (RM) and their use as carriers of antioxidant biomolecules	Size of RM containing GA increased, indicating its encapsulation, contrary to hydroxytyrosol; RM formation is spontaneous	22
Sodium octyl sulfosuccinate (anionic)	IBP (sodium salt)	Effect of electrolytes in the interactions between IBP and surfactant	Salt reduces repulsions between ionic head groups due to a screening effect; Increasing molar fraction of the drug increase hydrophobic attractive interactions	23
Pluronic F127; Cremophor RH40 (polymeric; non-ionic)	Ketoconazole	Nano emulsions using surfactants and essentials oils to improve the release of ketoconazole in *Candida albicans* infections	Nano emulsions are stabilized by the surfactants; Ketoconazole influences droplet size distribution; Non-ionic surfactants are more compatible, stable and less toxic; Nano emulsions allowed faster release	24
Octylguanidium chloride	Parabens; benzophenones	Extraction of personal care products in cosmetics using an IL with low cytotoxicity	IL has a lower critical micellar concentration (CMC) than other commonly used surfactants; IL cations with shorter alkyl chains are less toxic; Lower volumes of IL and extraction times for an efficient process	25
Tributyltetradecyl phosphonium chloride	Cobalt, iron, platinum and nickel ions	Extraction of metal ions from acidic solutions using an acidic aqueous biphasic system (ABS) based on an IL	Temperature-driven phase separation, reducing acid use; Multiple ions may be obtained from a one “pot” procedure, directly from waste devices leachates.	26
Imidazolium-, ammonium-based acetate ILs	Acyclovir; methotrexate	Use of ILs in oil microemulsions for drug delivery of acyclovir and methotrexate	Surfactants reduce viscosity of the system; Drug capacity varies with IL due to different dimensions; ILs stabilize the system	27
1-dodecyl-2,3-dimethylimidazoliu	Amitriptyline hydrochloride	Surface tension and conductance studies to assess the mixed micellization behaviour of	Mixed micelles have a lower CMC than the pure amphiphilic drug; This trisubstituted SAIL has a more effective micellization and	28

m chloride		a surface active IL (SAIL) with amitriptyline hydrochloride	drug binding than disubstituted ones, due to higher H-bonding and amphiphilicity	
1-hexyl-3-methylimidazolium dodecylsulfate	Tetracaine	Interaction and micellization studies of tetracaine in a SAIL	Physiological pH implies the cationic form of the drug, forming a highly bonded complex with IL due to electrostatic interactions; Increasing drug concentration delays the self-assembly of the IL, with the drug penetrating into the micellar core	29
1-octyl-3-methylimidazolium chloride	Doxorubicin hydrochloride	Assessment of the aggregation behaviour of doxorubicin hydrochloride in this SAIL	Drug/SAIL aggregates are formed spontaneously as large spherical vesicles; H-bonds are also determinant in this process; Salt addition induces phase transition to rod-like fibrils and aggregates with the individual components	30

Most studies presented in **Table 1** focus on the development of systems that allow not only solubilization of hydrophobic drugs in biological environments, but also to be able to control the release of the active pharmacological agent. The behaviour of amphiphilic molecules can be controlled by modifying the physicochemical characteristics of the solution such as the temperature or the pH among others, which is particularly important in drug delivery applications.[31] For instance, IBP can be released from an amphiphilic-crosslinked hydrogel host by a pH change in the environment.[16] In this work, Fan and collaborators[16] used cyclodextrin combined with modified cellulose to deliver the IBP. This technique demonstrated the versatility of amphiphilic agents, being used as a structural stabilizer of the actual drug carrier. Amphiphilic drugs can also act as carriers for other pharmaceuticals, reducing the number of steps in clinical trials, as described by Hu and colleagues.[20] In this study, the amphiphilic character of irinotecan hydrochloride self-assembles with insoluble drugs to form disperse nanoparticles in water. This fact bypasses the use and clinical approval of exogenous materials as only drugs are used for these formulations and the loaded drugs become sufficiently more soluble. Another interesting application presented by Pasquino and team[21] is the reverse strategy of combining surfactants and drugs; modification of the surfactant behaviour by using these structures as salts in the way that they can induce a phase transition. The design of tailor-made systems for specific applications can be further achieved by a proper selection of the amphiphilic character of the surfactant to fine tune the driving forces of the self-assembly besides the solvent characteristics. The addition of inorganic molecules in these mixtures and how these affect the phase behaviour is a matter of great interest, especially in pharmaceutical applications. These forces and their effects are explored in the following section.

In this project, GA is the target molecule not only due to its antioxidant activity and wide presence in biological sources and processes, but mainly owing its use as a proxy molecule to determine the probable behaviour of a wide range of pharmaceutical compounds in various systems. For example, this molecule was used by López-Martinez and Rocha-Uribe[15] by combining it with surfactants in order to avoid oxidation of fatty acids in food emulsions.[15] That was achieved since GA is moderately hydrophobic, providing a good combination with a surfactant, being held in the micellar core, and improving its effect as an antioxidant. Regarding the application of GA as a probe molecule, a recent work was developed to assess the effect of some chloride-based ILs ABS of polyethylene glycol (PEG) and salt, using the GA and other molecules as probes for partition studies.[32] IL use as adjuvants in PEG + salt ABS was beneficial for biomolecule partition, being the GA partition coefficient highly increased, with tetrabutylammonium chloride ([N_{4444}]Cl) as one of the most competent ILs for this purpose.[32]

c) Hydrotropes

Hydrotropy is frequently defined as a mechanism in which the solubility of a particular compound in a solution, usually aqueous, is drastically increased by adding a sufficient amount of a second solute, named hydrotrope,[33] being considered for the first time as a specific class of compounds by Neuberg.[34] Their functionality results from its general structure; an amphiphilic molecule different from surfactants in the way that it does not intrinsically self-assemble into organized supramolecular structures since the hydrophobic chain is not long enough.[35] Although the mechanism of action of hydrotropes to increase solubility of a solute is not fully understood, these molecules may aggregate incrementally up to a point where interactions between the hydrotrope and the solute may aid its solubilization.[36] The main mechanisms proposed to describe hydrotrope action involve the complexation between the hydrotrope and the target solute or disruption or enhancement of the solvent structure.[37] A schematic representation of these hypothesis are presented in **Figure 3**. Briefly, the former is based on the arrangement of the hydrotrope molecules around the hydrophobic solute with a relatively short hydrophobic chain and a hydrophilic component, often charged, which interacts with water. Although these molecules are not capable of self-assembly as surfactants, the use of a minimum hydrotrope concentration (MHC) is being currently applied to define a threshold concentration at which the first mechanism is stable.[38] The influence of hydrotropes on solvent structure might occur due to the formation of open spaces in the solvent, which can be occupied by the solute to enhance the solvation.

Hydrotrope properties are useful in drug solubilisation and delivery since most APIs have low solubility in aqueous media, improving the efficiency of drug delivery and further release. For example, Sintra *et al.*[38] described the use of catanionic hydrotropes, one of them [N_{4444}]Cl, used in this work, in the solubilisation of IBP.[38] Similar studies are paving the way for the use of ILs as hydrotropic agents, particularly catanionic, being promising in drug solubilisation and delivery of APIs. Some of those are presented in **Table 2**.

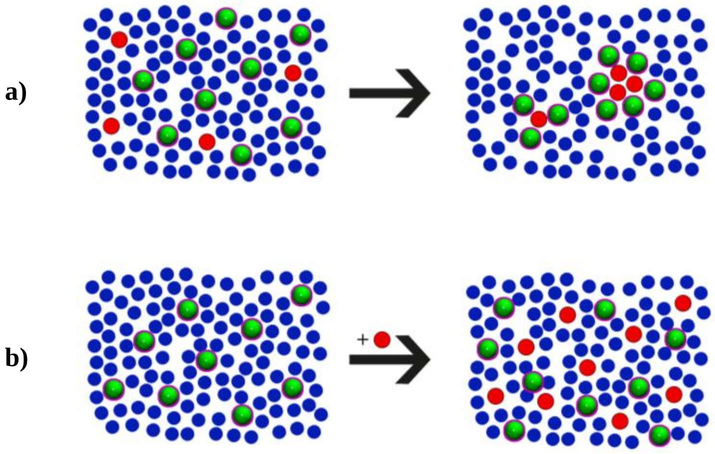


Figure 3 – Schematic representation of some of the main proposed hydrotrope mechanisms of solubilisation. Hydrotrope molecules are represented in green, water in blue and the target molecule for solubilisation in red. a) Complexation of the hydrotrope with the target molecule, forming aggregates that increase the solubility of the latter; b) Solvent structure disruption by the hydrotrope. Holes in the solvent structure are filled by the target molecule, while the hydrotrope then continues the same process.

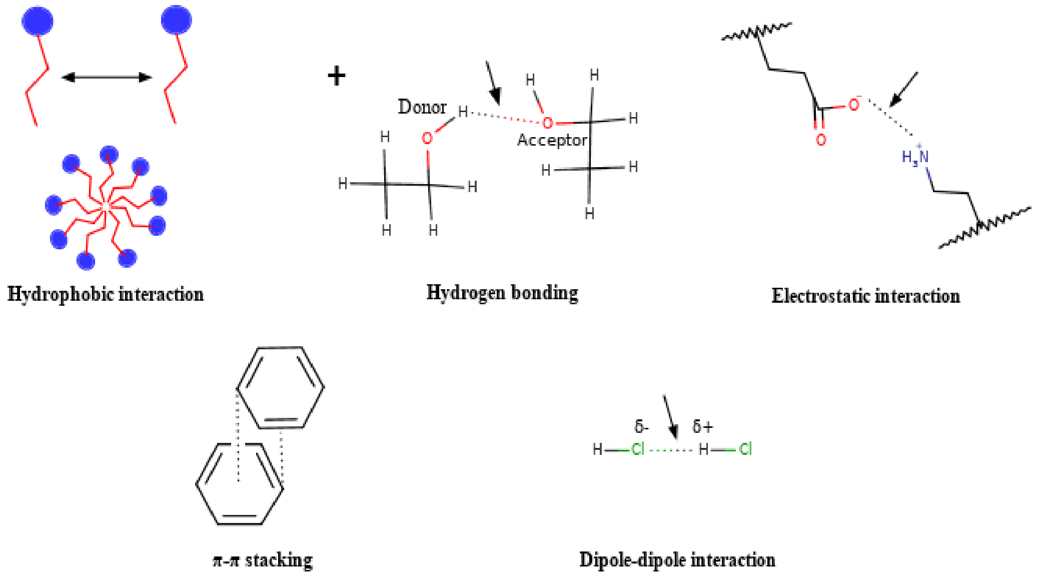
Table 2 - Recent studies on applications of hydrotropes in drug delivery.

Hydrotrope	Drug	Application	Relevant findings	Ref.
Tetramethylene-1,4-bis(*N*,*N*-dimethyl-*N*-butylammonium)bromide	IBP	Enhancement of IBP-surfactant interactions	Hydrotrope promotes aggregation of individual surfactant and IBP and between them	23
Sodium deoxycholate	Curcumin	Use of an hydrotrope as a stabilizer of nanoparticles used in drug loading	Stabilization of drug delivery vessel by controlling the size of the nanoparticle and the interactions involved in aggregation	39
Sodium salicylate; sodium benzoate	5-fluorouracil	Hydrotropic solutions to promote transdermal drug delivery	Main mechanism of hydrotrope action might be related with solvent breakage; Hydrotrope action promoted transdermal delivery	40
Sodium cumene sulfonate	Griseofulvin	MD simulations to understand hydrotrope mechanism of action on drug solubility	Hydrotrope aggregates into clusters, in which drug partitions; van der Waals (vdW) interactions are the main ones involved	41
Diclofenac sodium	Diclofenac sodium	Concentration effect of diclofenac sodium in the phase behaviour of a mixture with a SAIL	This drug has an hydrotrope action, affecting the phase behaviour of the complex; The mixture moves from prolate ellipsoids to micelles and to the former again, with increasing drug concentration	42
Various	Vanillin; GA	Screening of various hydrotrope ILs action on the solubility of vanillin and GA	First depiction of IL use as hydrotropes; Solubility increase in water due to aggregation between ILs and the biomolecules; Efficiency increases with IL cation alkyl chain length, anions with hydrotrope capacity, temperature increase	5
Various	IBP	Use of ILs as hydrotropes to improve IBP solubility and the underlying mechanism	All ILs demonstrated a highly efficient hydrotrope action and solubility increase, including the ammonium and phosphonium based ones; This action was driven by the formation of IBP-IL aggregates by non-polar interactions	38

1.2 Aggregation in amphiphilic systems

a) Forces that drive the aggregation process

Herein, a set of physical interactions that determine not only the shape but also the behaviour and self-assembly organization of amphiphilic molecules is described in detail. Aggregation into specific structures of surfactants is achieved by action of various non-bonded forces, such as hydrophobic, hydrogen bonding, electrostatic, π-π stacking interactions as illustrated in **Figure 4**.[31,43]



Hydrophobic interaction

Hydrogen bonding

Electrostatic interaction

π-π stacking

Dipole-dipole interaction

Figure 4 - Examples of non-bonded interactions involved in self-assembly of amphiphilic molecules.

The balance between each force can lead to different final morphologies. The main driving force of the assembly of individual molecules/smaller aggregates into larger ones is the hydrophobicity of the molecule alkyl-chain tails and the repulsive/attractive forces of hydrophilic head groups, being either closely related to hydrogen bonding in non-ionic surfactants, or also with electrostatic interactions in ionic surfactants.[8,44]

These phenomena are not associated to a single type of interaction. Cooperation of multiple bonds in the same type of molecules occurs by balancing the number of interactions and their strength, determining the structural outcome of the system.[45] Main interactions and their locus within the amphiphilic molecule 1-*n*-decyl-3-methylimidazolium ($[C_{10}mim]^+$) are summarized in **Figure 5** as an example. The $[C_{10}mim]^+$ cation is composed of an alkyl chain (C_{10}), and a methylimidazolium head (mim), which comprise a possible cation of a SAIL, requiring an anion to counter balance the electrostatic charge. Heuristically, these interactions lead the system to a more stable structure by decreasing the Gibbs free energy, either enthalpically of entropically, even though intermediate states may form.[44]

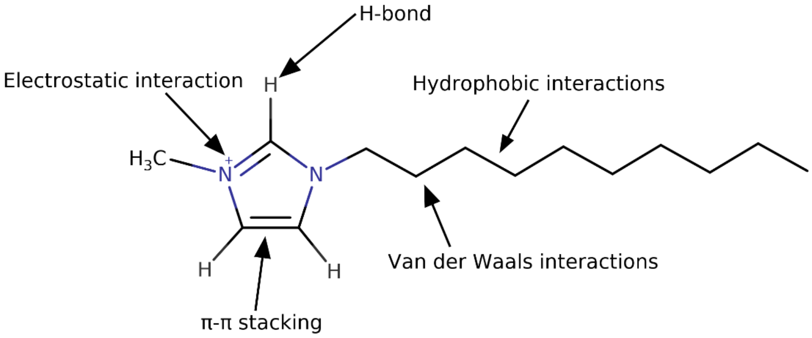


Figure 5 - Multiple non-bonded interactions within one amphiphilic molecule. In this example an imidazolium-based IL cation, 1-*n*-decyl-3-methylimidazolium ($[C_{10}mim]^+$) is used.

Hydrogen bonds are important non-bonded interactions since these are present in many amphiphilic systems and the strength is typically comparable to weak covalent bonds.[46] Hydrogen bonding is the attraction between a proton donor, usually an electronegative atom to which the hydrogen is connected, and another atom with electronic density, the hydrogen bond acceptor.[47] Electrostatic interactions are specific to charged head groups in ionic surfactants. Structures in equilibrium are optimal when a Gibbs energy minimum is achieved and when the repulsion between the same ions and attractions between counterions are counter balanced. In this situation, head groups repel each other until they are properly stabilized and solvated by their counterions.[48] Hydrophobic interactions are non-directional and lead the system into a more entropically favourable state in aqueous solutions.[49] As the contact with the solvent is reduced, the entropy increases due to hydrophobic chain and free polar solvent molecule movement. This directly entails the Gibbs free energy equation (G), which is shown in **Equation 1**, which allows the monitoring of a progression of the system as the different local minimums of G are reached at a constant temperature, as a function of established interactions and their energies, which in turn affect enthalpy and entropy.

$$\Delta G = \Delta H - T\Delta S \qquad \textbf{Eq. 1}$$

Thus, an increase in the entropy reduces the free energy of the system, promoting the assembly of more energy favourable micellar structures.[50] Such forces are important in drug delivery applications, as most APIs have poor solubility due to their relative hydrophobicity.[50,51] Moreover, novel requirements for the application of amphiphilic systems as carriers of not only hydrophobic but also hydrophilic drugs are emerging, demonstrating the potential of amphiphiles in this field.[52] vdW interactions are widely present in self-assembly, being both cumulative and non-pair additive. This means that multiple vdW interactions established upon molecular contact compensate the solo bond strength,[53] while attraction and repulsion occur not only between two particles, but into a large array of molecules.[45,47] For example, oil-water (OLW) emulsions are stabilized by vdW attraction forces. Since larger systems compel more interactions, any repulsion between hydrophilic particles and oil droplets are screened by the vdW interactions.[54] Specific interactions, such as π-π stacking, are present in molecules with aromatic rings. These interactions result in the one-dimensional incremental growth of aggregates, with the aromatic rings arranged in parallel with each other and shielding these apolar regions from polar solvents.[55]

b) Phase behaviour and its relation to the surfactant structure

The phase behaviour of a system can be characterized by parameters such as the CMC, the aggregation number (AN) or the critical packing parameter (CPP). The CMC is the concentration at which individual surfactant molecules in solution start contributing to the formation of micellar aggregates.[56] Beyond this concentration, a second CMC may take place, where micelles grow more rapidly, due to a phase transition.[57] The surfactant nature as well as thermodynamic variables, such as temperature, yield different CMC values. For instance, the CMC can be decreased by using alternative compounds and parameters, which could be beneficial to processes in which a rapid micellization is of interest.[58] Thus, the higher the hydrophobicity (by increasing the alkyl-chain tail length) of the compound the lower the CMC is, fact which enhance the molecule packing or the π-π stacking of aromatic counterions, for example.[59] Moreover, the addition of electrolytes such as salts, also decreases the CMC. Thereby, the repulsive forces between hydrophilic groups are screened, with a CMC decrease proportional to the salt concentration, above a certain threshold.[60] Another important parameters is the micelle AN, which is the number of molecules per micelle, and closely related with the CMC.[61] Finally, the CPP relates the polar head group and alkyl-chain tail areas as described in **Equation 2**. These parameters characterise the size and shape of the aggregate.[10] Thus, the CPP is obtained with the volume (V) and length (l) of the hydrophobic alkyl-chain tail as well as the hydrophilic head group (A) area.

$$\mathrm{CPP} = \mathrm{V}/(\mathrm{A} \times l) \qquad \textbf{Eq. 2}$$

Equation 2 shows that if the polar head groups are significantly larger in volume than the alkyl-chain tails, the surfactant molecules depict a conical shape and thus a spherical micelle will be formed when the $\mathrm{CPP} < 1/3$.[10] As the ratio between the hydrophilic and hydrophobic moiety increases, the molecule shifts to a cylindrical shape micelle, then to planar bilayers and from a CPP of 1 the system transits to RMs.[10,62] Thus, this parameter is demonstrated as an useful tool to predict the phase behaviour of systems containing amphiphilic molecules.

c) Phase separation in amphiphilic systems

An extreme case of aggregation is phase separation, which occurs when the solute assembles to a phase fully separated from the solvent with an interface between them.[10] The phase behaviour can be tailored by using salt to promote a phase separation,[63] or even by a temperature change.[64] Thus, a good knowledge of the phase diagram is important to relate the system properties. Phase diagrams are useful to map the composition of multiphasic systems according to important thermodynamic parameters, such as temperature or pressure, by verifying specific points in the solubility curves. For example, the temperature at which the system phase separates is known as the cloud point (CP).[65] The upper and lower critical solution temperatures (UCST and LCST, respectively) are other interesting parameters. UCST and LCST are defined as the temperatures that compose lower and upper boundaries, respectively, of the complete miscibility region[66] as it can be seen in **Figure 6**.[66] It must be noticed that other variables such as the pH may significantly affect if the system has either an UCST or LCST point. The influence of these variables in phase separation in amphiphilic systems has been reviewed before, for various types of amphiphiles, including surfactants and drugs,[67] and even polymers[68]. The latter have been recently been widely used as drug delivery vehicles. These are important for extraction and drug delivery purposes due to the ability of simply tuning certain parameters or the composition of solutions, including those containing ILs. For instance, the biphasic region may be reached at lower temperatures, enhancing the separation of the target component from the overall system.

IL systems with a LCST behaviour have been associated with tetraalkylammonium and similar cations[69] as in Saita and colleagues work.[70] The $[N_{4444}]^+$ and Cl^- are some of the ions assessed to study the effect in the LCST behaviour of tetrabutylphosphonium trifluoroacetate ($[P_{4444}]CF_3COO$), both increasing the phase separation temperature.[70] Another work developed by Schaeffer *et al.*, using tributyltetradecylphosphonium chloride ($[P_{44414}]Cl$), similar to one of the ILs used in this project, demonstrated the presence of a LCST point.[69] This rendered the use of the ABS in a more efficient way for separation processes, avoiding the addition of inorganic salts and the alteration of other system properties. An UCST behaviour was also reported for bis(trifluoromethylsulfonyl)imide ([Hbet][Tf2N]), which was used as a draw solution in an osmosis process.[71] In this case, a temperature increase allowed the miscibility of the IL solution with salt while cooling the mixture promoted phase separation for an almost salt-absent water and a reusable IL.[71]

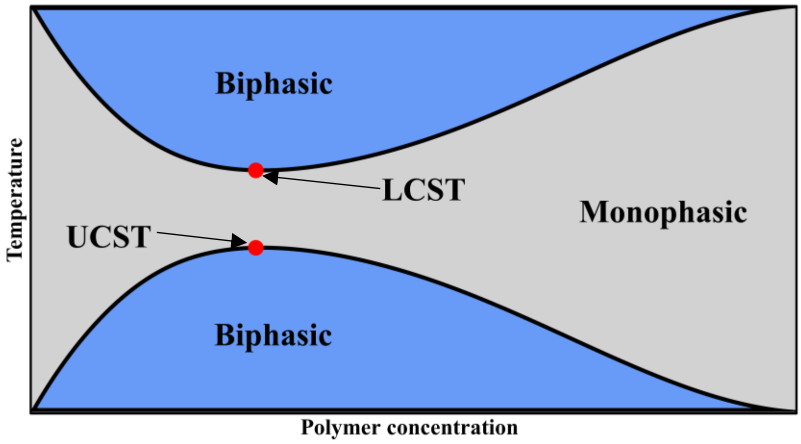


Figure 6 – Example of a phase diagram relating the concentration of a polymer aqueous solution and the temperature. The phase behaviour is depicted by either monophasic (grey) or biphasic (blue). The critical boundaries of the solubility curves, UCST and LCST points, are represented by red dots.

1.3 Ionic liquids

a) Tuneable solvent

ILs, in particular those which are liquid at room-temperature, have been reported since Walden’s work with ethylammonium nitrate, in 1914.[72] ILs are described as salts that are liquid at temperatures below an arbitrary threshold of 373 K.[2] One significant property of these systems is the negligible vapour pressure, which results in lack of volatile organic compounds emission into the atmosphere. This property is an important contributor to their labelling as so-called “green solvents”.[73] Their potential reutilization further cements this concept.[73] Other important characteristics are their negligible flammability, chemical, thermal and electrochemical stability[74] and their capacity to solvate a wide range of polar and apolar solutes.[75] The large number of possible combinations between cations and anions to build the IL system has rendered these highly tuneable “designer solvents”,[2] with the most common cations and anions used being illustrated in **Figure 7**. Finally, various ILs are amphiphilic molecules, as their constituents may assemble into surface-active structures, being usually defined as SAILs.[59] In this work, ammonium-based ILs, in particular alkyl ammonium surfactants, with a general formula of $[N_{xxxy}]^+[Cl]^-$ (x being either 1 or 4 and y either 14 or 4) will be discussed, along with their phase behaviour before and after the addition of salt.

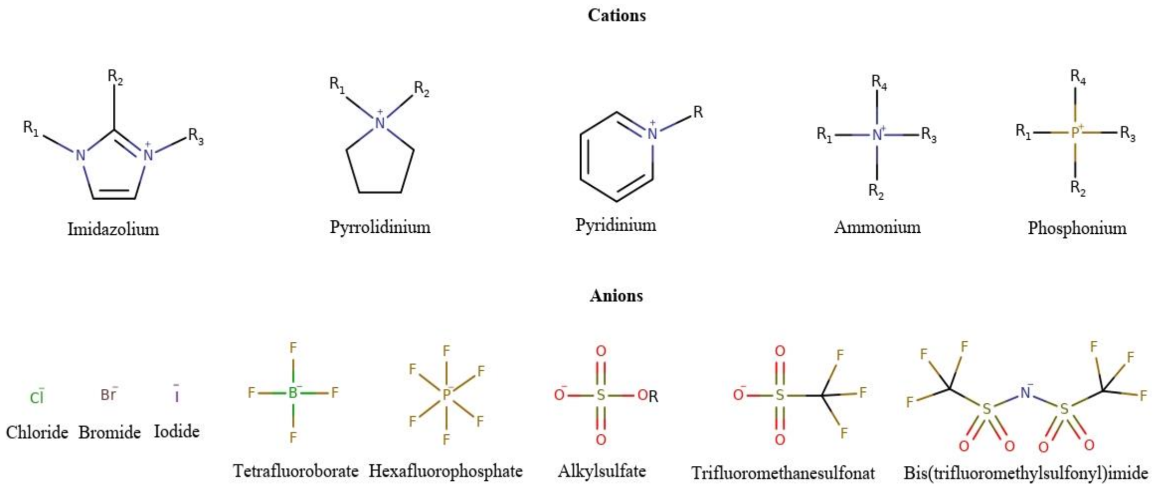

Cations
Imidazolium
Pyrrolidinium
Pyridinium
Ammonium
Phosphonium
Anions
Chloride
Bromide
Iodide
Tetrafluoroborate
Hexafluorophosphate
Alkylsulfate
Trifluoromethanesulfonat
Bis(trifluoromethylsulfonyl)imide

Figure 7 – Archetypical cations and anions commonly used as constituents of IL systems.

The IL phase behaviour is generally poorly characterized when compared to aqueous surfactant systems.[76] It must be highlighted that ionic liquids aggregate in both aqueous solutions and IL-based mediums.[76,77] ILs aggregate in various structures, depending on the driving forces implied. Model examples are depicted in **Figure 8**, ranging from ionic bulks to mesoscale structures.[78]

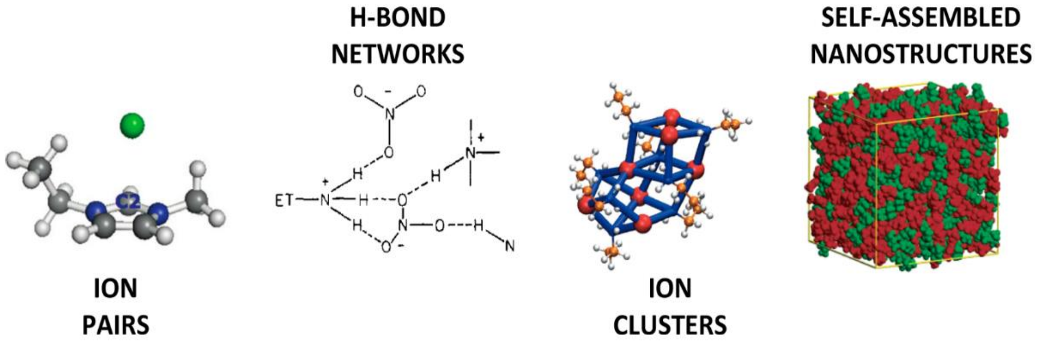


Figure 8 - Models for bulk structures of ionic liquids, from single ionic interactions (ion pairs), to small IL networks (H-bond networks) and clusters of self-assembled structures.[78]

Temperature is one of the possible driving factors of phase transition of ILs, which establish a category of temperature responsive ILs *per se*.[79] Furthermore, these systems prove valuable by being reversible, as changing temperature alters the phase behaviour of the system without resorting to other factors.[80] This is more efficient, especially energetically, if the phase transitions occur at room temperature (RT).[80] Various aqueous systems with ILs have demonstrated UCST action, being described in the previous section.

b) Applications

Current examples of IL applications are gathered in **Table 1** and **2**. ILs have been frequently investigated for extraction and separation purposes – to solve environmental concerns related with the use of organic solvents or concentration of pollutants in aqueous solutions. In Pacheco-Fernández and colleagues work,[25] octylguanidium chloride, an IL-based surfactant, was used as a less toxic solvent to remove personal care products such as parabens from cosmetic samples. The use of a long alkyl-chain decreased the CMC, turning the process into a more sustainable since less amount surfactant is required. On top of these advantages, the extraction efficiency was not damaged.[25] Various SAILs have been also used as self-aggregating components in drug delivery. For example, Vashishat and colleagues[29] were able to shed light on the molecular mechanism between an amphiphilic drug and a SAIL. Electrostatic forces as well as hydrophobic interactions and H-bonds intervene in the drug arrangement in the surfactant micelle.[29] Interestingly, surfactants comprising both an IL and an API as a counter ion, have been developed, resulting in a novel compound. In Qamar *et al*. work,[18] sodium valproate replaced bromide in CTAB, with lower drug concentrations with hydrophobic interactions as dominant over the hydrogen bonding forces.[18] This work breakthroughs the opportunity to control drug loading and further release. Some drugs may act as hydrotropes, as in the report by Singh *et al*.,[42] by directly affecting the phase behaviour of the IL-drug complex, through drug concentration changes, which triggers the API release.[42] ILs themselves have only recently been explored as hydrotropes, with a wide variety of ILs promoting the solubility of vanillin and GA as shown by Cláudio and team.[5] In this extensive assessment, it was proved that these ILs are efficient cationic hydrotropes, increasing the biomolecule solubility in water. The $[N_{4444}]Cl$ possesses the most hydrotropic effect, on par with the phosphonium counterpart; this effect is proportional to the IL cation alkyl chain size, until a certain point whereas the chloride anion has a negligible effect on hydrotropicity.[5] All these applications render trimethyltetradecylammonium chloride ($[N_{11114}]Cl$) and tributyltetradecylammonium chloride ($[N_{44414}]Cl$) as possible SAILs for drug delivery applications or phase behaviour studies involving APIs. In addition, the $[N_{4444}]Cl$ system is called to have an important contribution as hydrotrope for solubility tuning of biomolecules of interest.

2. Computational Chemistry

Computational chemistry is broadly defined as simulation of chemical systems through physical forces and mathematical calculations to provide insight into a specific issue.[81] Since the appearance of computational chemistry and MD, multiple molecular models have emerged as a result of the application of computational methods. Among those, the simulation of biomolecules is currently one of the main applications of MD simulations, such as drug delivery and underlying folding mechanisms of proteins and their functions.[3] In **Table 3**, a brief review of specific applications of biomolecule MD simulations is presented.

Table 3 – An example of recent MD simulations involving biomolecules.

Biomolecule(s)	Application	Relevant findings	Ref.
Histone H1; prothymosin-α	Binding model simulation of two proteins that remain structurally disordered	High-affinity interactions between disordered proteins is possible by electrostatic interactions; Phase separation is avoided by hydrophobic and π-stacking interactions not occurring	82
RNA tetranucleotides	Simulations in water and their complications to achieve thermodynamic agreement between MD and experimental data	MD simulations alone over stabilize dynamic conformations; Combination of both methods is required for efficient atomistic description	83
β-galactosidase; cysteine residues	Biomolecule functionalization using enzymes	Enzyme orientation in biomolecular surface affects enzyme activity; Simulations give information on the interaction dynamics that explain the previous effect	84
Rifampicin; isoniazid	Dual encapsulation of opposite drugs in a nanocarrier using MD simulations as part of the study	Drug loading occurs almost simultaneously; The driving force of self-assembly is repulsion between water molecules and drugs	85
IBP	Mechanism of drug solubility using aqueous solutions of ILs	Co-aggregates between IBP and IL are formed; ILs allow interaction between IBP and the aqueous phase (hydrotrope effect)	86
GA; bovine serum albumin	Interaction studies between GA and a common protein	Docking and MD simulations confirmed binding stability of GA at a protein site; GA binding promotes native protein stability	38
Caffeine; GA	Thermodynamic analysis of caffeine, GA and their cocrystal in water	Cocrystal formed by stacking of both molecules; This aggregate is hydrated, but maintained by weak electrostatic interactions	87
IBP	IBP encapsulation profile in pluronic L64 micelles (dissipative particle dynamics	IBP is successfully encapsulated in the polymeric micelles, mainly in their core; Micelles shrink upon loading, due to	88

	(DPD))	water expulsion and better organization	
DNA	Multiscale MD analysis of DNA condensation and phase separation	DNA was successfully represented at a mesoscale level, although sequence specificity and non-bonded interactions are not precise at a supramolecular level	89
Various intrinsically disordered proteins	Phase diagram and temperature analysis of protein phase separation	Correlations between condensed and dilute phases indicate possibility of assessing phase separations using diluted systems	90

The MD simulations opens the door to improve the development of novel pharmaceutical carriers and understanding important mechanisms such as drug loading and release, as well as separation and purification processes. The use of computational methods has the potential to reduce time and experimental resources in developing and optimizing these processes.[91] In general, previous studies demonstrated the ability of MD simulations as essential tool to provide a microscopic point of view of many complex processes.[92] A good knowledge of the interactions involved in each mechanism helps to guide research focus.[93] Still, challenges such as high resolution and precise recapitulation of biomolecules while being able to simulate long and large phenomena in MD simulations remain, requiring the employment of multiscale strategies.[94] Novel force fields (FF) and MD simulation tools and codes must be developed to overcome the inherent experimental technique limitations.

2.1 Multiscale approach

The differences between diverse computational simulation approaches from Density Functional Theory (DFT) quantum calculations to AA and CG classical MD models are presented and discussed in the context of a multiscale strategy. In this logic higher simulation scales are built upon and validated with lower ones, while still fulfilling a specific goal.[95] This application to the scales of our interest are depicted in **Figure 9**, as in this project, a multiscale approach is followed from previous quantum mechanics calculations with DFT to full atomistic and novel CG models. This leads to an integrated procedure in which the outcome of the MD simulation can enhance our acquittance in many experimental processes.[95]

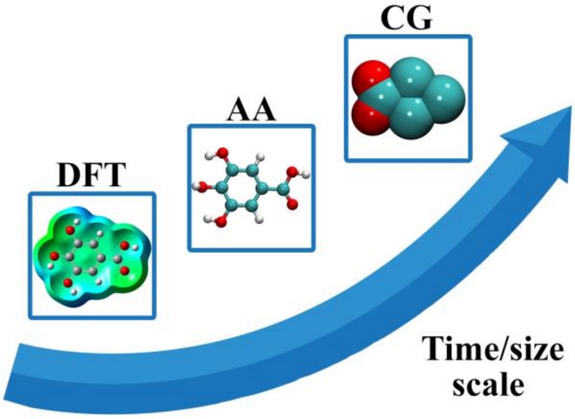


Figure 9 – Multiple scales involved in this work, using a GA molecule as an example. The multiscale strategy description applies to the development of each subsequent level in this scale.

DFT calculations explicitly tackle all electrons and allow the determination of energies at the ground-state by using electronic density distribution as the functional.[96] DFT simulations rely on solving the Schrödinger's equation and provide more accurate results than other methods based on classical mechanics. However, the large number and localization of electrons in more complex molecular systems as well as the complexity of the calculations, renders DFT calculations too demanding for applications in simulations containing more than a few components.[97] However, DFT calculations can provide an optimised starting molecular geometry point for further AA models for MD simulations. Thus, DFT is useful to study *ab initio* properties of an electronic structure and reactions, while MD allows further research on these systems due to the transition to the study of the models to a broader scale, basically the system size and time scale.[98]

The AA level for MD simulations is based on single atom descriptions in the framework of the classical Newton's equation of motion. Nevertheless, all AA models are limited in size and time scale when the mesoscale needs to be addressed due to the computational demand exponentially increases with these parameters.[99] In this regard, the CG models overcome the AA limitations and allow the study of larger sizes and time scale simulations since the number of interaction centers involved are decreased.[100] The CG level loses the atomistic resolution since each CG interaction center includes several heavy atoms.[100] Finally, the mesoscale level can be described using other levels of theory such as DPD, using larger and less particles than CG-MD (and consequently less resolution),[101] and eventually evolved into the macroscale by using novel continuum approaches. Since phase behaviour studies are the main scope of this project, CG-MD is the main approach carried out in this dissertation.

In this work, the CG models are built and validated from previous AA models following the same protocols as previously published.[92,102-103] A summary of the advantages and disadvantages inherent to the AA and CG models is presented in **Table 4**. Thus, the multiscale strategy presented in this project provides a good insight into the IL mesoscale phase formation in aqueous solutions and their interactions with biomolecules.

Table 4 - Advantages and disadvantages of using AA and CG models.

AA		**CG**	
Advantages	Disadvantages	Advantages	Disadvantages
Allows atomistic scale studies and reproduces relatively well some thermodynamic properties such as density of viscosity	Size and time scale limitations which do not allow to go further small aggregates or micelles	Allows the study of mesoscale structures	Lack the atomistic resolution and limits some atomistic detail information
Large amount of detailed information	The study of long range ordered structures are out of AA capabilities	Reduced computation time in one order of magnitude compared with the AA counterpart	The restricted energy matrix of interactions limits the molecule resolution

2.2 Molecular Dynamics

MD provides important details of atomistic interactions on the dynamic progression of a system, complementing the inherent experimental limitations in the characterization of molecular interactions.[104] Since classical mechanics are used to describe the movement of atoms in the MD simulations, some approximations are taken into account and the development of AA models requires the use of quantum mechanics to obtain geometry optimized molecules as well as the electrostatic partial charges.[105] The atomistic interactions in the MD simulation are evaluated through a FF which describes both the non-bonded and bonded interactions. Some of the most common FF currently used in MD simulations are AMBER,[106] Chemistry at HARvard Macromolecular Mechanics (CHARMM)[107], Optimized Potential for Liquid Simulations (OPLS),[108] and GROMOS.[109] All of these are included in the MD simulation package used in this project, the GROningen MAchine for Chemical Simulations (GROMACS).[110] For CG simulations, the MARTINI[111] FF is taken since it provides the parameters for interactions between the different CG interaction centers. MD package, FF, ensembles and their collaborative work towards our simulations will be described in the next section.

2.2.1 Simulations in the GROMACS package

The classical MD simulation package, GROMACS, was specifically designed for biological systems, being later proved as accurate in simulating diverse systems. In the MD simulation, the initial system coordinates are provided, and the FF is accordingly chosen to calculate the forces between atoms through the Newton's equations of motion, including the non-bonded and bonded contributions to the energy function[112] as it can be seen in **Equation 3**. The particle coordinates and velocities are subsequently updated and saved every simulation step, which in total form the overall trajectory along the simulation. The final structure, trajectories and energies obtained in the MD simulation are used to estimate the thermodynamic properties.[112] Thus, the algorithm used in GROMACS allows to compute not only the coordinates and velocities but also temperature, kinetic energy, pressure and total potential energy among others.[110]

The non-bonded interactions encompass the Lennard-Jones (LJ) potential and a Coulombic contribution, represented by **Equation 4** and **5** respectively.[113] The LJ function is an empirical potential which uses two length scale parameters, namely r^6 and r^{12}, for inter-atomic interactions (i and j).[110] The first one is related with attraction between dipoles, while the second one is the repulsion-related term, acting between electron clouds. The LJ potential uses system dependent atom pair parameters (C) to calculate the resulting energy balance between attractive and repulsive forces. **Figure 10** illustrates the interaction energy according to the distance between particles.[110] The Coulombic term takes into account the electrostatic interactions between charged particles (i and j) separated by a distance r.[110,113]

Bond-stretching, angle-bending, improper and proper dihedrals are considered for bonded interactions.[113] The bond-stretching and angle-bending are commonly governed by harmonic functions. Bond-stretching is essentially the harmonic bond length fluctuation between two covalently bonded atoms, while angle-bending is the angle variation of three covalently bonded atoms. Dihedral is the angle between the two planes, each one including three bonded atoms, formed by four covalently bonded atoms. Several dihedral potential functions can be selected to reproduce the proper geometry of the molecule.[113]

$$E_T = E_{bonds} + E_{angles} + E_{dihedrals} + E_{non\text{-}bonded} \qquad \textbf{Eq. 3}$$

$$V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{(12)}} - \frac{C_{ij}^{(6)}}{r_{ij}^{(6)}} \qquad \textbf{Eq. 4}$$

$$V_C(r_{ij}) = f\,\frac{q_{ij}q_{ij}}{\varepsilon_r r_{ij}} \qquad \textbf{Eq. 5}$$

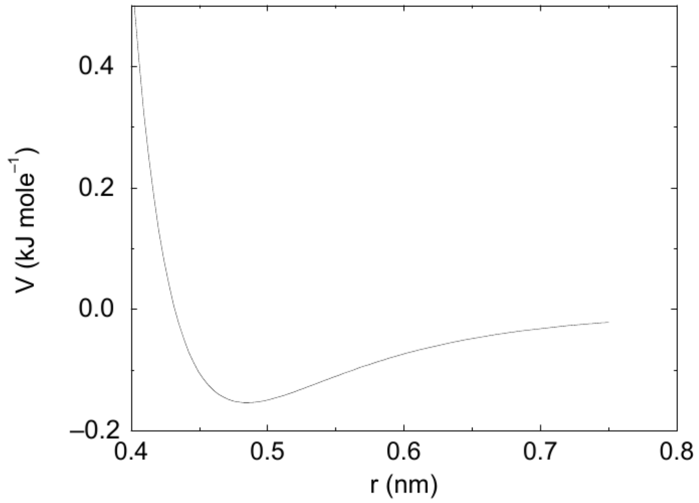


Figure 10 - Lennard Jones potential, relating the energy of interaction (V) and the distance between particles (r). The minimum value indicates the optimal balance between the repulsion and the attraction terms.[110]

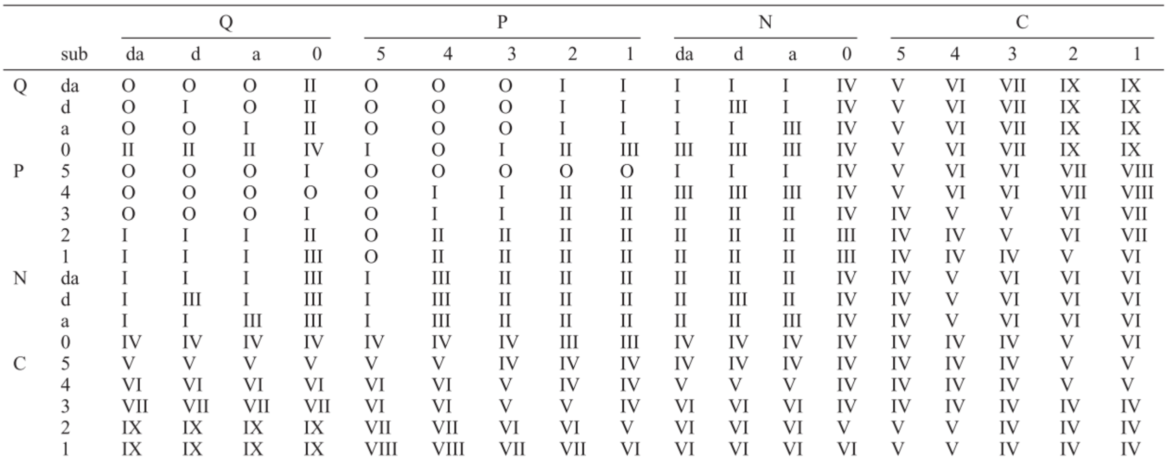
In the MD simulations, an energy minimization (EM) and equilibrium protocol is followed prior the production runs; the EM step is required to avoid any close contact between atoms or molecules in the initial configuration.[114] Rearrangement of the system into an optimized, energy-minimized state is made to position the molecules in the way that a minimum, realistic distance between atoms is achieved.[114] The NVT and NpT equilibration ensembles are commonly used to set temperature and pressure of the system, respectively. The thermodynamic equilibrium is essential to compare the MD results to real systems, in which further modification of the initial MD parameters might be done as a validation step.[116] In the NVT equilibrium step, the simulation box volume is fixed and the temperature coupling is performed to establish the proper kinetic energy (temperature).[117] Then, in the NpT ensemble, the pressure is stated, and the correct simulation box density is found.[117] The initial equilibration steps ensure that the proper thermodynamic equilibrium is reached for the NVT, NpT or other ensembles production runs.

2.2.2 Coarse-Grained models

CG frameworks are a progression in the multiscale strategy applied in this work to reach the mesoscale level. This model becomes relevant in systems where self-assembly occurs in long range ordered structures where a large system size is required. The CG mapping procedure departs from an atomistic structure used as a reference and then, the number of interaction centers are properly reduced. The CG mapping procedure consists of including several atoms into one interaction center (bead) which mimics the essential physico-chemical characteristics of the included group of atoms.[118] This procedure considerably reduces the number of degrees of freedom and computer resource requirements.[118] CG models require the use of specific force fields in MD such as MARTINI.[111] In this case, each bead represents four, three or two heavy atoms connected by bonds and angles while dihedrals are usually not necessary.[111] Overlapping between CG beads is occasionally required to realistically reproduce certain characteristics and the molecule geometry.[111]

The MARTINI FF includes an empirical energy matrix to establish the interactions between different beads.[111] This interaction matrix is shown in **Figure 11** and the MARTINI model mainly includes four types of beads, namely Q, P, N and C for charged, polar, non-polar (half polar-half apolar) and apolar beads, respectively, with different sub-types each as described in **Figure 11**.

The final system modelling is achieved by selecting the proper interaction energies from the matrix of interactions and taking the AA behaviour as a reference as well as the experimental data counterparts. The bonded parameters in the MARTINI model are described by harmonic functions, with a constant force being used. By default, the bond length between interaction centers is established at the default CG bead diameter, $\sigma = 0.47$ nm while the force is 1250 $\mathrm{kJ\ mol^{-1}\cdot nm^{-2}}$.[111] Angles are also described by harmonic potentials, with bond angles between beads being 180° for simple aliphatic chains and trans-unsaturated bonds and 120° for cis-unsaturated bonds.[111] Force constants are set at 25 $\mathrm{kJ\cdot mol^{-1}}$ and 45 $\mathrm{kJ\cdot mol^{-1}}$ for saturated and unsaturated bonds, respectively.[111] Aromatic rings are particular models, as a different mapping ratio of 2 or 3 to 1 is required to reproduce the CG ring.[111] In this case the suffix "S" indicates a special mapping with reduced σ to 0.43 nm and the energy of self-interaction is 75% of the initial value.[111] In certain cases, constraints such as dihedrals might be necessary to assure the ring remains planar. Although these values are the default ones that provide optimal circumstances for initial CG-MD simulations, these can be altered to better resemble the AA structure from which the CG mapping procedure was previously conducted.



		Q				P					N				C				
	sub	da	d	a	0	5	4	3	2	1	da	d	a	0	5	4	3	2	1
Q	da	O	O	O	II	O	O	O	I	I	I	I	I	IV	V	VI	VII	IX	IX
	d	O	I	O	II	O	O	O	I	I	I	III	I	IV	V	VI	VII	IX	IX
	a	O	O	I	II	O	O	O	I	I	I	I	III	IV	V	VI	VII	IX	IX
	0	II	II	II	IV	I	O	I	II	III	III	III	III	IV	V	VI	VII	IX	IX
P	5	O	O	O	I	O	O	O	O	O	I	I	I	IV	V	VI	VI	VII	VIII
	4	O	O	O	O	O	I	I	II	II	III	III	III	IV	V	VI	VI	VII	VIII
	3	O	O	O	I	O	I	I	II	II	II	II	II	IV	IV	V	V	VI	VII
	2	I	I	I	II	O	II	II	II	II	II	II	II	III	IV	IV	V	VI	VII
	1	I	I	I	III	O	II	II	II	II	II	II	II	III	IV	IV	IV	V	VI
N	da	I	I	I	III	I	III	II	II	II	II	II	II	IV	IV	V	VI	VI	VI
	d	I	III	I	III	I	III	II	II	II	II	III	II	IV	IV	V	VI	VI	VI
	a	I	I	III	III	I	III	II	II	II	II	II	III	IV	IV	V	VI	VI	VI
	0	IV	IV	IV	IV	IV	IV	IV	III	III	IV	IV	IV	IV	IV	IV	IV	V	VI
C	5	V	V	V	V	V	V	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	V	V
	4	VI	VI	VI	VI	VI	VI	V	IV	IV	V	V	V	IV	IV	IV	IV	V	V
	3	VII	VII	VII	VII	VI	VI	V	V	IV	VI	VI	VI	IV	IV	IV	IV	IV	IV
	2	IX	IX	IX	IX	VII	VII	VI	VI	V	VI	VI	VI	V	V	V	IV	IV	IV
	1	IX	IX	IX	IX	VIII	VIII	VII	VII	VI	VI	VI	VI	VI	V	V	IV	IV	IV

Figure 11 - MARTINI energy matrix of interaction.[111] Each level of interaction is used as the energy of interaction (ε) in the LJ calculations and the bead size (σ) is 0.47 nm by default. Beads are categorized as charged (Q), polar (P), non-polar (N) and apolar (C). In charged Q beads, d, a, da or 0 indicate the hydrogen bond capabilities; donor, acceptor, donor and acceptor and none, respectively. 1 to 5 indexes indicate increasing polar or apolar affinity being 1 the lowest polar/apolar character. From O to IX, the interaction strength is: 5.6 kJ/mol, 5.0 kJ/mol, 4.5 kJ/mol, 4.0 kJ/mol, 3.5 kJ/mol, 3.1 kJ/mol, 2.7 kJ/mol, 2.3 kJ/mol, 2.0 kJ/mol, 2.0 kJ/mol (σ = 0.62 nm at the IX selection).

3. Objectives

The aim of this work is to study and characterize several systems using molecular dynamics with the GROMACS package. The use of this method allows not only the visualization of the progress of the systems, in particular the phase transitions that may occur, but also the assessment of the possible driving interactions behind each phenomenon. A proper validation of ammonium-based IL solutions, GA and IBP systems, in which novel CG representations were developed, is essential for future simulations and proper focus of experimental research. The application of a multiscale strategy ensures the proper model scaling. This approach is critical in studies such as the presented in this project in which an adequate CG representation of the molecules must be ensured. Moreover, the time and size scale at which phase transition and separation are detected properly demand the use of this CG level. The main contribution of this work is the proper understanding of the molecular mechanisms that drive the phase behaviour and the biomolecule partition of API in IL-based aqueous solutions. In particular, our aim is to determine the specific interactions driving phase behaviour of GA and eventually IBP in micellar and hydrotropic systems, their differences and similarities. This knowledge allows the optimization and novel development of drug delivery systems comprising both ILs, APIs and similar molecules. Besides, this project proposes to discuss the potential of these systems for alternative applications, such as separation and purification of these APIs from aqueous solutions. Other secondary objectives, such as the achievement of novel API-based ILs, through efficient processes, can be explored. Overall, the obtained results can also establish guidelines for future studies using different biomolecules and even other solvents. Once a well-established framework is obtained, one can use the procedure as a platform for other molecules. In particular, the APIs considered in this project, which are scarcer at the CG level. For the sake of time and extent of this dissertation, only the development and validation of the CG model of IBP is presented. However, future work is already underway regarding the use of this API in the systems designed for this project. A flowchart with the general process of this project is schematically illustrated in **Figure 12**.

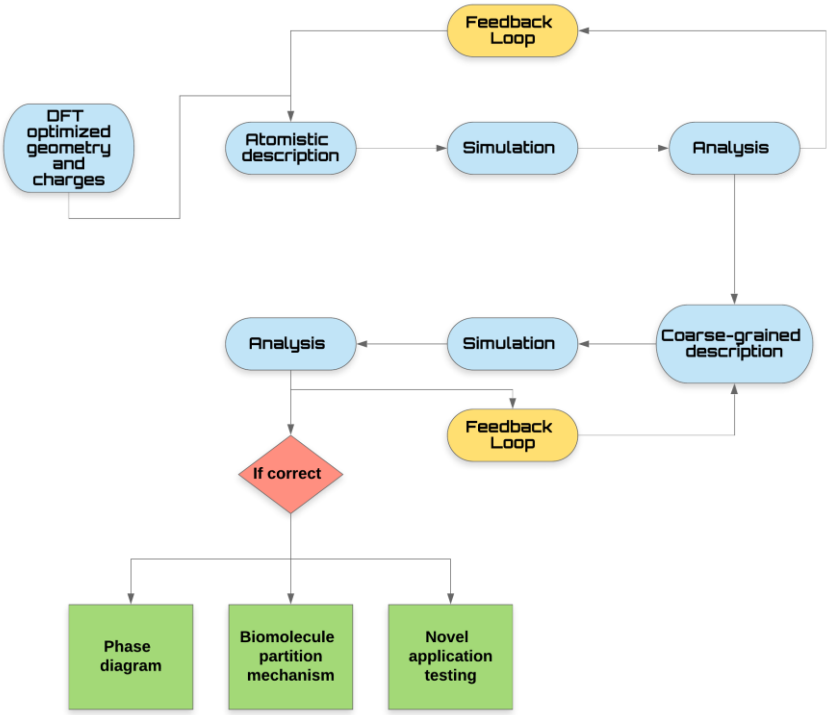

Feedback Loop
DFT optimized geometry and charges
Atomistic description
Simulation
Analysis
Coarse-grained description
Simulation
Analysis
Feedback Loop
If correct
Phase diagram
Biomolecule partition mechanism
Novel application testing

Figure 12 - Flowchart of the general line of thought used in this project and the main objectives.

Chapter II: Methodology

1. Simulations details

All MD simulations were carried out with the GROMACS package (version 5.1.5)[110] in both AA and CG descriptions. Each specific framework will be described in the following sections including the CG model validations. Initial simulation boxes were built using the GROMACS *gmx insert-molecules* tool to place all molecules in the simulation box with random starting positions. The MD production runs, otherwise stated along 0.5 μs of simulation time, were developed prior an equilibration protocol previously described.[92,102-103]

The specific parameters of each ensemble, as well as their topology, are compiled in **Table 7**. The system numbers follow the indexation presented in the table. The simulation time is defined as the number of simulation steps multiplied by the time step, 2 and 10-30 fs for AA and CG, respectively. The leapfrog algorithm[119] was used in the equilibrium and production runs. The pressure was fixed at 1 bar with an isotropic coupling and Parrinello-Rahman barostat[120] whilst the temperature was established either with the velocity-rescaling[121] in the equilibrium steps and the Nose-Hoover[122] thermostat in production runs.

In this work the cut-off radius for LJ and Coulombic interactions was set as 1.2 nm, with Particle-Mesh-Ewald (PME),[123] and potential-shift functions being used to calculate long-range electrostatic interactions, as in previous works developed by Pérez-Sánchez *et al*.[92,102-103] Periodic boundary conditions (PBC) in all directions were considered. In the LJ interactions, the force-switch function was used smoothly approaching to zero from ~ 0.9 nm to the 1.2 nm cut-off. The LINear Constraint Solver (LINCS) was used to constrain all bonds.[124] Rectangular boxes were used in systems containing IL and GA to ease visual analysis of the density profiles, while cubic and dodecahedric boxes in big systems were used for the rest of the simulations, particularly for validation and final runs, respectively. Dodecahedric boxes provided optimized simulations as less water molecules are necessary to fill the unit cell and represent the bulk solution in PBC.[125] Cubic boxes were used for the validation simulations of GA and IBP due to the small system size.

In all of MD simulations carried out in this project, the equilibrium was monitored following two steps: (i) visual analysis of the trajectories to ensure that the system was not suffering any phase transitions or unexpected variations and (ii) using the GROMACS *gmx* *energy* tool to monitor the total energy of the system throughout the trajectory, which should be defined when the plateau is reached to assure an equilibrium state.

The MD input and resulting outputs files are summarized in **Table 5** and further information can be consulted in the GROMACS manual.[110]

Table 5 - Input and output files used in the GROMACS MD simulations.

Input	Action	Output	Action
.gro	Molecular structure coordinates	.gro	Final molecular structure coordinates (resulting from simulation)
.itp	Specific topology of a system component, such as atoms and restraints (bonds, angles, pairs, dihedrals). Included in a .top file.	.trr	Trajectories obtained in the simulation: Coordinates, velocities, forces and energies
.top	Topology of the system, containing the number of molecules and individual topology files (.itp)	.xtc	Compressed version of .trr files (only coordinates)
.mdp	Contains the run parameters for the simulation (time steps, thermostat, barostat, etc.)	.edr	Storage of energies written during simulation
.tpr	Compilation of all previous information required to start the simulation	.log	Information gathered throughout the full simulation process

2. All-atom molecular dynamics (AA-MD) systems

In this work AA-MD systems were used as a direct computational validation tool of novel CG descriptions of GA and IBP. The same logic was applied to these ILs based on previous studies, validating the quaternary ammonium surfactants also used throughout this work.[92,102-103] Overall, the optimal partial charges were calculated using the CHarges from Electrostatic Potentials using a Grid-based method (CHELPG),[126] with the Becke three-parameter Lee-Yang-Parr (B3LYP) exchange-correlation functional used with the 6-311+G(d,p) basis set for DFT calculation verification of the structures, according to Vueba *et al.* previous work.[127]

2.1 Ammonium-based ILs

The $[N_{11114}]Cl$, $[N_{44414}]Cl$ and $[N_{4444}]Cl$ ILs were considered in this project mainly to explore the effect of the IL headgroup and the effect of the alky-chain length, thus the solubilization mechanism involved. The simulations were carried out for at least 20 ns with a time step of 2 fs, at 298K, using the force field parameters from the OPLS-AA FF version developed by Canongia Lopes *et al.*[128,129] for the cationic surfactant and from Åqvist[130] for the chloride ions whereas the SPC/E model[131] was chosen to model the water.

2.2 GA and IBP

A new CG model for GA and the IBP molecules were developed in this project, which required the AA description for validation purposes. It must be noticing that the pH is essential in many biochemical phenomena.[132] However, in the MD simulations, the pH must be fixed since the protonation states are stationary throughout the simulation.[133] Thus, two AA versions of these molecules were generated; one with the carboxylic acid group protonated (COOH) and another deprotonated (COO^-). The addition of the species used to measure pH, hydronium and hydroxide, would not correctly induce a specific pH. Thus, we designed separate simulations for each protonation state to be able to address the effect of pH, as it will be explained in the **Results and Discussion** chapter. To achieve this, on each initial configuration of the simulation, sodium counter cations were added to assure the neutrality of the solution. The number is equal to the GA/IBP amount multiplied by the number of dissociated protons on each molecule, since these acids dissociate when dissolved in water. This corresponds to the acid-base titration principles in which in this case, sodium hydroxide would be used to neutralize acidic molecules such as GA and IBP, as described in **Equation 6**. The M in this equation indicates either the molecule with the group where proton dissociation may occur (no stoichiometric ratios were considered).

$$MH + NaOH \rightarrow NaM + H_2O \qquad \textbf{Eq. 6}$$

This provides results of GA and IBP behaviour at pH values broadly below and above its first dissociation constant at 298K. A second dissociation constant was also analysed for GA, in production runs also comprising ILs, in which the hydroxyl groups were also deprotonated. A labelling of “Prot. GA”, “Deprot. GA (-1)” and “Deprot. GA (-3)” was used to describe GA from below the first dissociation constant, between both constants and above the latter, when convenient. A similar labelling as GA was used to distinguish IBP at pH below and above the pKa ("Prot. IBP" and "Deprot. IBP"). Reference values of interest for the protonation state, such as pKa, and of solubility limits used in this work, as well as the octanol-water partition coefficient (K_{ow}), of interest in partition studies, are displayed in **Table 6** for GA and IBP.

Table 6 – Reference values related to pH, solubility and partition used for GA and IBP.

	GA	IBP
pKa	4.20[134]/4.44[135] (COOH); 8.54[134]/8.45[135] (OH); 10.05[135]; 11.30[135]	4.40[136] (COOH)
Solubility limit (wt.%)	1.470 (pH 6.9)[137]	0.23 (pH 7.4);[138] 0.005 (pH 1.2);[139] 0.20 (pH 6.8)[139]
Log(K_{ow})	0.70[140]	3.97[140]

The initial protonated GA structure was obtained as Protein Data Bank (PDB),[141] while the IBP structure was obtained from previous DFT calculations. The final optimized AA models were obtained recurring to this computational level in both molecules. MD simulations were carried out using the parameters described in **Table 7** (systems 1 and 2 for GA and 5 and 6 for IBP). The general procedure was explained previously, with a 2 fs timestep and at concentrations below the solubility limit of GA and below or near to the IBP experimentally reported value at around 298K. The AA structures of GA and IBP as well as the partial charges used in these production runs are illustrated in **Figure SI:1** and **Figure SI:2**, respectively. The electrostatic potential of each section of these molecules is presented as a contour map in **Figures 13** and **14**. These were used as a validation tool to assure that the novel CG models have bead polarities matching an adequate charge distribution when compared to the AA level.

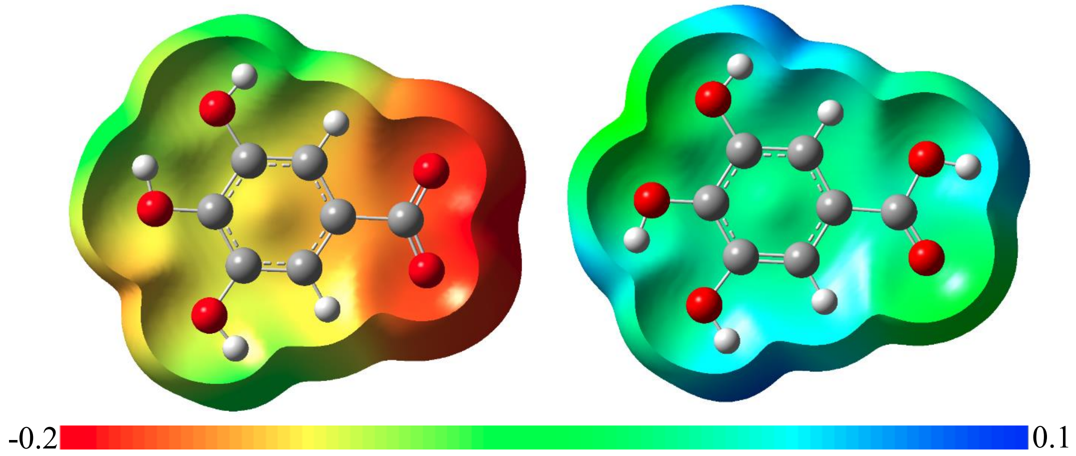


Figure 13 - Charge distribution comparison between the deprotonated (left) (system 2) and protonated (right) (system 1) version of GA, related to the carboxylic acid group. These were calculated with the B3LYP/6-311+G(d,p) level of theory. Red colour indicates negatively charged sections while blue colour illustrates positively charged regions.

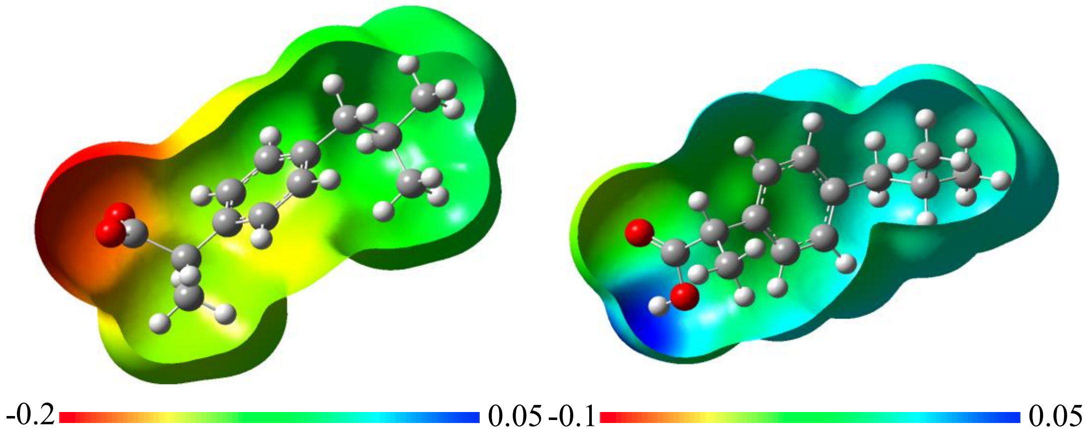


Figure 14 – Charge distribution comparison between the deprotonated (left) (system 6) and protonated (right) (system 5) version of IBP. These were calculated with the B3LYP/6-311+G(d,p) level of theory. Red colour indicates negatively charged sections while blue colour illustrates positively charged regions.

3. AA mapping to CG and validation

A mapping procedure was followed to obtain the novel CG model from the AA reference. In this work, the MARTINI FF (version 2.2)[111] was used in the CG MD simulations. The CG mapping used for each molecule in this work is depicted in **Figures 15-20**. In each figure, the AA groups that are converted to CG beads have the same colour, with the selected MARTINI polarity being depicted in blue. The centers of interaction for the CG beads are presented in red, with the atoms being grouped contained in a blue area. The mapping employed for the ILs is based on previous works, for both $[N_{11114}]Cl$[102], $[N_{44414}]Cl$[69] and $[N_{4444}]Cl$[69].

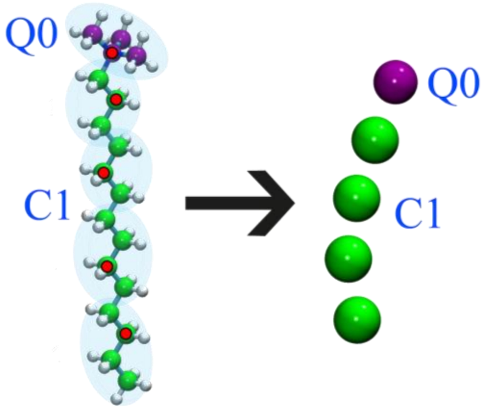


Figure 15 – AA to CG mapping for $[N_{11114}]^+$, with a charged polarity being selected for the surfactant head (Q0), and an apolar one for the alkyl tail (C1).

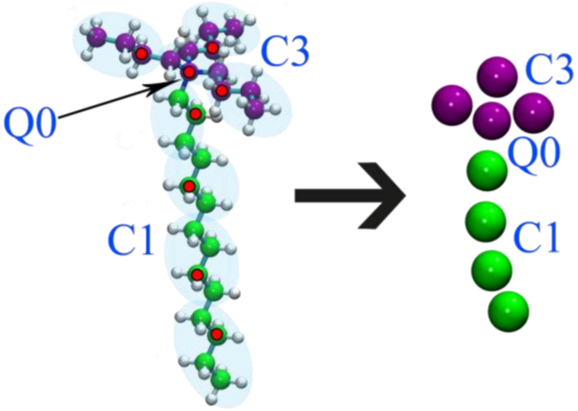


Figure 16 – Mapping from AA to CG of $[N_{44414}]^+$. The center bead of the polar head is set as charged (Q0), while the butyl chains of head are mildly apolar (C3). The alkyl tail is described as apolar (C1), exactly as in $[N_{11114}]^+$.

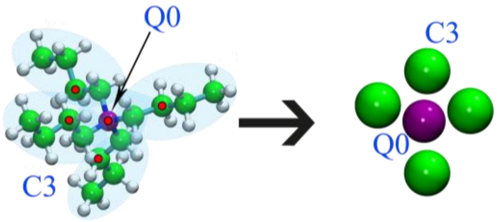


Figure 17 – AA mapping to CG of $[N_{4444}]^+$. Similarly to $[N_{44414}]^+$, the central ammonium atom is set as charged (Q0), along with the initial carbon atoms surrounding it. The butyl chains are set as mildly apolar (C3).

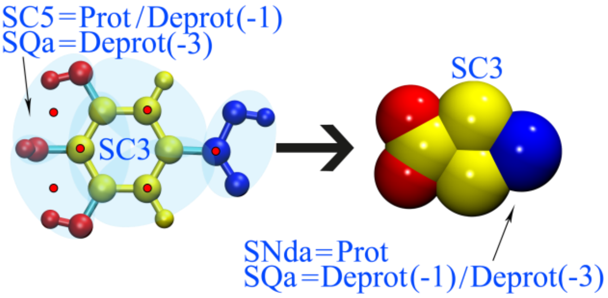


Figure 18 – GA CG mapping from the AA structure reference. A special mapping was employed, labelled by an "S" preceding the actual bead polarity. The aromatic ring was set as mildly apolar (SC3), while the carboxylic acid and hydroxylic groups had different polarities depending on the simulated pH, and consequently their protonated states. The former ranged from nonpolar (SNda) to charged (SQa) while the latter ranged from weakly apolar (SC5) to charged (SQa).

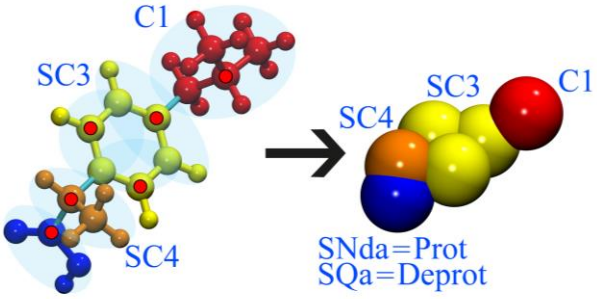


Figure 19 – IBP mapping from AA to CG. While the most apolar region is defined by a default mapping, apolar (C1), the rest of the procedure uses the special mapping, also applied to GA. The aromatic ring is again mildly apolar (SC3), while the intermediate alkyl region between it and the carboxylic acid group is more polar (SC4). The latter group is either set as nonpolar (SNda) if protonated or charged (SQa) if deprotonated.

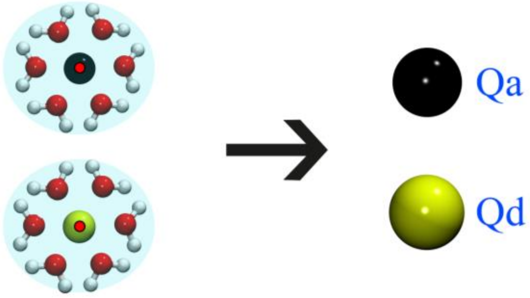


Figure 20 – Counter ion mapping procedure from AA to CG. Chloride was used as anion, for the ionic liquids used, being described as charged and hydrogen bonding acceptor (Qa), while sodium as used as cation, to charge balance deprotonated GA and IBP, being set as charged and donor in hydrogen bonds (Qd). Both CG mappings have six implicit water molecules.

Water followed a similar CG mapping, in the way that multiple water molecules were combined into a single bead. However, incorrect interactions would often be established between water and other particles, due to lack of adequate polarization of this solvent in the default CG MARTINI water model. Thus, the PW model was used in this project in order to capture the water dipole effect to maintain the behaviour of water as close as possible to the real water.[142] The PW water model, depicted in **Figure 21**, allows an adequate screening of electrostatic charges, although a higher computational demand. However, the PW model resembles better the real water compared with the regular CG water model. It must be noticed that the polarizability may affect the phase behaviour since water dipoles can overestimate their interaction with the charged groups of amphiphilic molecules.[143]

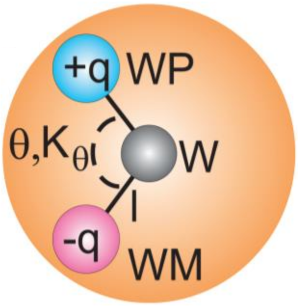


Figure 21 – Polarizable CG water PW, containing three particles: a central center of mass (W) and two charged dummy interaction centers, a positive (WP) and a negative (WM) to mimic the orientational polarizability of water. Both are connected to W at a distance of *l* and can harmonically vibrate with a spring constant of K_θ (harmonic function).[143]

The CG mapping is established on the base of AA descriptions and experimental results, resulting frequently in an exhaustive procedure where several mappings are attempted until the CG model matches the AA reference behaviour. Generally, the CG molecule geometry is firstly analysed in vacuum to ensure the molecular stability and compared with the AA. The main angles and distances between interaction centers are verified and modified in case these do not reproduce the AA counterpart. Once the geometry is properly reproduced in the CG model, MD simulations are carried out under similar conditions for a direct comparison with literature data. Final validation of these systems was done visually and resorting to GROMACS tools such as radial and spatial distribution functions, which are explained in the **Analysis** section. The CG descriptions of the quaternary ammonium surfactants also involved the use of density profiles and even micelle aggregation numbers. Since the validation of those systems was already carried out in previous works, only the novel CG descriptions developed in this work will be discussed, in the **Results and Discussion** chapter.

4. Coarse-grained molecular dynamics (CG-MD) systems

The MD-CG simulation carried out in this project are presented in **Table 7**. Previous developed ILs CG models for [N_{11114}]Cl, [N_{44414}]Cl and [N_{4444}]Cl were used in this work. The phase behaviour of these ILs in aqueous solutions and how the salt alters the phase behaviour (for the first two) is explored. More importantly, ILs aqueous solutions with GA are considered to further understand partition of GA besides other compounds of pharmaceutical interest, such as IBP.

All simulations of ILs with water and salt (**Table 7**: systems 9-14) were conducted at 390K and 1 bar. Three simulations at 298K were used as controls for systems containing both ILs and GA (**Table 7**: systems 15, 22 and 29). These were required for a direct comparison between the systems with ILs and with or without GA, as the ones did for [N_{11114}]Cl and [N_{44414}]Cl ran at 390K (**Table 7**: systems 9 and 12).

5. Analysis

Visual and statistical tools were used as a post-processing analysis of the results obtained in the MD simulations. The final coordinates of the molecules were visualized using the Visual Molecular Dynamics (VMD) software package.[144] Each snapshot or visual representation directly obtained with VMD used a graphical drawing of either vdW or CPK (for CG and AA descriptions respectively). The MD rendered simulation snapshots were obtained with the VMD program as well as the simulation movies and processed with the Blender open-source software. The usual GROMACS analysis tools, the radial distribution functions (RDFs), the solvent-accessible surface areas (SASA) and the density profiles were also used to analyse the MD simulations. In these situations, an index describing specific groups in the molecules of the system or even groups of molecules was previously created with the GROMACS *gmx make_ndx* tool for further assays.

The RDF functions, $g(r)$, were obtained with the *gmx rdf* tool of GROMACS. These calculations indicate the probability of finding the selected particles at a distance r of the reference groups, *i.e.*, the density (ρ) of a particular region in the system compared to the bulk density of the solution, as shown in **Equation 7**.

$$g(r) = \frac{\rho(r)}{\rho^{bulk}} \qquad \textbf{Eq. 7}$$

The RDF profile not only servers for a qualitative analysis of the molecular organization of the system, but also the coordination number around the reference particles, by integrating the RDF function at a defined distance *r*. A representative scheme of the detection of a defined target particle or group around the reference is depicted in **Figure 22**. In case of PW as a refence in the RDF calculation, the *g*(*r*) values were multiplied by 3, since the PW model contains three implicit water particles. In those situations, *g*(*r*) was replaced by *g**(*r*).

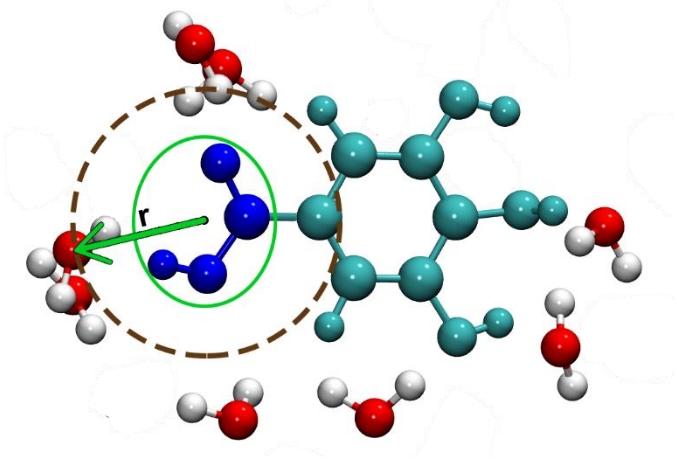


Figure 22 – Schematic representation of the RDF calculation using GA (in dark blue and cyan) and water (red and white) as example. In this case, the carboxylic acid group of the GA molecule is used (dark blue), from which center of mass (COM) of the selected water molecules (red and white) is performed at a given distance *r*. At each value of *r*, the probability of finding water around the carboxylic acid group of GA is calculated and described by *g*(*r*).

Commonly, a qualitative visual inspection of the system requires a quantitative point of view depicted by the statistical information obtained with RDFs. The spatial distribution functions (SDFs) provide a 3D scenario of the information gathered by the RDFs. Thus, the degree of solvation of a certain molecule and its association with other molecules in the system, such as the π-π stacking of aromatic rings and ionic interaction, can be explored in detail. This analysis was used to validate the new CG models of GA and IBP developed in this project. The SDFs were obtained with the TRajectory Analyzer and VISualizer (TRAVIS)[145] tool.

Finally, the SASA is calculated by the assessing the surface of the reference molecules regions that are available to solvent molecules. The GROMACS *gmx sasa* tool uses a double cubic lattice algorithm[146] which divides the tri-dimensional space in the simulation box into smaller cubic sub-divisions. Then, calculations are performed by considering the contact points between a probe molecule (solvent) to the atoms of a system, bearing in mind their vdW radii. The application of this algorithm is presented on **Figure 23**.

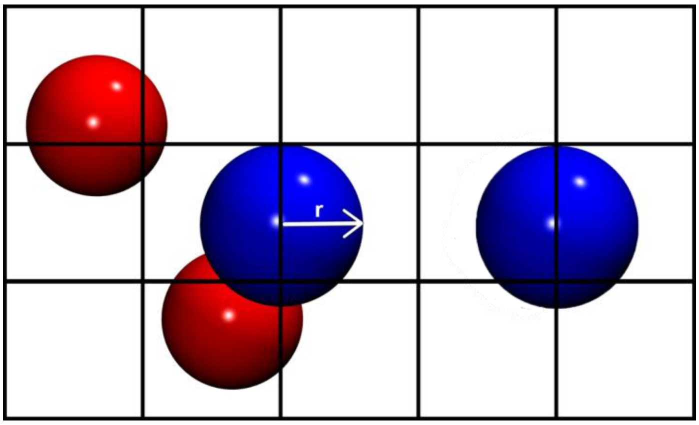


Figure 23 – Two-dimensional view of the double cubic lattice method to calculate SASA. A three-dimensional grid is displayed in the system, with *r* indicating the maximum atomic radius of the reference atoms (in blue). This radius is given by the sum of the vdW radius and the accessible surface to a probe molecule. The SASA is calculated using the disposition of the selection atoms (in red) towards the reference ones, considering the atomic radius.

The density profiles were also compiled to estimate the density, either exact or relative, of each component of the system alongside the simulation box. In systems with spherical micelles or micellar-like structures, a cluster counting code based on the Hoshen-Kopelman algorithm[147] was used. This code compiles the trajectories of the MD simulations to obtain the density of each selected group from the micelle COM. The micelle radius is the distance from the micelle core (end groups of the surfactant alkyl-chain tail) to the surfactant head groups. The code also allows to estimate the average aggregation number of the micelle.

Since this cluster counting code is limited to spherical or micellar-like structures, the *gmx density* tool of GROMACS was used to characterise the phase behaviour above the micellar region. Thereby, the simulation box is split in different slices in which the density profile is estimated alongside the cartesian directions.

Table 7 – Simulation parameters of each system, regarding the production run, with the respective notation. Box sizes are set as "3" for cubic boxes, "2" for rectangular ones and "dode" for dodecahedric boxes.

System	n_{IL}	[IL] (wt.%)	$n_{GA/IBP}$	[GA/IBP] (wt.%)	n_{H_20}	$n_{NaCl/Na}$	[NaCl/Na] (wt.%)	Temp. (K)	Sim. time (µs)	Final box size (nm)
Validation of CG models of GA and IBP										
(1) Prot.GA (AA) + H_2O	-	-	10	0.94	10000	-	-	298	0.04	7^3
(2) Deprot. GA (-1) (AA) + H_2O + Na	-	-	10	0.93	10000	10	0.13	298	0.04	7^3
(3) Prot.GA (CG) + H_2O	-	-	10	0.94	3333	-	-	298	0.1	7^3
(4) Deprot. GA (-1) (CG) + H_2O + Na	-	-	10	0.92	3333	10	0.13	298	0.1	7^3
(5) Prot. IBP (AA) + H_2O	-	-	10	0.23	50000	-	-	298	0.05	11^3
(6) Deprot. IBP (AA) + H_2O + Na	-	-	10	0.23	50000	10	2.55	298	0.05	11^3
(7) Prot. IBP (CG) + H_2O	-	-	10	0.23	16667	-	-	298	0.15	13^3
(8) Deprot. IBP (CG) + H_2O + Na	-	-	10	0.23	16667	10	2.54	298	0.15	13^3
Ammonium surfactants with or without salt addition										
(9) $[N_{11114}]$Cl + H_2O	4000	42.19	-	-	21608	-	-	390	2	20^2, 14 (dode)
(10) $[N_{11114}]$Cl + H_2O + low NaCl	4000	41.84	-	-	20008	400	0.84	390	1	20^2, 14 (dode)
(11) $[N_{11114}]$Cl + H_2O + high NaCl	4000	32.76	-	-	21608	2900	4.75	390	1	20^2, 14 (dode)
(12) $[N_{44414}]$Cl + H_2O	4000	46.06	-	-	28250	-	-	390	1	22^2, 16 (dode)
(13) $[N_{44414}]$Cl + H_2O + low NaCl	4000	45.76	-	-	26650	400	0.64	390	2	22^2, 16 (dode)
(14) $[N_{44414}]$Cl + H_2O + high NaCl	4000	38.91	-	-	26650	3675	4.72	390	2	23^2, 16 (dode)
Ammonium surfactants + GA										
(15) $[N_{11114}]$Cl + H_2O	500	10.02	-	-	23250	-	-	298	1	12^2, 23
(16) $[N_{11114}]$Cl + H_2O + low Prot. GA (CG)	500	9.99	25	0.29	23250	-	-	298	0.5	12^2, 23
(17) $[N_{11114}]$Cl + H_2O + low Deprot. GA (-1) (CG) + Na	500	10.01	25	0.29	23150	25	0.04	298	0.5	12^2, 23
(18) $[N_{11114}]$Cl + H_2O + low Deprot. GA (-3) (CG) + Na	500	10.00	25	0.28	23050	75	0.12	298	0.5	12^2, 23

(19) $[N_{11114}]Cl$ + H_2O + high Prot. GA (CG)	500	9.79	200	2.28	23250	-	-	298	1	12^2, 23
(20) $[N_{11114}]Cl$ + H_2O + high Deprot. GA (-1) (CG) + Na	500	9.66	200	2.24	23150	200	0.30	298	1	12^2, 23
(21) $[N_{11114}]Cl$ + H_2O + high Deprot. GA (-3) (CG) + Na	500	9.37	200	2.13	23050	600	0.89	298	1	12^2, 23
(22) $[N_{44414}]Cl$ + H_2O	500	10.02	-	-	33750	-	-	298	1	
(23) $[N_{44414}]Cl$ + H_2O + low Prot. GA (CG)	500	10.00	25	0.20	33750	-	-	298	1	13^2, 26
(24) $[N_{44414}]Cl$ + H_2O + low Deprot. GA (-1) (CG) + Na	500	10.01	25	0.20	33650	25	0.03	298	1	13^2, 26
(25) $[N_{44414}]Cl$ + H_2O + low Deprot. GA (-3) (CG) + Na	500	10.08	25	0.20	33250	75	0.08	298	1	13^2, 26
(26) $[N_{44414}]Cl$ + H_2O + high Prot. GA (CG)	500	9.86	200	1.60	33750	-	-	298	1	13^2, 26
(27) $[N_{44414}]Cl$ + H_2O + high Deprot. GA (-1) (CG) + Na	500	9.76	200	1.58	33650	200	0.21	298	1	13^2, 26
(28) $[N_{44414}]Cl$ + H_2O + high Deprot. GA (-3) (CG) + Na	500	9.63	200	1.53	33250	600	0.64	298	1	13^2, 26
(29) $[N_{4444}]Cl$ + H_2O	500	10.05	-	-	22000	-	-	298	1.5	11^2, 23
(30) $[N_{4444}]Cl$ + H_2O + low Prot. GA (CG)	500	10.02	25	0.31	22000	-	-	298	1.5	11^2, 23
(31) $[N_{4444}]Cl$ + H_2O + low Deprot. GA (-1) (CG) + Na	500	10.00	25	0.30	22000	25	0.04	298	2.1	11^2, 23
(32) $[N_{4444}]Cl$ + H_2O + low Deprot. GA (-3) (CG) + Na	500	9.95	25	0.30	22000	75	0.12	298	1.5	11^2, 23
(33) $[N_{4444}]Cl$ + H_2O + high Prot. GA (CG)	500	9.81	200	2.40	22000	-	-	298	1.5	12^2, 23
(34) $[N_{4444}]Cl$ + H_2O + high Deprot. GA (-1) (CG) + Na	500	9.64	200	2.35	22000	200	0.32	298	1.5	12^2, 23
(35) $[N_{4444}]Cl$ + H_2O + high Deprot. GA (-3) (CG) + Na	500	9.30	200	2.22	22000	600	0.92	298	1.5	12^2, 23
Other assays										
(36) $[N_{11114}]Cl$ + H_2O + Deprot. GA (-3) (CG) + Na (water density test)	5000	9.37	2000	2.13	230500	6000	0.89	298	1	25^2, 50

Chapter III: Results and Discussion

1. Biomolecule model validation

In this work, two biomolecules of interest in the pharmaceutical industry, GA and IBP, were approached in aqueous solutions containing ILs with either surfactant, hydrotrope or an intermediate character. These drugs were selected for several reasons:

(i) widely available[148,149] and well-established in the medical industry as probe molecules to study more complex drugs,[150,151] including partition assays;[15,152]
(ii) different protonation states, pH effect in their solubility and partition.

With the aim to rationalize the partition of these APIs in different solutions, novel CG models for GA and IBP were developed to establish a transferrable platform for the development of more efficient drug delivery systems and extraction processes. Due to the previously mentioned characteristics of these molecules, these new descriptions were created for different protonation states, depicting a wide pH range. This extent, often overlooked in MD simulations, is in this case easily applicable to most experimental, real-world applications, especially in drug delivery. Thereby, an analysis of the impact of each mechanism used to increase drug solubility was intended. $[N_{11114}]Cl$ was selected as a surfactant; $[N_{4444}]Cl$ was used as a hydrotrope and $[N_{44414}]Cl$ was chosen as it combines both characteristics, a hydrotrope-like polar head and a long alkyl tail. This allowed the coverage of the main solubilisation mechanisms throughout the continuum ranging from micellar solubilisation to hydrotropy. Thus, it is possible to determine which mechanism enhances the drug solubilisation or conversely, if it is a combination of both that provides the best outcome.

The CG models of the ILs mentioned before were either already developed or based on the mapping procedure and validation by Pérez-Sánchez *et al.*[103] Validation of novel GA and IBP models was accomplished by running both AA and CG MD simulations of these molecules in water. The same concentration and conditions were used for a direct comparison of the RDFs as well as a visual analysis in both scales. Note that the process of obtaining an adequate CG description of a compound is an incremental trial and error procedure by testing different MARTINI bead types until the CG model matches the AA reference. For instance, in the GA at least 11 different attempts were made, counting different bead polarity levels. The same procedure was followed for the IBP CG model.

The RDFs of the carboxylic acid group in the GA, which in turn encompass the protonation/deprotonation state, are presented in **Figure 24**. The RDFs of the aromatic ring and hydroxyls reference groups are available in the **Supporting Information** (**Figures SI:3-4**). The coordination numbers (CNs) of all RDF profiles are displayed in **Table 8.** These values were only provided for groups depicting peaks with a $g(r)$ or $g^*(r)$ values above 1 or 3, respectively. This denotes the solvation of those groups and avoiding the water that has a lower statistical probability of surrounding those groups than the water in the bulk solution. These include the aromatic ring in both molecules and the ethyl benzene group (alkyl chain between carboxylic acid group and aromatic ring) in the IBP. In **Table 8**, CNs at different distances correspond to separate coordination spheres.

Table 8 – CN of the RDF profiles used for validation of CG models of GA and IBP. Number in brackets represent the coordinates of the peak (*r* in the RDF profile in nm) where the CN is estimated.

Reference group	GA				IBP			
	AA		CG		AA		CG	
	Prot.	**Deprot.**	**Prot.**	**Deprot.**	**Prot.**	**Deprot.**	**Prot.**	**Deprot.**
Carboxylic acid bead	0.1 (0.17); 1.33 (0.26)	0.6 (0.17); 3.7 (0.26)	12.4 (0.51)	26.6 (0.51)	0.1 (0.17); 1.7 (0.26)	0.6 (0.17); 3.7 (0.26)	12.3 (0.50)	25.6 (0.50)
Hydroxylic beads (GA) / alkyl tail (IBP)	-	-	10.1 (0.51)	17.4 (0.51)	-	-	-	-

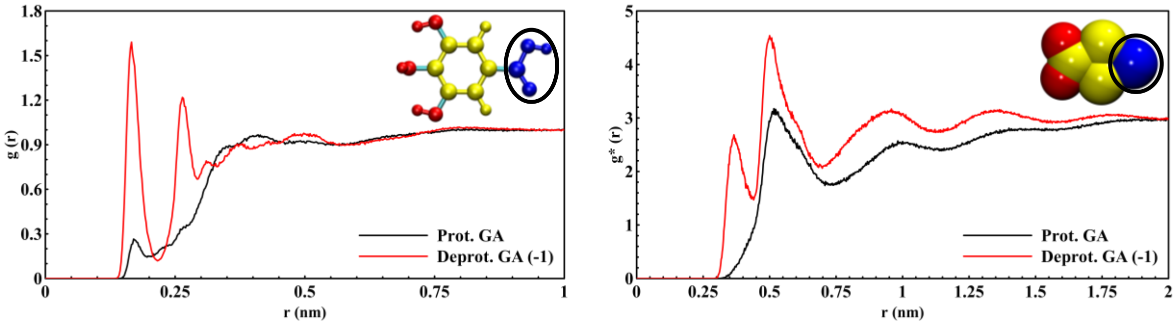


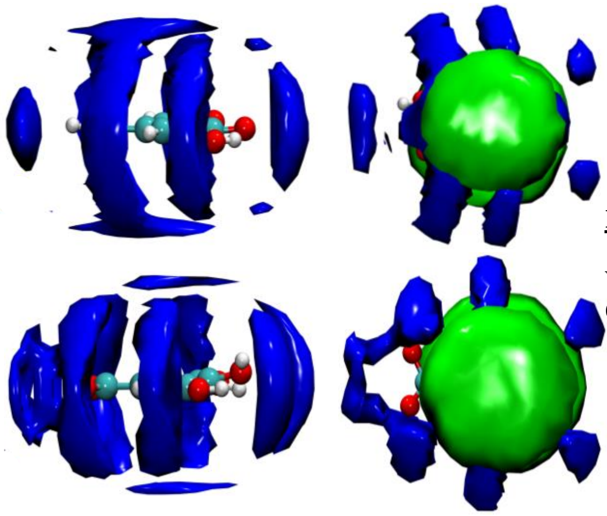
Figure 24 – RDF profile of AA (left) and CG (right) GA with the carboxylic acid group selected as a reference with water, both in protonated (system 1,3) and deprotonated (system 2,4) states, black and red color, respectively.

A qualitative comparison of both RDF profiles demonstrates a significantly higher hydration shell of the deprotonated carboxylic acid group. The sinusoidal behaviour indicates two coordination spheres in the Deprot. GA. The GA AA model, given by the first two peaks of the profile in **Figure 24**, has a CN near 1 and 4, respectively. In the Prot. GA, there is only a first peak in the same location as the Deprot. GA, with statistically no water molecules present, corroborated by the CN. The different peak intensities in the AA profile are expected due to the vdW radii of these atoms. The first coordination sphere has a radius slightly larger than those radii, resulting in a higher probability of water presence. Moreover, the well-defined sinusoidal profile in the AA RDF can be related with the water arrangement towards the COO^-, since hydrogens of water molecules establish H-bonds with the oxygen atoms of this group. In the CG model, the first RDF peak at around $r \approx 0.37$ nm of the Deprot. GA is below the CG bead radius (0.47 nm) thus must not be taken into account. This may interact to a larger extent with charged regions (such as the deprotonated carboxylic acid groups) which have more polar character than the neutral counterpart.[143] Indeed, the MARTINI FF parameters indicate that the energy of interaction is higher for the PW with SQa particles (I – attractive) than between PW and SNda beads (IV – intermediate).[104,143] The values for the coordination spheres surrounding the carboxylic acid group in CG are significantly higher than the ones obtained with AA-MD simulations, not only due to the use of PW but also to the increase of the available area surrounding the CG bead. Three atomistic water molecules are included in one CG PW bead and there is a lack of atomistic resolution of the hydrogen bond interactions. Hence, the comparison of these RDF profiles is mainly qualitative. Bearing in mind this limitation, the CG description is well correlated to the AA model, with more solvation of the COO^- than the COOH, also demonstrated by the CNs. This is expected since the deprotonation of the carboxylic acid increases the number of interactions of this group with water, mainly due to ion-dipole interactions and hydrogen bonding interactions.[153]

The absence of water around the aromatic ring is reflected by the RDFs (**Figure SI:3**), with a difference between both protonation states in the first shoulder of the RDF profile ($r \approx 0.27$ nm for AA and 0.52 nm for CG) around 2-fold higher for the Deprot. GA (-1). These values concord with the apolar character of the aromatic ring.

A different scenario is depicted by the hydroxylic groups, with some differences in the AA and CG RDFs (**Figure SI:4**). While the former OH groups have an apolar character as for the aromatic ring, in the latter, there are three distinguishable peaks in which these regions are polar. Nevertheless, not as much as the carboxylic acid group, namely at $r \approx 0.51$ nm, $r \approx 0.99$ nm and $r \approx 1.40$ nm. Overall, the $g^*(r)$ of the CG RDF is also higher, which is probably related with the lack of atomistic detail inherent in CG models. In the GA CG model the new interaction centers are positioned in the mean distance between each two AA OH groups, as described in the previous chapter. This implies an overlapping with the aromatic ring, being necessary to better resemble the AA structure and avoid its overestimation. It also brings these groups slightly closer to the most polar region of the molecule, which could in turn increase the possibility of interaction with water molecules, apart from their structure comprising two atoms that establish hydrogen bonds.[154] Interestingly, the charge distribution of the AA structure, as depicted in **Figure 13**, confirms this fact and possibly the higher hydration of OH groups when the carboxylic acid group is deprotonated, as two of the hydroxyl groups, which comprise a CG bead, have negative charge. This distribution and RDF data correlate well with another computational study by Cappelli *et al*[154], in which similar $g(r)$ values were reported for the first peaks in the carboxylic acid and hydroxyl RDFs and even the CN for the former.[154]

The SDFs of the AA Prot. GA and Deprot. GA (-1) are presented in **Figure 25.** The SDF images display the water presence around the GA and the π-π stacking arrangement promoted by the GA organic rings.



Isovalue

Water – 47.9
Other GA – 0.4655

Figure 25 - SDF of water (blue) and other GA molecules (green) surrounding a Prot. GA molecule (top) and a Deprot. GA (-1) molecule (bottom). A side view (left) and a top view (right) of the SDF are provided. The GA orientation is set with the carboxylic acid group on the left and the hydroxyls on the right, for all images.

These representations corroborate the conclusions obtained in the RDF analysis, with a clear π-π stacking arrangement of GA molecules.[155] In fact, this interaction is relevant in partition studies because it promotes the aggregation and is present in both protonation states. In summary, the carboxylic acid group generally has more water surrounding it and hydroxyls are also solvated, even though the amount of water is closer to the aromatic ring character than the carboxylic acid group. Comparing both protonation states, the hydration was improved substantially, with coordination spheres between the COO^- group and the ring being pulled towards the carboxylate group, which corroborates the RDF results. Another important aspect is the location of water solvation around the hydroxyl groups, as it is placed in the direction of the positions selected as interaction centers in the CG models and the most charge dense AA hydroxyl regions (**Figure 13**). The SDFs were not thought for CG simulations as the TRAVIS tool was developed bearing in mind the AA level with no recognition of the CG structure. This occurs due to the overlapping of beads innherent in CG models, which influence the distances between them and, consequentely, the ability of TRAVIS to effectively segregate groups of beads. Finally, it is important to note that the GA CG model is quite different from a previously published model.[22] In this regard, only 3 CG beads were used to represent the entire molecule, one for the carboxylic acid region and the two others for the aromatic ring/hydroxyl regions. The former bead had a SP3 polarity while the two more apolar beads were set as SP4.[22] Besides the simplistic CG mapping carried out in this study, furthermore, the polarity differences between each chemical environment present in the molecule are not thoroughly segregated. Thus, our novel model was developed to emphasize the polar differences in the GA molecule and the multiple protonation states present in a realistic solution. In turn, these characteristics are crucial to fully understand the partition of this biomolecule in any solution.

The same AA-to-CG validation procedure was followed for the IBP. The RDF profiles for carboxylic acid with water are presented in **Figure 26** whereas for the other regions are shown in the **Supporting Information** (**Figures SI:5-7**).

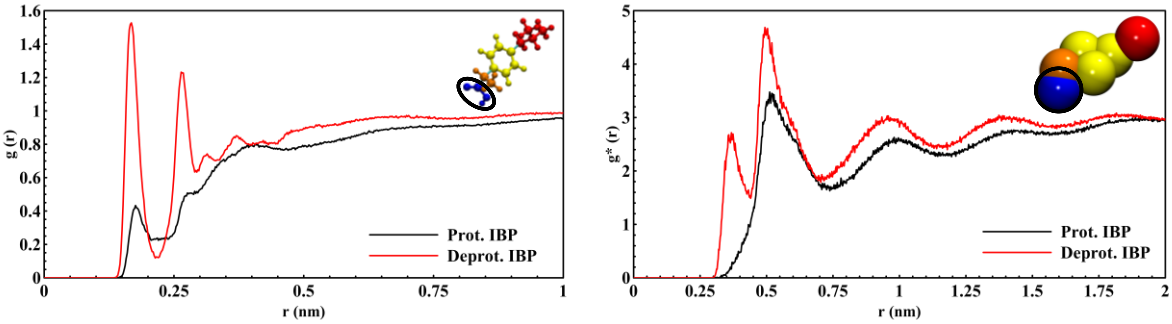


Figure 26 – RDF profile of AA (left) and CG (right) of IBP with the carboxylic acid group selected as a reference with water, both in the protonated (system 5,7) and deprotonated (system 6,8) state, black and red respectively.

These RDFs profiles depict a similar pattern found in GA for the same group since the GA and IBP molecules are similar. The CNs indicate the presence of water in the protonated and the deprotonated IBP states, with a higher intensity for the latter. The first peak in the CG RDF at around 0.37 nm also represents the same effect described in the GA.

The IBP CG model for the carboxylic group reproduced very well the AA behaviour depicted by the RDFs. The rest of the IBP structure is apolar, due to its composition of alkyl groups and the aromatic ring, which resulted in the expected $g(r)$ and $g^*(r)$ values near the reference group, below 1 (3 in the CG) as expected in bulk solution. Nonetheless, slight differences must be noticed in these regions: the alkyl section between the carboxylic acid group and the benzene ring has higher probability of water than the other hydrocarbon groups. This is due to the proximity (and partial overlapping of the CG beads) with the carboxylic acid. The alkyl chain farthest from the carboxylic acid group, which has slightly more water around than the aromatic ring, is possibly caused by the steric hindrance of the π-π stacking between the aromatic rings, reducing the available area for water solvation.

The SDF images of water and π-π stacking presence around IBP molecules are depicted in **Figure 27**.

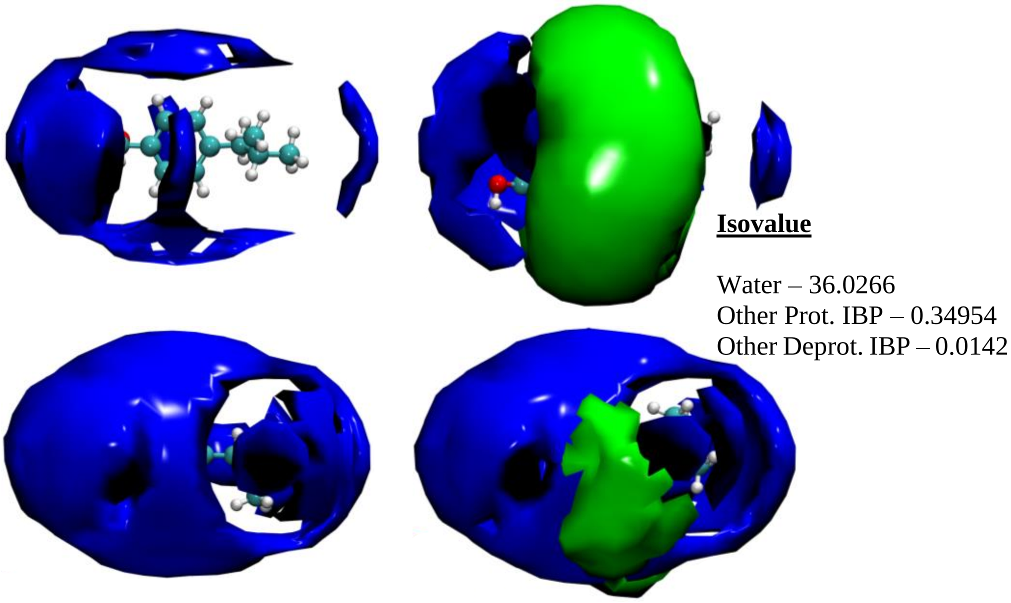


Figure 27 – SDF images of water (blue) and other IBP (green) structures surrounding a reference IBP molecule, both in the protonated (top) and deprotonated states (bottom). The reference structure is placed with the carboxylic acid group pointing towards the left side of the image.

The SDF distributions shown in **Figure 27** indicate that the deprotonated IBP is overall more solvated than its protonated counterpart. However, an increase of other IBP molecules in the protonated version around aromatic ring area is also noticed. This indicates the formation of π-π interactions between rings of different IBP molecules is enhanced in the protonated state. Furthermore, the distribution of the latter totally around the aromatic ring, below and above the ring (as in GA), denotes some degree of aggregation between IBP molecules. Indeed, a visual inspection of the last frame of the simulation (**Figure SI:8**) depicted a clumping between two IBP molecules. In fact, these simulations were conducted either above or near the solubility limit of IBP, which can explain this behaviour. The SDF representation of the IBP arrangement around the organic ring required the selection of different isovalues in the protonated and deprotonated versions for a proper visualization. The RDF profiles differences between Prot. IBP and Deprot. IBP shown in **Figure 26** are also in agreement with another work.[152] In this study, two CG IBP models were compared,[88,156] with a high resemblance between the one developed in this work and the currently published ones being noted. Even though, our model is more realistic in terms of differentiation between the regions that characterize the molecule, particularly with a single center of interaction describing the alkyl region between the aromatic ring and the carboxylic acid group.

The validation of the novel CG models for the GA and IBP molecules was successfully achieved. The RDF profiles were used to validate the behaviour of the CG versions of GA and IBP in the presence of water taking the AA models available in the literature as a reference. The larger specification of the different regions of each molecule in our CG model compared with the previously published ones provides a more realistic picture of these molecules. Furthermore, not only the pH is considered, but also the influence of each characteristic region of the molecule when analysing the partition.

2. ILs phase behavior

The phase behaviour and the effect of the SAIL polar head size of the two SAILs used in this work, $[N_{11114}]Cl$ and $[N_{44414}]Cl$ in aqueous solution, were analysed at the CG level. The effect of salt addition, NaCl, at different concentrations and at 390K, were also tested. The “Presalt”, “Post-salt – low” (0.84 wt.% for $[N_{11114}]Cl$ and 0.64 wt.% for $[N_{44414}]Cl$) and “Post-salt -high” (4.75 wt.% and 4.72 wt.% for $[N_{11114}]Cl$ and $[N_{44414}]Cl$ respectively) were used as labels for the simulations without salt and with low and high concentration of NaCl, accordingly. The final frame of each simulation (**Table 7:** systems 9-14) is shown in **Figure 28** whilst **Table 9** summarises the relevant information discussed below. **Table 9** only shows the CNs for the IL head groups since no water molecules were found in the vicinity of the alkyl-chain tails. The RDF analysis was also compared with the SASA obtained in the simulations.

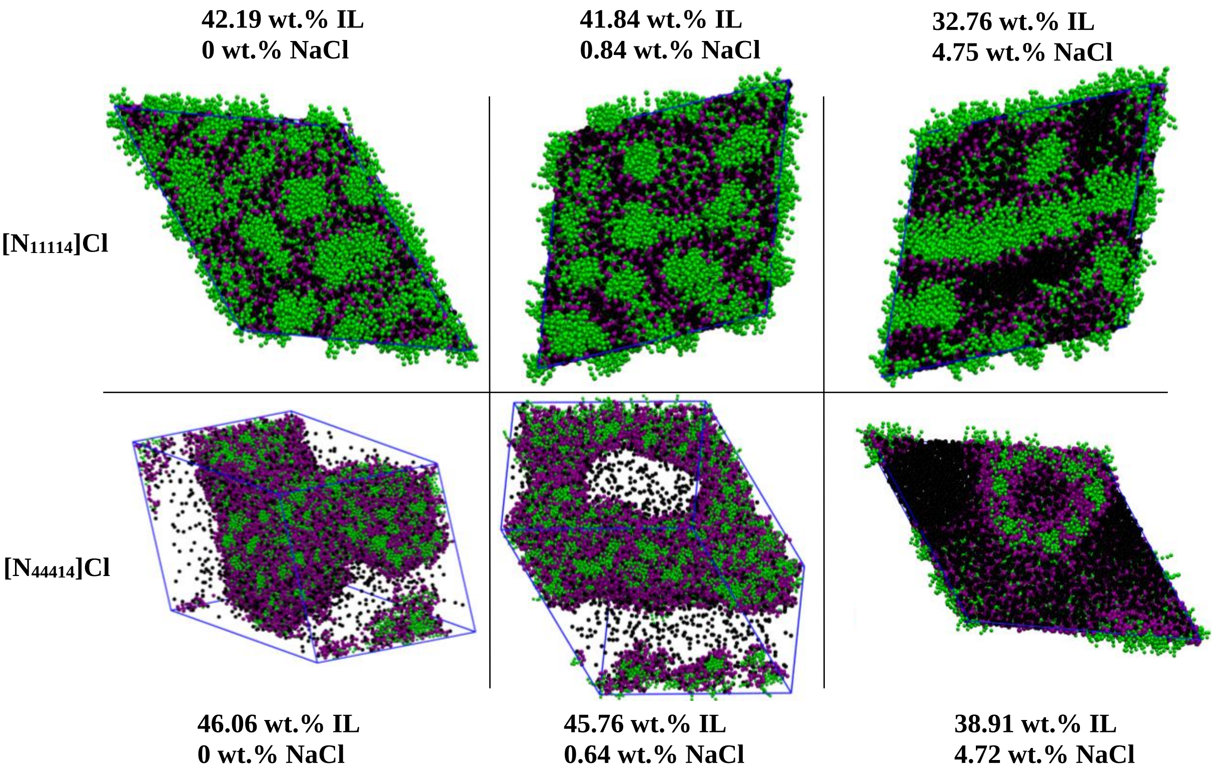

42.19 wt.% IL
0 wt.% NaCl
41.84 wt.% IL
0.84 wt.% NaCl
32.76 wt.% IL
4.75 wt.% NaCl
[N11114]Cl
[N44414]Cl
46.06 wt.% IL
0 wt.% NaCl
45.76 wt.% IL
0.64 wt.% NaCl
38.91 wt.% IL
4.72 wt.% NaCl

Figure 28 – Visualization of the [N_{11114}]Cl and [N_{44414}] mesophases before and after the addition of NaCl at lower and higher concentrations (systems 9-14). Cation IL heads are depicted in purple, with the IL alkyl-chain tail in green and chloride ions (both the IL anion and salt component) in black. Water and sodium molecules were removed for clarity.

Table 9 – Summary of the phase behaviour, CNs and SASA (area) values obtained for systems 9-14. The coordinates at which the CNs were obtained relatively to the reference group are presented in brackets. The SASA is the result of a running average, with the SASA per bead of the group being presented in brackets when applicable.

System	SAIL	[SAIL] (wt.%)	[NaCl] (wt.%)	Phase behaviour (main)	Reference group		
					Polar heads		Alkyl tails
					CN	SASA (nm^2)	SASA (nm^2)
9	[N_{11114}]Cl	42.19	0	Hexagonal	12.9 (0.50); 126.0 (0.95)	4357	9900 (2475)
10		41.84	0.84	Hexagonal	13.0 (0.50); 126.2 (0.95)	4354	9780 (2445)
11		32.76	4.75	Transition to lamellar	10.5 (0.50); 107.8 (0.95)	4342	9240 (2310)

12	[N$_{44414}$]Cl	46.06	0	Bicontinuous	5.3 (0.50); 54.9 (0.96)	12720 (3180)	11200 (2800)
13		45.76	0.64	Bicontinuous	4.3 (0.50); 46.7 (0.96)	12500 (3125)	11280 (2820)
14		38.91	4.72	Phase separated	3.1 (0.50); 38.0 (0.96)	12400 (3100)	11300 (2825)

The [N$_{11114}$]Cl aqueous solution at 40 wt. % depicts an hexagonal phase, in line with the expected phase at this concentration and based on the experimentally reported phase diagram of a similar surfactant.[103] This phase progressively moves towards a lamellar phase after addition of NaCl. This hexagonal-to-lamellar transition through an intermediate lamellar state is further confirmed mainly by a density analysis, shown in **Figure 31** and in agreement with the RDF and SASA data summarized in **Table 9**. Regarding the density profile, the periodic wave-like profile of water and the alkyl chains of [N$_{11114}$]$^+$ indicate the hexagonal arrangement of the surfactant, since peaks of each of these constituents intercalate each other along the simulation box. **Figure 31** also shows the IL head-group (purple) coupled by the chloride anions (black). The system with low NaCl concentration (system 10), has a similar profile, indicative of an intermediate phase between hexagonal and lamellar. This is an overall effect of NaCl addition which promotes a phase transition and the disruption of the initial structure. When the NaCl concentration is increased, this results in a more compact and organized phase closer to a lamellar state, with the presence of larger and non-spherical aggregates. Interestingly, at this salt concentration, the density profile of the chloride anion and the surfactant head begin to decouple, with less surfactant-chloride interaction in favour of sodium-chloride interactions. This is expected as chloride is the only component of the SAIL anion and it has the same polarity as the chloride added with NaCl, thus being able to establish the same interactions.

In the [N$_{44414}$]Cl solutions, all simulations depicted a bicontinuous regime and move towards phase separation with increasing salt concentration. This is shown in the simulations snapshots of systems 12 and 13 (**Figure 28**), where the SAIL occupies the entire simulation box, while it clearly condenses in system 14 (high salt), becoming phase separated. The addition of salt seems to slightly compact and organize the IL phase. Although, not as noticeable as in the case of [N$_{11114}$]Cl due to constant presence of the bicontinuous regime. The [N$_{44414}$]Cl density profiles confirm the progressive phase separation of the system with increasing NaCl concentration, see **Figure SI:9**. Curiously, in the $[N_{44414}]Cl$ system, with a higher concentration of salt, the system was totally phase separated and a water channel was formed in the IL-phase. This channel is composed by IL cation walls filled with water, sodium and chloride ions. For clarity, a detailed picture of this ion channel is depicted in **Figure SI:10**. This liquid electrolyte channel could have interesting applications, as it was pointed out,[157] as well as some works regarding the effect of ILs in ion channels in a biological system[158]. Indeed, the ILs can also serve as templates for mesoporous structures used in catalysis and increase the number of possible reaction sites in the designed material.[159]

The impact of NaCl addition is even more drastic in $[N_{44414}]Cl$ since it originates phase separation at higher concentrations. To further assess the salt addition effect, the RDFs between each region of the IL cation and water were calculated, as well as the corresponding SASA profiles. This data is presented in **Figures 29-30** for the polar heads and in **Figures SI:11-12** for the alkyl tails.

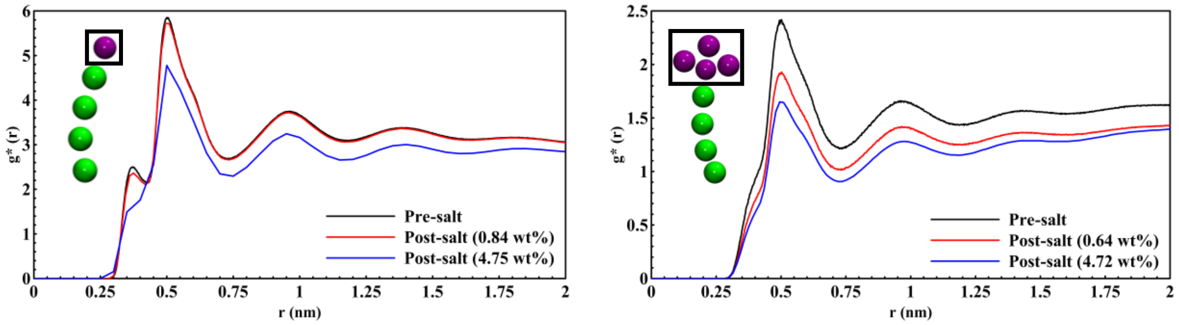


Figure 29 – RDF profile of $[N_{11114}]Cl$ (systems 9-11) and $[N_{44414}]Cl$ (systems 12-14), using the polar head of the cation as a reference and the PW as the selection.

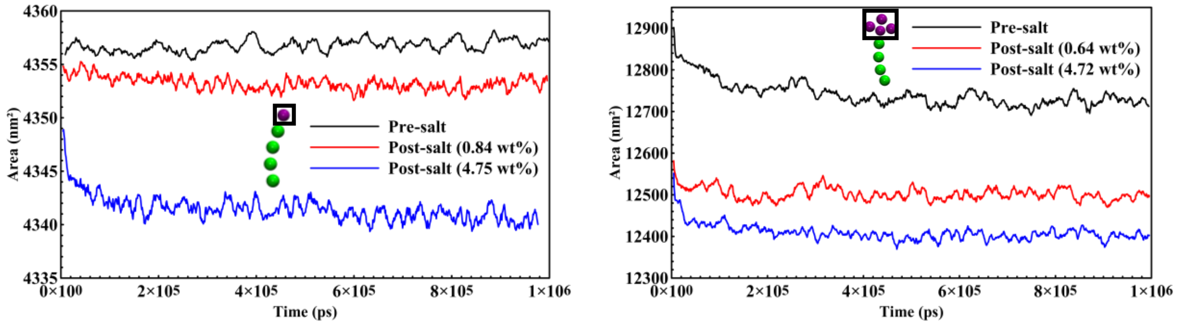


Figure 30 – SASA profiles for the polar head group of the $[N_{11114}]Cl$ (left) and $[N_{44414}]Cl$ (right) cation (systems 9-14).

The RDFs of the polar head and the alkyl tail regions of the IL cation show that the addition of salt induces a 'salting-out' effect, by reducing the number of water molecules able to interact with the components of the IL.[62] Such phenomenon occurs due to the competition of sodium and chloride ions with the surfactant for water solvation. This ionic hydration is set by the density charge, larger in the NaCl ions, with water (PW) molecules typically aligning themselves according to the electrostatic interaction in play. In this case, this is larger for PW-sodium or PW-chloride (I – attractive) than between PW-polar head (II – almost attractive for the central ammonium and VIII – repulsive for the alkyl beads, in $[N_{44414}]^+$).[111] In practice, repulsion between polar heads is reduced, resulting in a more compact organization. All of this is shown in the RDF profiles due to the decrease of the $g^*(r)$ values with the increase of salt concentration in the solution, in both ILs and groups of the cation. In the $[N_{11114}]$Cl, the polar head of the cation still demonstrates the hydration given by the peaks in the RDF profile, with the alkyl tail with no water around due to the hydrophobic core of the surfactant micelle, mainly at a higher NaCl concentration. In the $[N_{44414}]$Cl, even at low NaCl concentration this affects the RDF profiles. In general $g^*(r)$ is kept below the bulk solution plateau, even with polar heads. This is caused by the apolar character of the butyl chains at the surfactant head, which screens the interaction with water. The SASA calculations confirm these results, with the polar head area being reduced as salt concentration increases. This deepens when NaCl concentration is 4.75 wt.% than when comparing the "Pre-salt" and "Post-salt – low" simulations, for $[N_{11114}]$Cl. For $[N_{44414}]$Cl, the SASA for the polar head groups of the IL cation corroborates the RDF profile. Interestingly, SASA analysis of the $[N_{44414}]$Cl alkyl tail indicates that the available area slightly increases with the addition of salt, being similar at both concentrations of NaCl. It would be expected that salt addition, especially when phase separation occurs at a higher NaCl concentration, would result in a decrease of SASA for alkyl tails, as these are more tightly packed and shielded from water, as the RDF profiles suggest. One distinct aspect, however, is the overall lower available area for the alkyl tail of $[N_{11114}]$Cl than that of $[N_{44414}]$Cl, which indicates that this group is protected to a larger extent from water in the former IL. This is related to their phase behaviour, as the former has a more organized hexagonal packing while the latter has a bicontinuous phase with both, the polar head and the alkyl tails, shielded from water to some extent. The polar heads of $[N_{44414}]$Cl are mainly screened by their apolar beads, but also the phase behaviour. This means that not all polar heads are in the outer side of the IL-rich phase with the alkyl tails in the core, as in [N_{11114}]Cl. These are the differences caused by the different polar head size.

In terms of designating these SAILs for drug delivery applications, it is safe to assume that only [N_{11114}]Cl is demonstrated as a viable option, as it maintains a stable hexagonal phase with micelles that could be used as drug vehicles in aqueous solutions. Even at a higher NaCl concentration, the transition to a lamellar phase is more useful than an heterogeneous mesophase, as in [N_{44414}]Cl. However, this bulkier SAIL is more interesting for extraction purposes, as quickly reaches the biphasic state, favourable for the extraction of compounds retained in the IL phase, for example.

From these results, one can conclude that:

(i) The CG mapping adequately captures the differences between the surfactants as well as salt effect;

(ii) Salt addition leads to phase separation of the [N_{44414}]Cl system but not [N_{11114}]Cl under the studied conditions, in line with experiments;

(iii) Significant impact of cationic head polarity on self-aggregation, with a larger polar head resulting in loss of structured phases;

(iv) Phase separation is advantageous or disadvantageous based on the application of interest (drug delivery – bad, extraction and recovery – good).

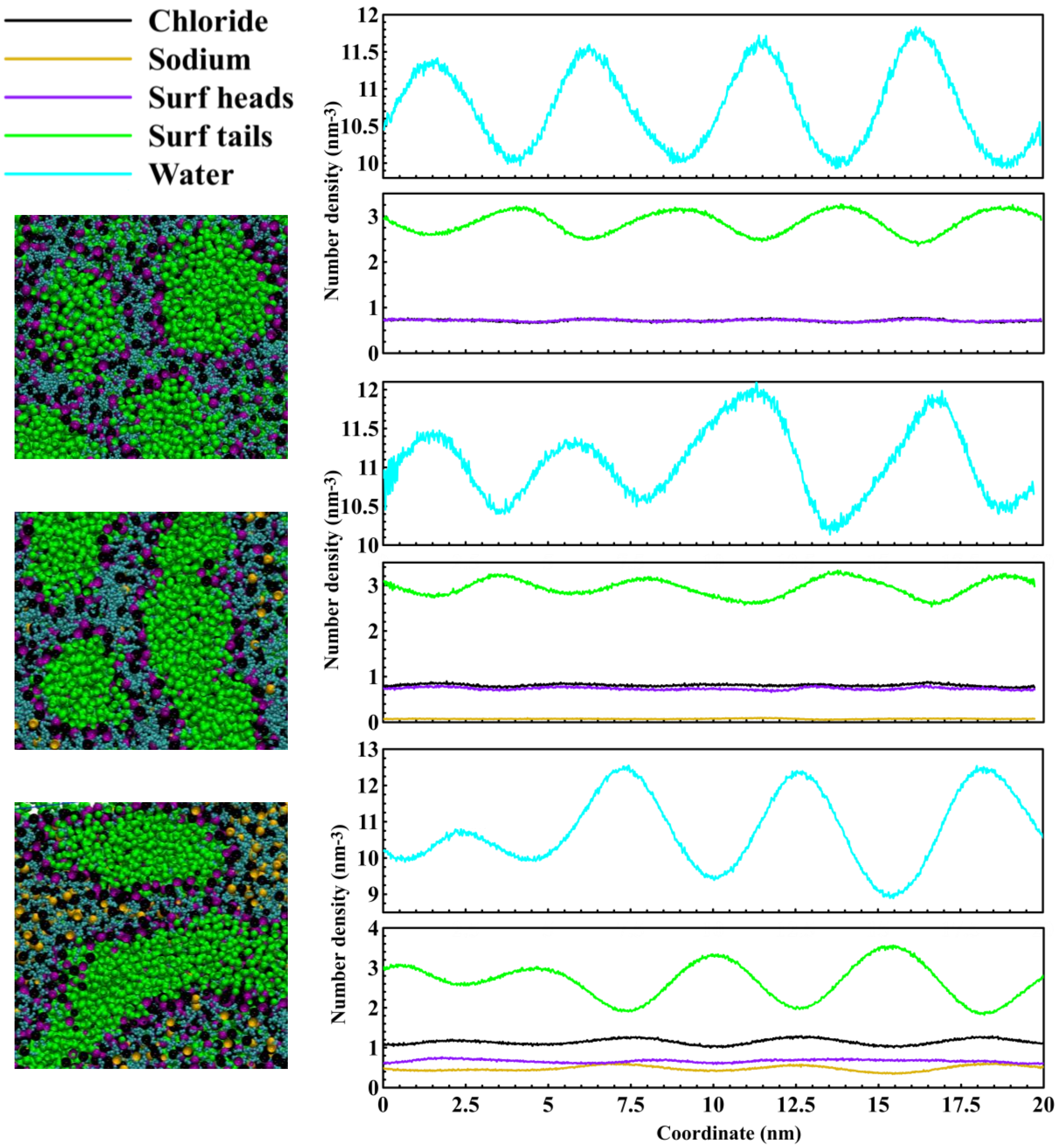


Figure 31 – Density profiles of $[N_{11114}]Cl$ before the addition of salt (top), with NaCl at 0.84 wt.% (center) and at 4.75 wt.% (bottom) (systems 9-11). The SAIL cation heads are represented in purple and tails in green. Chlorides, both from the SAIL anion and NaCl, are depicted in black. Sodium is presented in dark yellow and water in cyan. In the right side, the water density is shown in the top while the rest of the components are depicted in the bottom for clarity. The profiles were generated along the Y axis of the simulation box. Snapshots of a cross-section of the simulation box are presented on the left side of the figure, following the Y direction shown in the density profiles.

3. GA in aqueous IL solutions

The partition of GA was assessed as a function of:

(i) the IL used, *i.e.*, the differences of using adjuvants with differently sized cation polar heads and with micellar solubilization or hydrotropism as the solubilization mechanism;

(ii) the pH, which translates to how the protonation state of GA may induce diverse orientations of the biomolecule towards the IL phase and its influence on phase behaviour in general;

(iii) the GA concentration, in which either the biomolecule concentration is below or above the solubility limit.

Finally, a discussion on the role of these effects will be presented. The systems discussed in this section range from 15 to 35 in **Table 7**, with all simulations running at 298K (T).

Three different ILs were combined with GA in an aqueous solution, $[N_{11114}]Cl$, $[N_{44414}]Cl$ and $[N_{4444}]Cl$. The first two, already analysed before, were compared on the size of their cation polar head region, and both compared with the third IL on the mechanism of action involved in the partition of GA. The IL concentration was set around 10 wt.% for all systems to allow direct comparison between all ILs. The low GA concentration systems (**Table 7**: systems 16-18, 23-25, 30-32) have a GA concentration of ~ 0.20 – 0.30 wt.%, below its solubility limit. This was done to first study the systems in the absence of potential additional effects resulting from elevated GA concentrations. This allowed for the obtention of a spherical-like aggregate suitable for cluster counting analysis. However, most of the analysis will be focused on the high GA concentration (**Table 7**: systems 19-21, 26-28, 33-35) at ~ 1.50 – 2.40 wt.% as the results obtained with the lower ones are similar but more pronounced. The low GA concentration results can be found in the **Supporting Information** whenever not present in the main document and of interest. In certain cases, a “Control” IL was added to the profile in study, corresponding to the IL + water system under identical conditions but devoid of GA (**Table 7**: systems 15, 22 and 29). A summary of the main observations obtain from these tools is presented in **Table 10**.

The effects studied in this work, mentioned in the beginning of this section, will be discussed in an integrated manner, with the latter being organized as follows:

(i) Visual analysis of the final snapshots of each system – phase behaviour and GA partition;
(ii) Density profiles of each system with GA for partition confirmation and orientation verification;
(iii) GA hydration resorting to RDF profiles of the main groups affected by pH change;
(iv) GA influence on the phase behaviour of ILs, using RDF and SASA profiles;
(v) Comparison of each solubilisation mechanism and potential for drug delivery and/or extraction.

a) Gallic acid partition

The final frames of each simulation containing the full simulation box are presented in **Figure 32**. A visual comparison of the different simulations conducted in this work allows for the distinction of all systems in terms of: IL phase behaviour and the relation to the expected solubilisation mechanism; GA partition depending on the protonation state and concentration. Starting with the IL phase behaviour, it is important to note that the two ILs previously analysed without GA addition, $[N_{11114}]$Cl and $[N_{44414}]$Cl, are at lower concentrations than in those assays. Thus, differences in their phase behaviour may be presented. Moreover, the labelling used for the IL cation was the same for all ILs, both the hydrotrope IL ($[N_{4444}]$Cl) as in the two SAILs ($[N_{11114}]$Cl and $[N_{44414}]$Cl). This was done to be consistent with the color code and labelling in the project, even though it is a fact that $[N_{4444}]$Cl is not a surfactant.

From a brief view of both SAILs behaviour in **Figure 32**, it is clear that the trend described in the preceding section is also present after GA addition, with some tuning caused by this biomolecule. The $[N_{11114}]$Cl is generally in a micellar regime, with the increase in GA concentration promoting the enlargement of micelles and their connection. This aggregation seems slightly more prominent as the deprotonation degree of GA increases. When examining systems with $[N_{44414}]$Cl, a more drastic phase transition was denoted, as in the systems where NaCl was added. The increase in GA concentration resulted in larger micelles and phase separation, albeit this enlargement was not so significant. Since phase separation only occurred with pH increase (deprotonation of GA), this process cannot be solely assigned to GA concentration but also to the increased sodium cation concentration. When GA is triply deprotonated (Deprot. GA (-3)), the IL depicted the largest micelles and connection extent, and consequently the largest difference from the two other protonation states.

Table 10 – Summary of the main observations obtained with the ILs + GA systems (systems 15-35). The coordinates at which the CNs were obtained relatively to the reference group are presented in brackets. The SASA presented is the results of a running average, with the SASA per bead of the group being presented in brackets when applicable.

System	IL	[IL] (wt.%)	[GA] (wt.%)	IL phase behaviour	GA partition	Reference group (PW as selection in CN from RDFs)							
						IL cation polar heads		IL cation alkyl tails	GA COOH		GA ring	GA OH	
						CN (0.50 nm)	SASA (nm^2)	SASA (nm^2)	CN (0.51 nm)	SASA (nm^2)	SASA (nm^2)	CN (0.51 nm)	SASA (nm^2)
15	[N_{11114}]Cl	10.02	0	Micellar	-	22.0	546	1180 (295)	-	-	-	-	-
16		9.99	0.29	Micellar	Inside IL micelle	21.3	546	1230 (308)	8.7	30	60 (20)	5.4	51 (26)
17		10.01	0.29	Micellar	Inside IL micelle	21.1	546	1229 (307)	15.7	30	60 (20)	3.6	51 (26)
18		10.00	0.28	Micellar	Outside IL micelle, near surface	20.4	546	1197 (299)	15.1	30	60 (20)	19.7	53 (27)
19		9.79	2.28	Micellar (larger)	At IL micelle core	20.5	546	1294 (324)	8.5	238	473 (158)	4.7	394 (197)
20		9.66	2.24	Micellar (larger, some connected)	Insite IL micelle, near surface	18.0	545	1247 (312)	15.1	242	474 (158)	2.82	386 (193)

21		9.37	2.13	Connected micelles	Outside IL micelle, some near surface	15.8	545	1094 (274)	17.8	242	480 (160)	22.2	425 (213)
22	[N_{44414}]Cl	10.02	0	Micellar	-	12.6	1639 (410)	1178 (295)	-	-	-	-	-
23		10.00	0.20	Large, connected micelles	Inside IL micelles	10.2	1600 (400)	1186 (297)	3.5	30	60 (20)	1.6	51 (26)
24		10.01	0.20	Large, connected micelles	Inside IL micelles, near polar heads	11.0	1621 (406)	1184 (296)	2.4	30	60 (20)	1.0	51 (26)
25		10.08	0.20	Large, more connected micelles	Near IL cation polar heads	9.7	1588 (397)	1170 (293)	2.4	30	60 (20)	6.9	53 (27)
26		9.86	1.60	Large micelles	In IL micelle core	10.6	1637 (409)	1322 (331)	4.4	238	471 (157)	2.0	393 (197)
27		9.6	1.58	One IL-phase separated, and a disperse large IL micelle	In IL-phase and micelle	7.8	1554 (389)	1310 (328)	2.5	242	474 (158)	0.7	387 (194)
28		9.63	1.53	Phase separated	Disperse in solution and outside the IL polar head surface	6.4	1531 (383)	1210 (303)	8.3	242	480 (160)	14.9	426 (213)
29	[N_{4444}]Cl	10.05	0	Monophasic solution	-	13.4	547	2281 (570)	-	-	-	-	-

30		10.02	0.31	Monophasic solution	Mainly disperse in solution	13.4	547	2271 (568)	10.9	30	58 (19)	8.0	49 (25)
31		10.00	0.30	Monophasic, small clustering of IL	Gathered in IL clusters	12.7	547	2266 (567)	9.8	30	58 (19)	4.9	47 (24)
32		9.95	0.30	Monophasic, small clustering of IL	Disperse in solution	13.1	547	2264 (566)	17.8	30	60 (20)	24.3	53 (27)
33		9.81	2.40	Large IL cluster, with some disperse molecules	In IL cluster	11.0	547	2232 (558)	5.4	230	418 (139)	2.8	358 (179)
34		9.64	2.35	More compact and dense IL cluster, less disperse molecules	In IL cluster	8.4	547	2073 (518)	4.7	242	456 (152)	1.8	355 (178)
35		9.30	2.22	Nearly phase separated, largest IL cluster	Disperse in solution, some molecules in IL cluster	8.6	547	1963 (491)	14.0	242	480 (160)	20.0	425 (213)

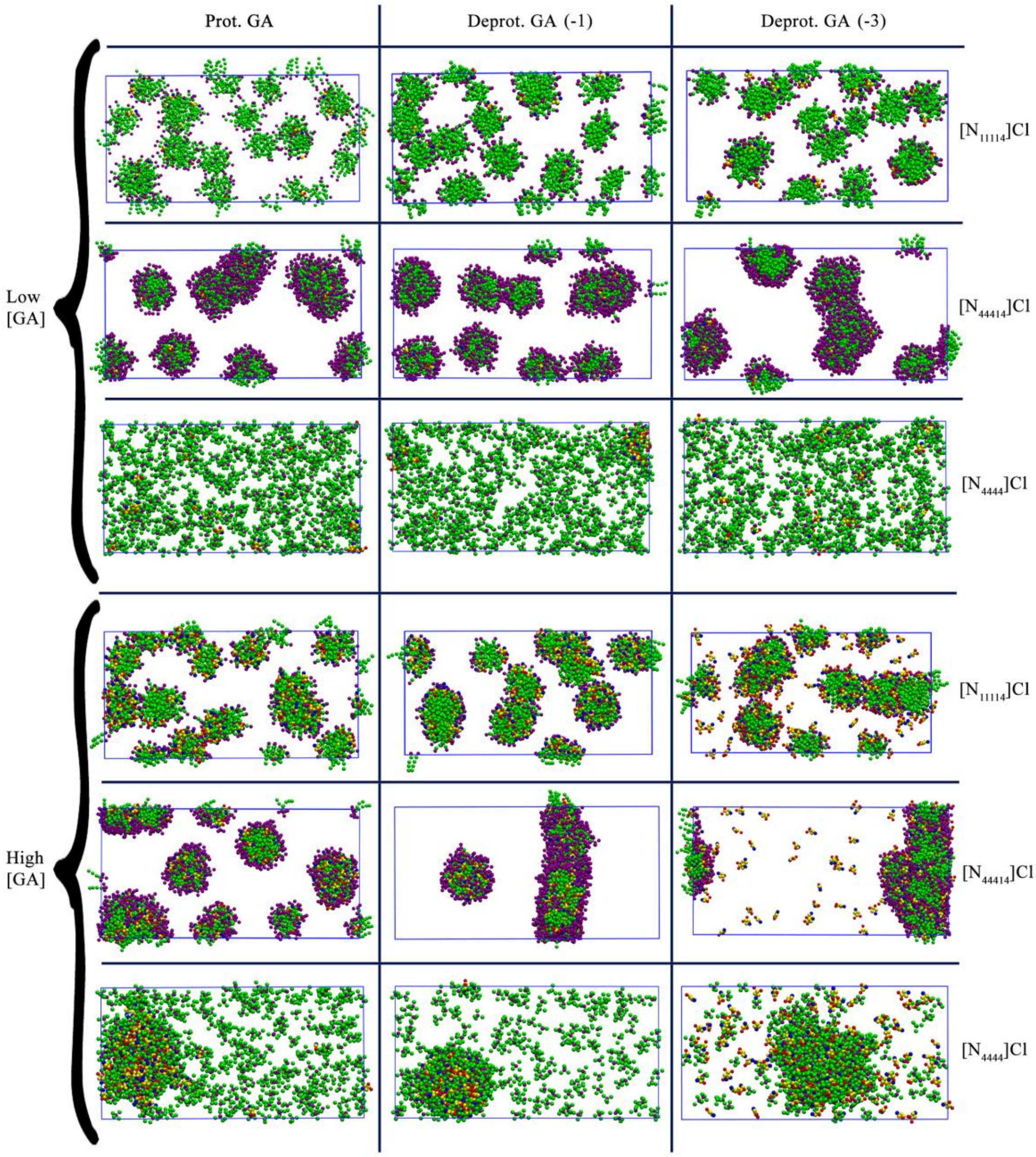

Prot. GA
Deprot. GA (-1)
Deprot. GA (-3)
Low [GA]
High [GA]
[N11114]Cl
[N44414]Cl
[N4444]Cl
[N11114]Cl
[N44414]Cl
[N4444]Cl

Figure 32 - Final snapshots of the production runs comprising [N_{11114}]Cl, [N_{44414}]Cl and [N_{4444}]Cl with GA at low and high concentrations and various protonation states (systems 16-21, 23-28 and 30-35). IL cation head groups are depicted in purple and alkyl chains in green, while the carboxylic acid group of GA is presented in blue, the aromatic ring in yellow and the hydroxyls in red. Both water, chloride and sodium counter ions were removed for clarity.

The hydrotrope IL, [N_{4444}]Cl, displayed a behaviour that was could not be compared to the results of the previous section, as it was not tested for NaCl addition. Still, the hydrotropic action was detected, especially in the systems with higher GA concentration (**Table 7**: systems 33-35). While this IL is dispersed in solution at lower GA concentrations (**Table 7**:

systems 30-32), large hydrotrope clusters are formed when the GA concentration is increased. These clusters move towards almost complete phase separation, present in **Table 7**: system 35, as the increasing GA deprotonation degree is associated to a larger and more compact IL cluster. When comparing the three ILs, it is visually noticed that their phase behaviour is affected by GA, in a pH-driven manner, with the GA concentration increase emphasising this effect.

The GA partition is also visually distinguishable in these systems and a general trend is found in all the ILs:

- when fully protonated (Prot. GA), GA is inside the micelle core or hydrotrope cluster, and closer to its core;
- when singly deprotonated (Deprot. GA (-1)), it is still partitioned inside the micelle or hydrotrope cluster, but seems to be closer to the polar heads of the cation;
- when triply deprotonated (Deprot. GA (-3)), it is either at the outer side of the micelle and hydrotrope cluster or disperse in solution at higher concentrations. In $[N_{4444}]Cl$, some molecules are inside the cluster.

The partition behaviour of GA is expected to be more affected by the pH than the concentration. The ILs are ionic molecules and the interactions between both compounds will determine their behaviour and distribution. However, in the hydrotropic systems, the concentration is also an important factor as the GA is dispersed in solution at lower concentrations. The similar partition verified in the two other ILs was only visually confirmed at higher GA concentrations.

A possible cause for the morphing of the IL phase behaviour and GA partition due to the protonation state of the latter is the exchange of the chloride anion of the IL with GA. As the carboxylic acid group is deprotonated, electrostatic interactions with the IL polar heads of its cation may be established. The polarity of these bead types, however, is not as promoting of this interaction as when this group is protonated. In terms of the CG MARTINI energy of interactions, the protonated COOH group affinity to the charged ammonium center in the polar head is set as almost attractive, while for the COO^- version it is almost repulsive. The hydroxyl bead interaction is intermediate in both Prot. GA and Deprot. GA (-1), and almost repulsive in Deprot. GA (-3). Thus, it is interesting on how there is a protonated GA where it is more attracted to the IL polar heads, and draws farther from the latter as GA deprotonation degree increases. Such a behaviour is demonstrated in the GA partition. Still, the visual similarities between the systems with Prot. GA and Deprot. GA (-1), as well as the confirmation of these observations and GA orientation in the IL micelles or phase require more tools.

b) Density profile analysis

The density profiles were obtained with the cluster counting algorithm for spherical-like aggregates and the *gmx density* tool included in GROMACS for non-spherical systems. Since the systems with higher GA concentration allow a good visual inspection, only the density profiles of these systems (19-21, 26-28, 33-35) are showed. Similar but less statistically relevant results were obtained for low GA concentration systems (16-18, 23-25, 30-32) and are available in the **Supporting Information** (**Figures SI: 13-15**). The density profile for the [N_{11114}]Cl systems (19-21) is depicted in **Figure 33**, while for [N_{44414}]Cl (26-28) and [N_{4444}]Cl (33-35) are presented in **Figures 34** and **35**, respectively.

It must be noticed that the cluster counting code used in this work only allows to differenciate two regions in the molecule, thus the organic ring and hydroxyl beads of GA were grouped in a single selection and the carboxylic for the other group. Nevertheless, this is enough to confirm the GA partition towards the SAILs micelles and also their orientation. Furthermore, the cluster counting code highlighted the underestimation of the density of water inherent in many CG models. This is a consequence of the MARTINI CG bead packing since the CG beads arranged in the simulation box keep some “empty” spaces in between. Thus, the cluster counting code, initially developed for AA models where the molecules are more realistically displayed, displays an unrealistic low water density below 1 $g\cdot cm^{-3}$. As a further test, a double-size system was built and tested to ensure that this is not an effect produced by the system size (**Table 7**: system 36).

The RDF profiles using the Ct3 bead (the bead before the last bead of the alkyl tail) of the SAIL as the reference and the three groups of the GA as the selection were obtained and depicted in **Figure SI:16**.

The RDFs of the [N_{11114}]Cl system show that the micellar arrangement was kept throughout the full pH range. In all states, the RDFs of the different GA groups, carboxylic, aromatic ring and hydroxyl moities respect the alkyl chains (Ct3) have their peaks well inside the micelle surface, denoted as the green dashed line in **Figure SI:16**, when protonated or singly deprotonated, and out of the micelle surface when triply deprotonated. A distinction between Prot. GA and Deprot. GA (-1) is noted as the peaks of organic ring/hydroxyls and the carboxylic acid are more decoupled from each other in the singly deprotonated version. This suggests that while Prot. GA displays a spurius orientation in relation to the micelle surface, the Deprot. GA (-1) was essentially arranged perpedicular to the surface, with the GA carboxylic group inserted between the surfactant cationic head and the aromatic ring lying deeper in the micelle core. Moreover, in the protonated version, the GA molecules are closer to the micelle core and from the micelle surface than in the singly deprotonated version. The Deprot. GA (-3) was arranged out of the micelle surface, but with an orientation different from the other protonation states. The organic ring/tail peak appears slightly after the carboxylic bead, indicating that the GA molecule is totally outside the SAIL micelle. The similar distance from the micelle COM for the GA carboxylic head and aromatic/hydroxyl groups further suggests that the GA molecules lie parallel to this micelle surface, as observed in micelle snapshot in **Figure 32**.

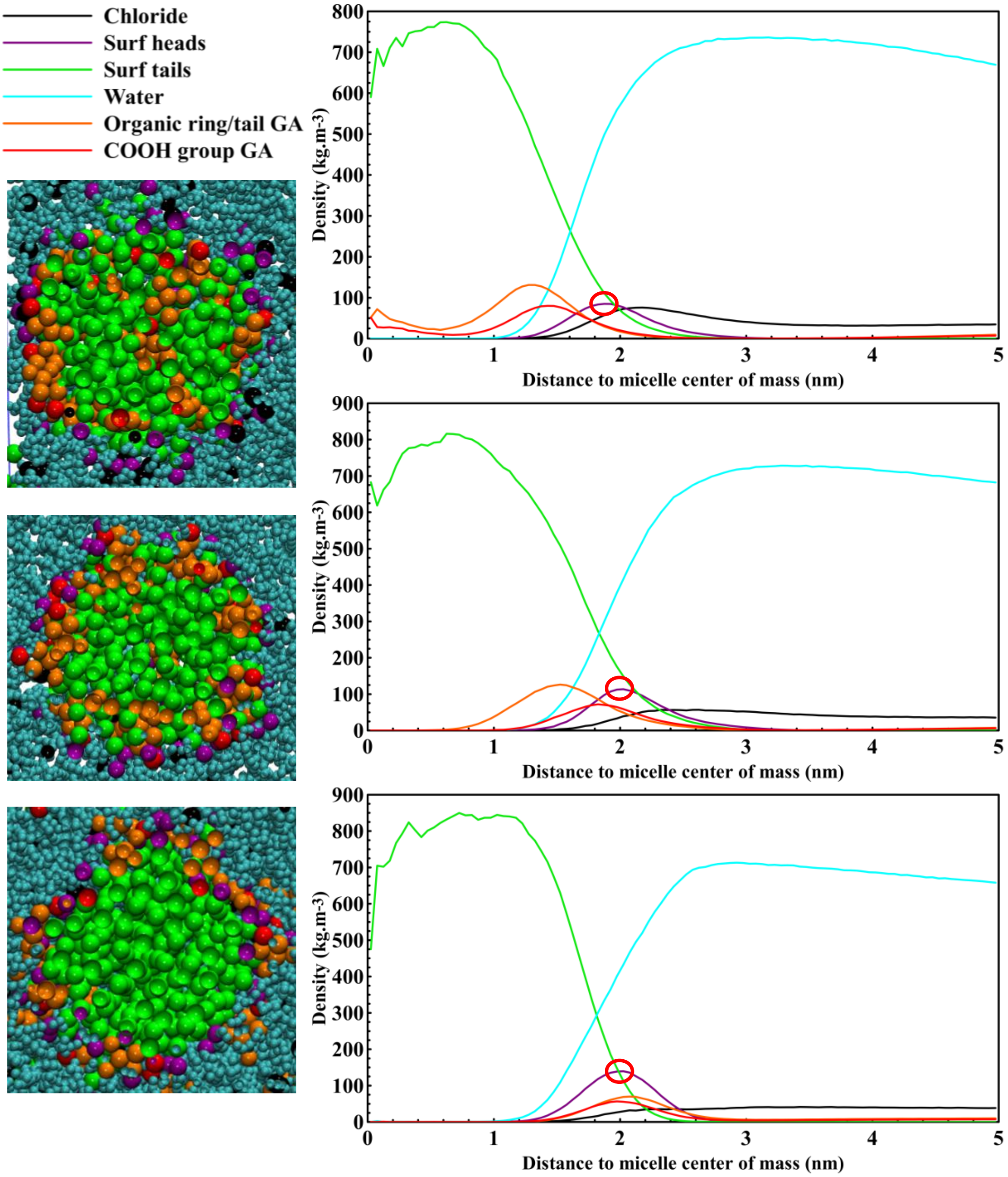


Figure 33 - Density profiles obtained with the cluster counting code for 19-21 systems, containing GA and $[N_{11114}]Cl$. The COM of the IL micelle was set as the reference center from with the density of each component was obtained. The snapshot on the left side corresponds to a representative micelle in each state (Prot. GA in the top, Deprot. GA (-1) in the middle and Deprot. GA (-3) in the bottom). The micelle surface is highlighted by a red circle.

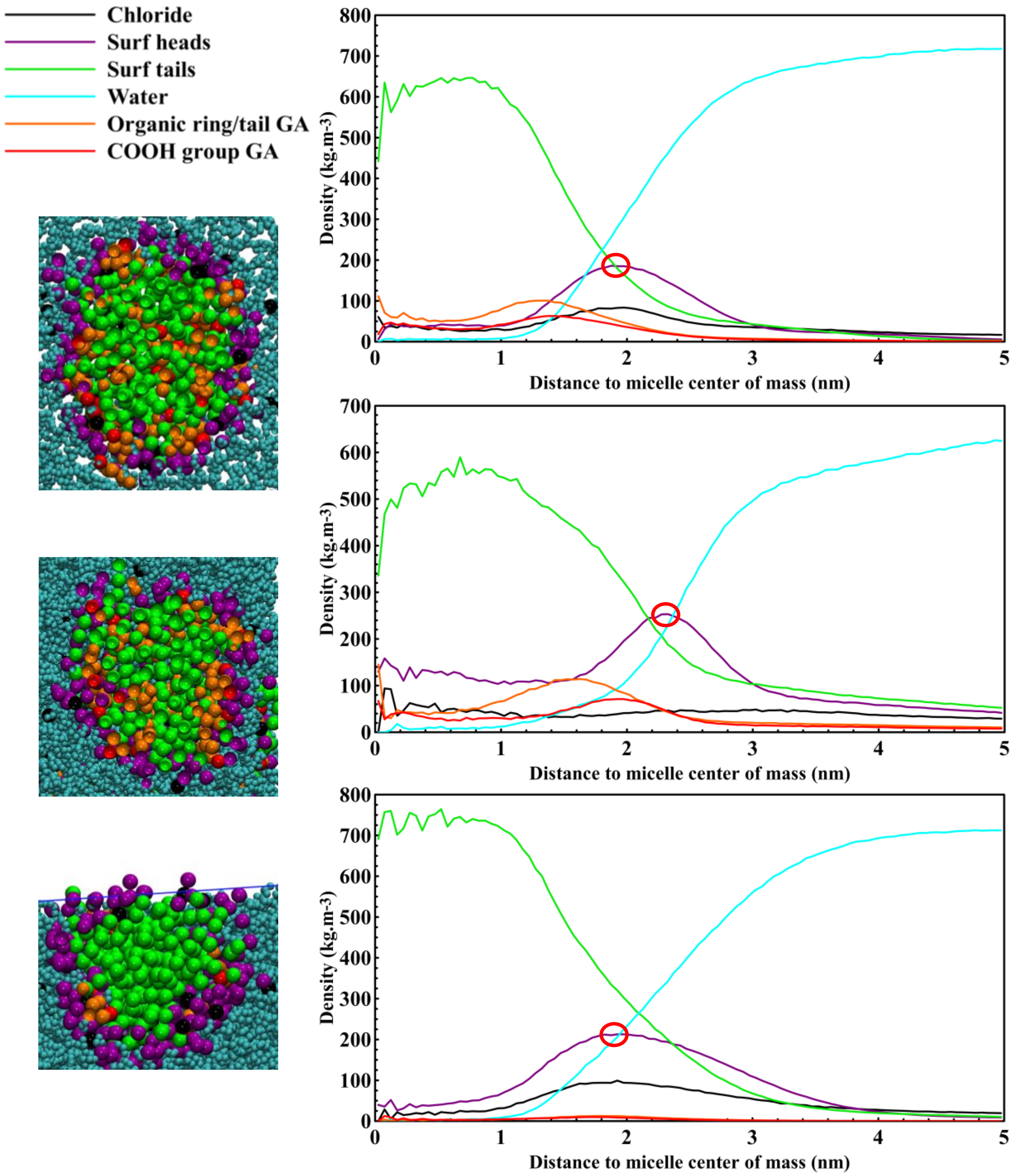


Figure 34 - Density profiles obtained with the cluster counting code for 26-27 systems, containing GA and [N$_{44414}$]Cl. The Deprot. GA (-3) at low concentration (system 25) was used since it maintains the micellar phase and allows the use of the cluster counting algorithm The snapshot on the left side corresponding to a representative micelle in each state or the full simulation box (Prot. GA in the top, Deprot. GA (-1) in the middle and Deprot. GA (-3) in the bottom). The micelle surface is highlighted by a red circle.

The [N$_{44414}$]Cl systems depicts some similarities with the [N$_{11114}$]Cl system, particularly in the protonated and singly deprotonated systems. In both, the partition and orientation of the GA molecule in the SAIL micelle is similar. However, the [N_{44414}]Cl differs in the sense that phase transites towards the phase separation, also demonstrated in the density profiles by the increased polar head density near the micelle COM. This is noted particularly in the singly deprotonated system, which is already close to the phase separation, as demonstrated in the final snapshots of the simulation runs. The full phase separated system is reached in the high concentration Deprot. GA (-3), which prevents the use of the cluster counting code for this reason. Thus, in **Figure 34**, the corresponding low GA concentration density profile is shown. This density profile indicates the presence of GA inside the SAIL micelle, conversely to the [N_{11114}]Cl. Such distinction is caused by the butyl chains in the polar head of [N_{44414}]$^+$ which shield the charged ammonium center from the negative hydroxyl groups in the GA molecule. Thus, the negative groups of GA (including the carboxylic acid) are not able to closely interact with the polar head of the IL, as in the systems containing [N_{11114}]Cl.

The similarities in the partition of Prot. GA and Deprot. GA (-1) in both SAILs are related to the carboxylic bead charge when deprotonated, which is attracted to the ammonium center of the IL polar head. In addition, the retention of this group at the micelle surface also contributes to the alignment of the rest of the GA molecule with the alkyl chains of the surfactant.

Finally, the [N_{4444}]Cl density profiles with GA were also obtained. In this case, the *gmx density* tool of GROMACS was used since it has an hydrotrope solubilisation mechanism that does not resort to a micellar phase. At low GA concentrations, these systems were monophasic, with an IL and GA dispersion throughout the full simulation box, except in the Deprot. GA (-1) system (31) where small GA clusters are observed, as seen in **Figure SI:15**.

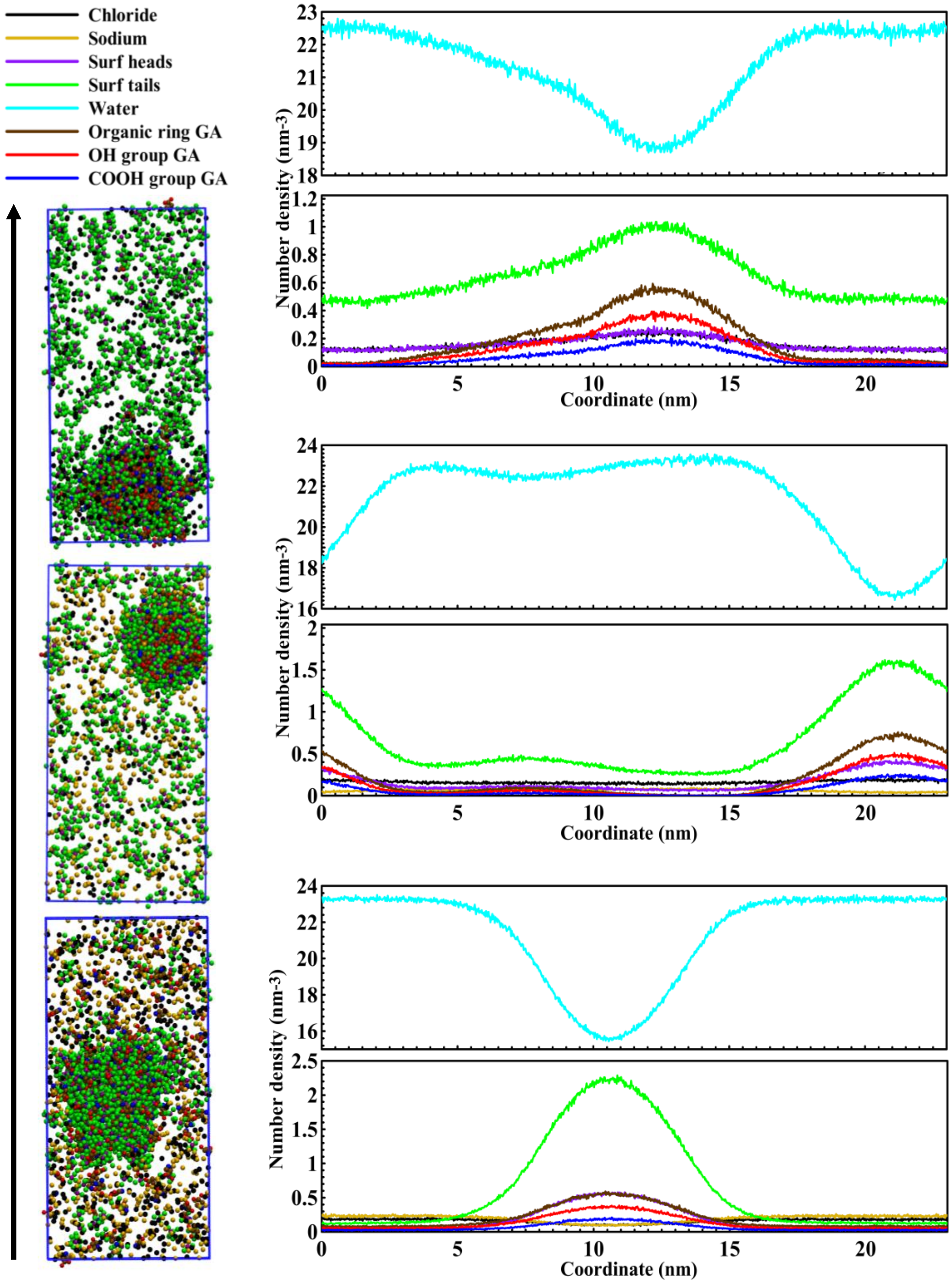


Figure 35 - Density profiles obtained with the *gmx density* gromacs tool for systems 33-35 containing GA and [N$_{4444}$]Cl. The density profile of each component was obtained along the simulation box in the *z* axis direction. The snapshot on the left side corresponds to the simulation box (Prot. GA in the top, Deprot. GA (-1) in the middle and Deprot. GA (-3) in the bottom). Water was removed in these for clarity.

At higher concentrations clustering of the IL arises in all cases, with more progress towards phase separation as the deprotonation degree of GA increases. This is confirmed in the density profiles by a noticeable alkyl chain peak. The [N_{4444}]Cl density is inversely correlated to the water density. It is clear that the GA is placed in the IL cluster in all of cases since the density peaks of each GA group appear in the same position as the alkyl tail. In terms of GA molecule orientation, it is not possible, with these density profiles, to confirm any specific orientation of this molecule towards the IL.

All of these orientation and partition conclusions are confirmed by the close-up snapshot next to the density profile and the RDFs depicting the prevalence of each GA group near the micelle core (**Figure SI:16**). A summary containing the CNs obtained from these RDFs and the aggregation number (AN) of SAIL micelles in all aplicable systems is organized in **Table 11**.

The RDF profiles relating the micelle core and each GA group (**Figure SI:16**) provide a qualitative confirmation of the density profiles. An example is presented in **Figure 36** for the [N_{44414}]Cl system and the hydroxyl beads of the GA (**Table 7**: systems 23-25). In each case, the position of the GA group matches the partition and orientation described in the previous section. For example, in the [N_{11114}]Cl system, the hydroxyl beads are closer to the micelle core when in the Deprot. GA (-1) compared with the Prot. GA, with this group and the aromatic ring arranged out of the micelle surface. In case of the [N_{44414}]Cl system, the same trend is observed. In terms of the micelle surface, a clear correlation between the CN and the RDF and density profiles is denoted. With the Prot. GA and the Deprot. GA (-1) in both SAILs, there is a higher presence of all groups near the micelle surface compared with

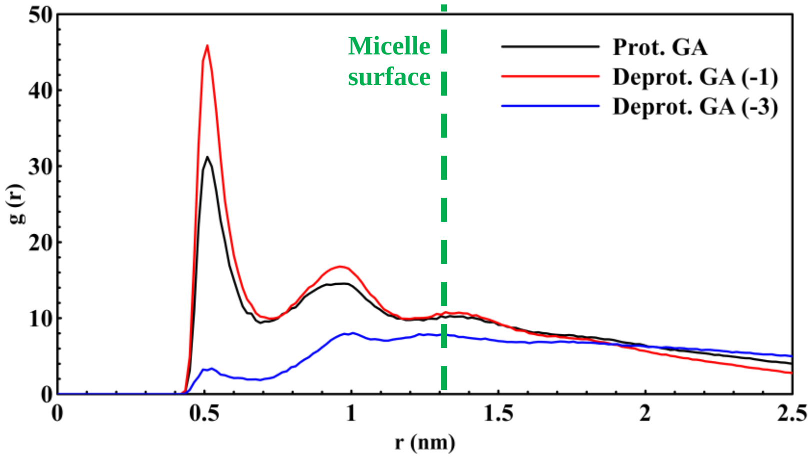


Figure 36 – RDF profile using the Ct3 bead of [N_{44414}]Cl as the reference and the hydroxyl beads of GA as the selection (systems 23-25).

Deprot. GA (-3). This is relevant for the ring and hydroxyls, which have their CNs drastically mitigated when Deprot. GA (-3) is used, as these are placed out of the IL micelle.

c) Gallic acid hydration

The hydration of GA was also assessed, with similar RDFs designed for each GA group. Since the aromatic ring was verified as not hydrated, as expected by the model validation and the polarity, the profiles for this group are displayed in the **Supporting Information (Figure SI:17)**, as well as the low GA concentration results (**Figures SI:18-19**). The RDFs regarding the carboxylic and hydroxyl moieties at high GA concentrations are presented in **Figure 37**. The SASA profiles complementing this information are depicted in **Figures SI:21-23.**

Endorsing the GA partition conclusions mentioned before, these RDF profiles indicate the hydration of the carboxylic region when the GA is deprotonated. It is either at the surface of the IL-phase (in micelles) for the Deprot. GA (-1), or out of the micelle surface for the Deprot. GA (-3). This is further envisaged for the hydroxyl moiety as it is either shielded from water inside IL micelles or by hydrotrope clusters (for Deprot. GA (-1)) or outside these and available for solvation (for Deprot. GA (-3)). This logic is followed in both ILs, although for the [N_{44414}]Cl, the $g^*(r)$ values of both IL cation moieties are below the 3 plateau, the reference threshold below which there is no significant hydration in the CG description. This is a consequence of the IL phase behaviour, which in these cases, can be either monophasic bicontinuous-like or phase separated, with the GA screened from water. A similar phenomenon occurs for the [N_{4444}]Cl system, indicating that this is a result of the hydrotrope mechanism. However, the hydroxyl groups of the Deprot. GA (-3) are hydrated, since in the fully hydrotropic system, the GA is mainly solvated conversely to the [N_{44414}]Cl behaviour. Althought the aromatic ring is not statistically solvated, these results also demonstrated the differences between the GA protonation states. This region depicted much higher $g^*(r)$ values in the Deprot. GA (-3), as these are the only systems where the ring is on the outer side of the IL phase, with the surface exposed to the water.

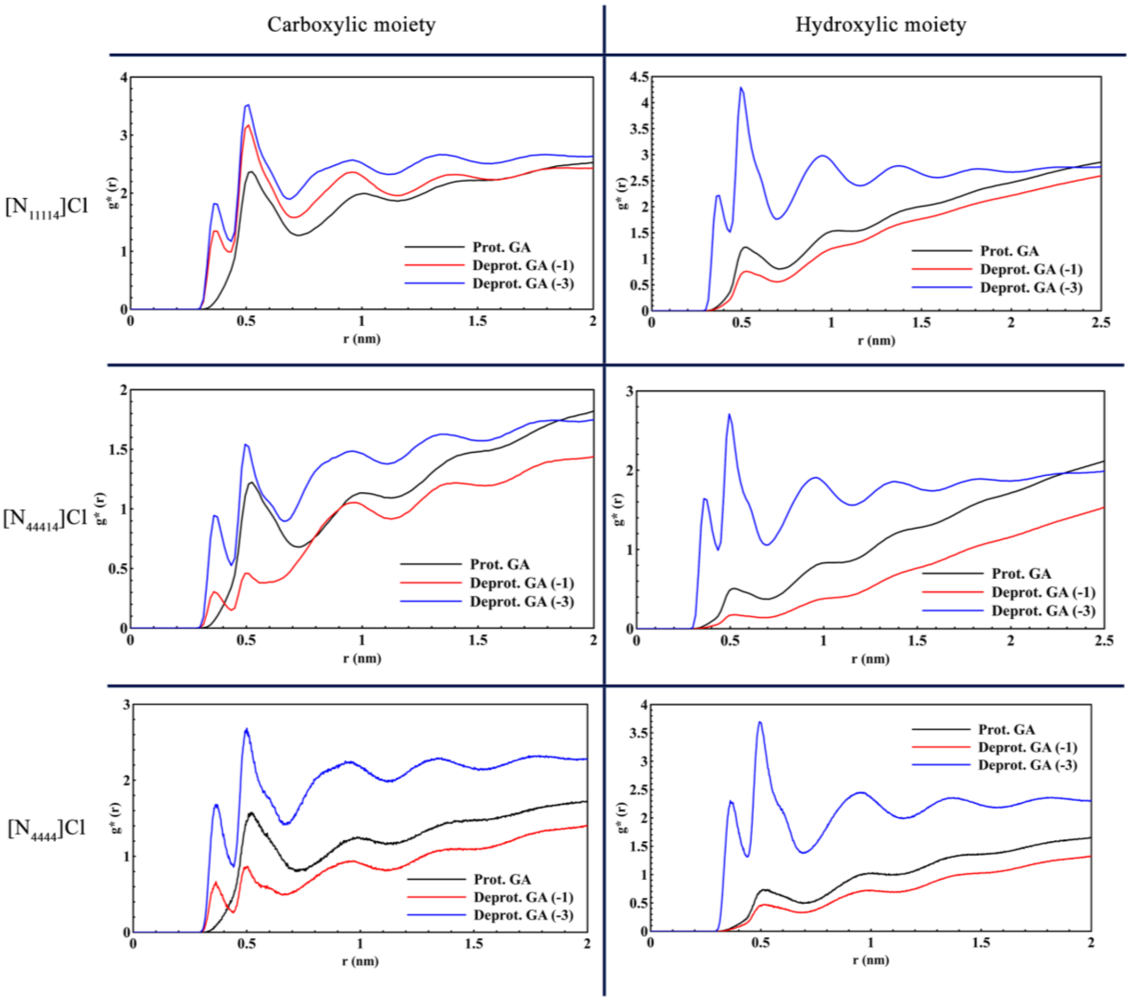


Figure 37 – RDF profiles using the carboxylic (left) and hydroxylic (right) moieties of the GA as the reference and the PW as the selection (systems 19-21, 26-28, 33-35).

To conclude the study of the partition of GA, it is also possible to validate these results with some previous works using similar systems (GA mixed with similar ILs). For example, Spigno and team[160] were able to prove the distribution of GA in a CTAB (similar to [N$_{11114}$]Cl) phase and also that the electrostatic interactions were the main driving forces in this partition.[160] Their work also reinforces the pH importance in GA extraction. Neves and colleagues[161] were able to indicate that more hydrophobic IL cations are related to the partition of GA in the hydrophobic polymer-IL rich phase in an ABS.[161] This may be applied to the comparison between both long-chain ILs used in this work, since the [N$_{44414}$]Cl system displays a slightly more hydrophobic cation and thus could explain the noticeable partition of GA in this IL phase. The hydrotrope behaviour of [N$_{4444}$]Cl stems from its chaotrope character according to the Hofmeister series,[162] as it is less susceptible to water interaction, in turn making GA more likely to partition.

d) Influence of GA on phase transition – micelle swelling

The AN presented in **Table 11**, demonstrate the effect of the concentration of GA in the IL phase behaviour. As it was mentioned before, the cluster counting can only estimate the density profile in systems that self-assemble in separated spherical aggregates. Thus, this code canot be used in systems such as aqueous solutions of $[N_{4444}]Cl$ or systems that are phase separated such as the 27 and 28 systems. The results shown in **Table 11** indicated that an increase in the GA concentration yields to an increase in the ANs attributed to GA saturation of the micelle core. However, for a given GA concentration, the AN varies accordingly with the pH change. The partition of GA deeper in the micelle core in the protonated state causes a decrease in the AN compared to the deprotonated GA. This may occur due to the disruption of the stable micellar structure induced this biomolecule, with the GA interacting with these alkyl chains and competing with other SAIL cations.

Table 11 – Aggregation number (AN) of the SAIL micelles and the CN of each group of GA respect the SAIL micelle core used as a reference (Ct3 bead).

System	IL	[IL] (wt.%)	[GA] (wt.%)	AN	GA net charge	CN at micelle surface – 1.31 nm)		
						GA COOH	GA Ring	GA OH
						CN	CN	CN
15	$[N_{11114}]Cl$	10.02	0	38	-	-	-	-
16		9.99	0.29	28	0	0.4	1.4	1.0
17		10.01	0.29	29	-1	0.3	1.3	1.1
18		10.00	0.28	33	-3	0.2	0.7	0.4
19		9.79	2.28	45	0	2.6	10.1	7.3
20		9.66	2.24	56	-1	1.8	9.7	8.5
21		9.37	2.13	56	-3	1.4	3.7	2.0
22	$[N_{44414}]Cl$	10.02	0	45	-	-	-	-
23		10.00	0.20	100	0	0.5	1.7	1.2
24		10.01	0.20	63	-1	0.4	1.7	1.3

25		10.08	0.20	125	-3	0.4	1.1	0.6
26		9.86	1.60	71	0	3.1	10.9	7.5
27		9.6	1.58	-	-1	2.8	11.4	9.1
28		9.63	1.53	-	-3	2.2	5.3	3.1

The electrostatic interactions influence the extent of aggregation. To better understand the GA interaction with the IL as a function of pH, the RDF between each GA group as the reference and the IL cation polar heads as the selection were obtained and are represented in **Figure 38**. Similarly to the previous analysis, the high GA concentration systems are deeply discussed whereas the low GA concentration are available in **Figure SI:20**.

In all of the IL studied, the carboxylic moiety is closer to the IL cation polar head when the GA is Deprot. GA (-1) followed by the Deprot. GA (-3). The hydroxyl region shows a similar trend in all of the ILs, with Deprot. GA (-3) being next to the polar heads and the two other versions farther from it. This corroborates that the Deprot. GA (-1) is inside the SAILs micelles with the carboxyl moiety near their surface and the hydroxyl group in the core, with this molecule aligned with the micelle building blocks. The protonated version is buried deep in the SAILs micelle cores with both groups being farther from the surface.

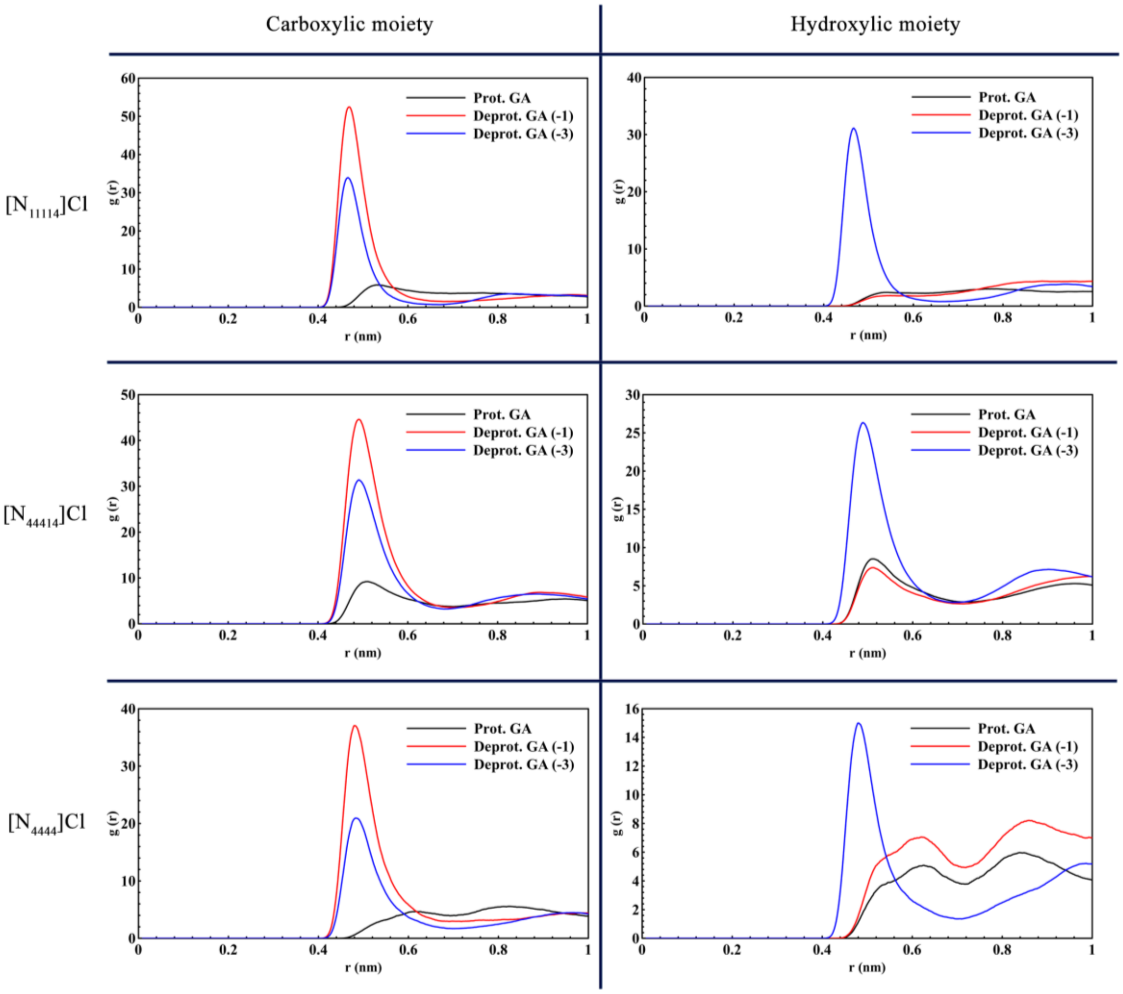


Figure 38 – RDF profiles using either the carboxylic or hydroxylic moieties of the GA as the reference and the polar heads of the IL cation as a selection (systems 19-21, 26-28, 33-35).

Interestingly, the profiles in the non-amphiphilic systems, with [N_{4444}]Cl, are highly similar to the ones with SAILs, particularly with the deprotonated versions of GA in both moieties. This implies that even with an IL that does not self-assemble into an organized structure, the GA promotes the aggregation. A brief visuallization of the simulation snapshopts shown in **Figure 32**, one can confirm that the same GA partition is applied to the hydrotrope IL. In the singly deprotonated version, the IL clusters with the GA inside the IL-rich phase, while in the Deprot. GA (-3) remains disperse in the aqueous phase, as in the systems with SAILs. This fact indicates that the relation between the IL cation polar heads and the GA is more significant to the resulting phase behavior of the system. In turn, this supports the importance of electrostatic interactions in these outcomes. As a hydrotrope IL, the [N_{4444}]Cl seems to be able to increase the solubility of GA in aqueous solutions (in the protonated and singly deprotonated versions of GA). The application of this IL could be extended to be used in separation purposes, even though the purification of GA would be an issue due to the lack of a fully structured IL-phase containing the biomolecule.

The Deprot. GA (-3) version is outside the micelle, with the carboxyl bead pointing towards the surface and close to it, and the hydroxyls even closer to this IL area, indicating the “laying on the micelle” disposition of this GA in SAILs. Moreover, these profiles were verified for the hydrotrope IL, [N_{4444}]Cl, indicating that the same GA orientation towards the IL molecules is present. These results reinforce the hypothesis that the partition of GA in these solutions is pH-driven, as the same trend is denoted in all ILs, no matter the solubilisation mechanism and with the same influence by its protonation state. Another factor contributing to this is that even at lower concentrations, this fact is observed, with thinner differences in the hydrotropic systems, as the concentration plays a more important role in this type of coadjuvant.

The greater degree of GA interaction with the IL charged group engenders a dehydration of the IL, the extent of which depends on the IL nature and the pH. The influence of GA in the IL phase behaviour was assessed by verifying the hydration of both the IL and GA in each system, to check any possible ion-exchange and dehydration phenomena. For the former, the RDF profiles relating the IL cation polar heads and the IL anion were developed. These RDFs provided information which concord with the previously mentioned GA-polar heads RDFs in the discussion of the decoupling of chloride and exchange with deprotonated GA as the new anion. For the latter, the RDFs regarding the hydration of each component and their SASA profiles were prompted, with the summary of these results gathered in **Table 10**. This was already discussed in the previous subsection. The RDFs that were mainly used to characterize the hydration of the IL are depicted in **Figure 39**. The RDF and SASA profiles obtained to complement this information are shown in **Figures SI:24-27**. Finally, the RDF profiles used for the discussion about ion-exchange are presented in **Figure 40**.

Regarding the hydration of the ILs, it is worth to notice that the IL solvation was altered by the addition of GA, althought its extent depends on the IL nature. In case of the [N_{11114}]Cl system, the addition of GA generally decreases the $g^*(r)$ values in both, the polar heads and the alkyl chains of the IL cation. However, the CN reduction is not associated with the changes in the RDF profile for these groups, as these follow the same trend in all protonation states. The SASA values also corroborate this fact by a decrease in the available area for solvation. Concentration also plays a role since not only the pH effects are intensified at higher GA concentrations. Although the alkyl chains of $[N_{11114}]^+$ seem to have inverse behaviours before and after the addition of GA at different concentrations, the CN and SASA values neglect this as significant. Finally, the $[N_{44414}]$Cl system depicted a similar trend, however with a more intense distinction between each GA protonation state as well as before and after the GA addition, in both IL cation groups. The higher GA concentration intensifies the effect of deprotonated GA, which reduce more significantly the hydration of $[N_{44414}]^+$. This could be caused by the larger number of GA molecules near the polar heads of this cation when deprotonated, in turn shielding the inner side of the IL-phase, composed by the alkyl chains. The Prot. GA affects more significantly the alkyl chains than the polar heads, as it is partitioned in the micelle core or deeper in the IL-phase. Then, again, the CNs and SASA values confirm these changes, except for the Deprot. GA (-1) system at low GA concentration, were a slight increase in the CN of water around the polar heads is noted, as well as the SASA profile. Still, this difference is not high enough to change the hydration of this group.

The hydrotrope IL also depicted similar RDF profiles than the other ILs, althought the polar heads and alkyl chains have an inverse behaviour. This is caused by the hydrotrope structure, which has the polar head in the center of the cation surrounded by the alkyl chains. This inevitably reduces the hydration of the polar head, even without the addition of GA. Still, the latter influences both groups, by partitioning in the IL cluster and shielding the hydrotrope molecules from water. The concentration effect is more prominent in this IL than in the others, since at higher GA concentrations, the IL clusters are formed and the partition of GA enhance the hydration of each component. Since the same GA distribution is noted for this hydrotrope as in the SAILs, althought the different supramolecular structure (IL-rich phase instead of a micelle), the same hydration effect is noticed. Nonetheless, it is important to note that at lower GA concentrations, the hydration of this IL is isotropic, differing from SAILs since the hydrotrope is disperse in the aqueous solution.

Overall, a dehydration of the IL cation is caused by the addition of GA and increased by the GA concentration and for the bulky [N_{44414}]Cl and [N_{4444}]Cl. No matter which solubilisation mechanism comes into play, this effect is mainly (i) pH-driven and (ii) dependent on the extent of charge shielding of the IL cation. The deprotonation of GA renders its partition in the IL cation polar head shielding positions, either directly in the SAILs or indirectly with the hydrotrope IL. Consequentely, the alkyl chains are indirectly affected in the former, and directly affected in the latter.

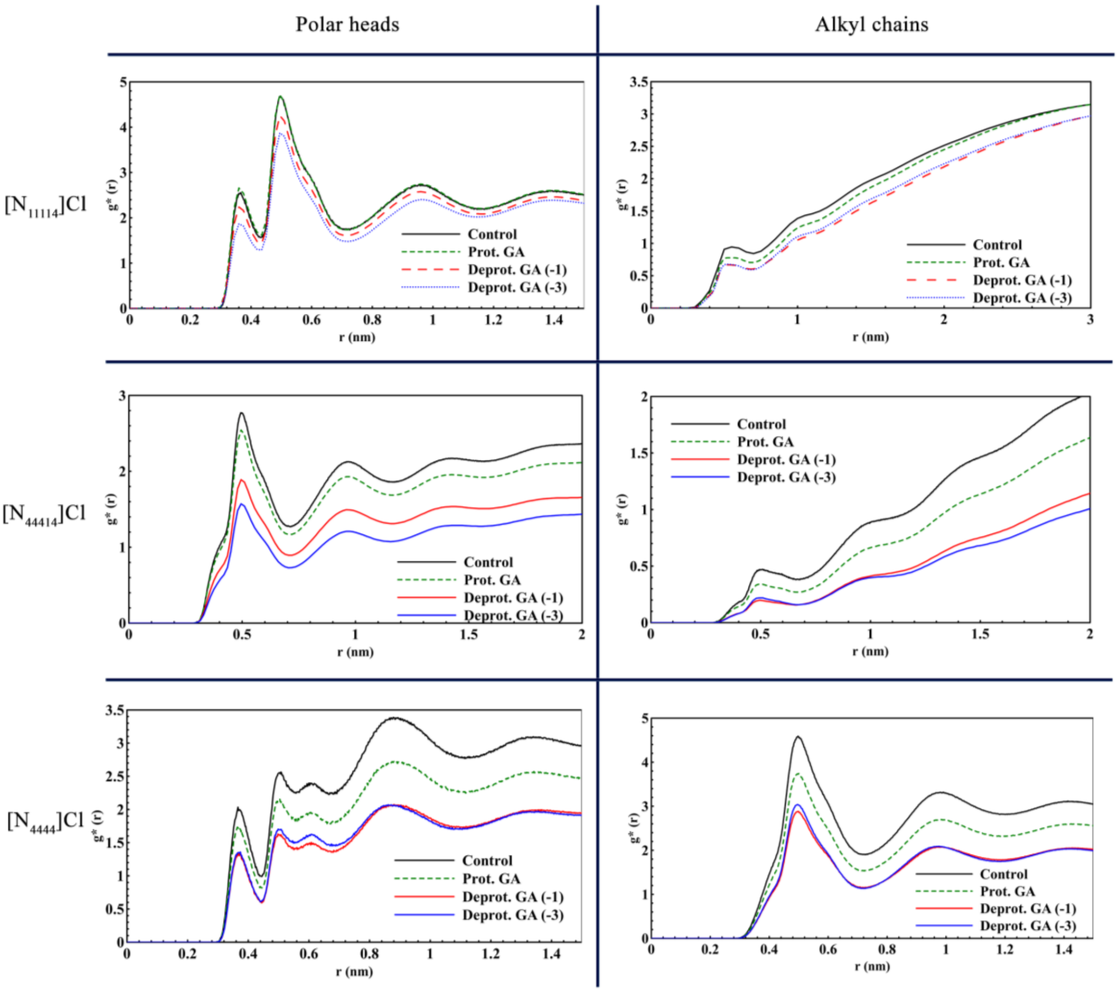


Figure 39 – RDF profiles using either the polar heads or the alkyl chains group of each IL cation as the reference and the PW water as the selection (systems 15, 19-21, 22, 26-28, 29, 33-35).

e) Influence of Gallic acid on phase transition –ion-exchange

The partition of the deprotonated GA demonstrated that the interaction between the polar heads and the GA at this pH could imply the substitution of the chloride anion of the IL by the GA. Such ion-exchange hypothesis was then explored with the previous RDF profiles of the cations polar heads with each GA group as well as further RDFs to confirm the decoupling of chloride from the IL cation as shown in **Figure 40**.

In general, the only similarity between the three profiles, each for a particular IL, regards the protonation state of the GA. In all of cases, the systems with Prot. GA have their IL cation polar heads more coupled to the IL anion, followed by a decrease in the $g(r)$ values when the GA is singly deprotonated and even more when the pH increases. Thus, it is possible to conclude that the GA partition of deprotonated molecules, either at the surface of the IL cation phase (Deprot. GA (-1)) or outside but shielding the polar heads (Deprot. GA (-3)) yields to an anion exchange mechanism. The IL cation is maintained, while the chloride anion is exchanged by the deprotonated GA. This matches with the conclusions obtained with the density profiles, the RDFs of each GA moiety and the IL cation polar heads.

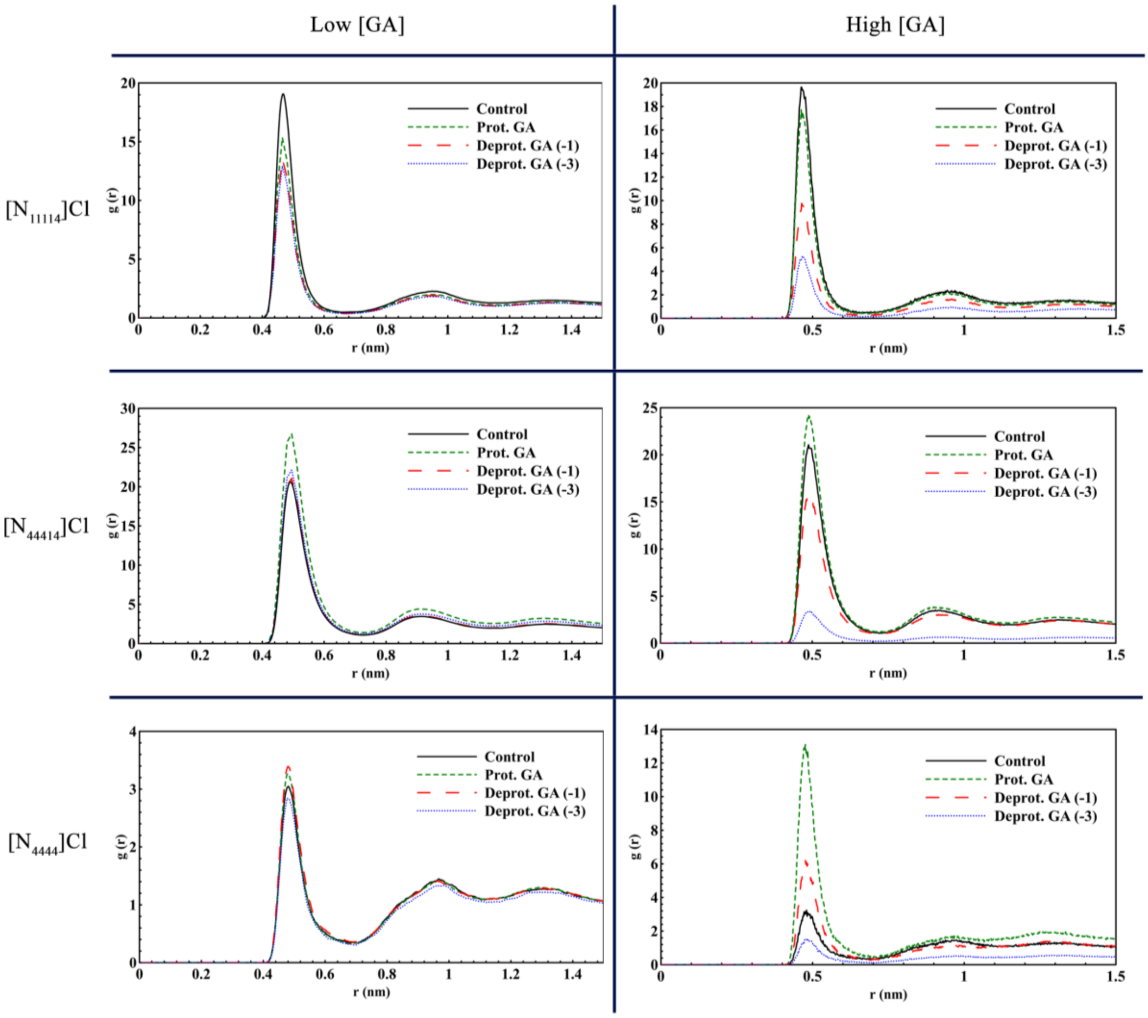


Figure 40 – RDF profiles using the polar heads of each IL cation as the reference and the respective IL anion (chloride) as the selection (systems 15-35).

Divergences occurs between the ILs when comparing the systems before and after the GA addition. In the [N$_{11114}$]Cl system, the addition of GA reduces the $g(r)$ values in all cases, implying that there is a decrease in the accumulation of chloride at the micellar interface. In the [N$_{44414}$]Cl system, only the Prot. GA increases the presence of chloride anions around the IL cation polar heads. However, in the [N$_{4444}$]Cl system, not only the Prot. GA but also the Deprot. GA (-1) system is above the one without GA addition. The $g(r)$ values for the Prot. GA seem more related to the hydrotrope mechanism of these ILs, as they were only higher than the control in the bulky cations ([N$_{44414}$]Cl shares this aspect as an hydrotropic contribution). Such values are caused by the neutral charge of the GA, which is able to interact with the butyl chains in the polar heads of the IL cation as they have an apolar character. Both ILs with a partial or full hydrotropic nature, differ in the singly deprotonated version of GA. This fact can be explained by the differences in the structural organization of both ILs. While in the ILs, the Deprot. GA (-1) partitions inside the IL micellar structures (even in a phase separated $[N_{44414}]$Cl which is preceded by micelles and their connection), in the $[N_{4444}]$Cl, this structuring is not present. The electrostatic interactions between the negatively charged carboxylate bead and the positive ammonium bead are not as strong as when $[N_{11114}]$Cl is used due to the alkyl chains in the polar head. Thus the GA is not able to shield the central ammonium center, allowing the interaction with the IL anion. Combining this with the mesophase of the ILs, the intermediate behavior of $[N_{44414}]$Cl is explained by the IL structure locking the GA, mitigating this effect towards the IL cation polar head. All of this is corroborated by the profiles of the Deprot. GA (-1), which follow the ascending order in terms of comparison with the control values: $[N_{11114}]^+ < [N_{44414}]^+ < [N_{4444}]^+$. Since the $[N_{4444}]$Cl hydrotrope character comes into play, the concentration plays a larger role than in the other ILs. The presence of the Deprot. GA (-3) reduces the polar head-chloride interaction, but not as significantly as in the SAILs. This indicates that the novel IL with the Prot. GA may only be truly established using the SAILs and not the hydrotrope IL. As it was pointed out, the alkyl chains in the hydrotrope cation shield the central ammonium center from the electrostatic interactions. However, the degree of structuring induced by the SAIL behavior seems also necessary for an effective exchange to occur. At lower GA concentrations, the GA addition effect is confirmed for all of the ILs, with the small differences due to pH.

f) Summary of findings

With all these results and conclusions in mind, it is possible to summarize the effects of the different solubilisation mechanisms and comparisons between all systems and their efficacy in particular applications. Thus these result in a few points:

- $[N_{11114}]$Cl phase behaviour, *i.e.*, micellar arrangement, is constant in all systems, demonstrating this IL as a potential vessel in drug delivery applications, as reported before for these type of cationic ILs,[163] as well as a low toxicity[164]; the GA is placed well inside the IL micelle at the pH corresponding to an acidic-to-neutral medium, being ideal for oral formulations; The biomolecule is aligned with the micelle constituents at neutral pH, being more stable due to electrostatic interactions between the negatively charged GA and the positive IL cation polar head; The affinity between the apolar tail of GA (aromatic ring and protonated hydroxyls) should also contribute to this, althought not as significantly as the charged moiety;

- $[N_{44414}]Cl$ depicts a phase transition from micellar-to-bicontinuous phase and eventually phase separation, being more applicable to extraction and purification purposes; Considered as an intermediate IL, sharing properties with the $[N_{11114}]Cl$ (surfactant) and the $[N_{4444}]Cl$ (hydrotrope), regarding the phase behavior and the GA partition; the GA partitions in the micelles proceeds accordingly with ILs, but the bulky cation polar head with a more apolar character yields a phase separation, with the GA partitioning within the same positions but in a more spurius orientation;
- $[N_{4444}]Cl$ exhibits a hydrotrope character which is useful to enhance the solubility of GA as demonstrated in a recent work[5]; This hydrotropy is associated both with the lack of a long alkyl tail and the bulkier cation polar head; The former would promote a higher structuring degree while the latter shields the charged moiety of the group from repelling each others, giving room for the aggregation; The dominant forces between this IL and the GA were closely related to the ones in $[N_{44414}]Cl$, indicating that the polar head is the main region of the IL interacting with the GA (due to the presence of a charged bead); The lack of long alkyl chain in the IL cation resulted in the GA partitioning in IL clusters rather than IL micelles ($[N_{11114}]Cl$) or micelles/IL-rich phase ($[N_{44414}]Cl$);
- The high prevalence of APIs with similar structures than those used in this project (hydrophobic or acidic character) improves the transferability of these assays to other drugs; Other cationic ILs, either surfactants or hydrotropes, can possibly be used due to the nature of the most prevalent interactions being their charge;
- A wide, realistically applicable pH range in experimental works was covered by the three different protonation states selected for GA, not frequently explored by CG-MD simulations; While for acidic-to-neutral pH values interesting results were obtained for drug delivery and biomolecule extraction, for neutral-to-basic pH values the possibility of forming an API-IL was assessed;
- A pH-dependent partition of GA was demonstrated, with the state of the biomolecule affecting not only the partition but also the IL phase behaviour; a similar trend was

verified for the partition of the GA in all of the ILs studied; the IBP, with a structure similar to GA, was also used to control the phase behaviour of the IL and consequentely the drug release profile in a previous work;[21] This renders a new application of these biomolecules: tuning the phase diagrams of ILs;

- This pH dependency induced that electrostatic interactions carried the main role in the force balance that influences the phase behavior of the system and the GA partition; The polar head of the IL cation, which contains the charged moiety of the IL and the ionizable regions of GA (carboxylic and hydroxylic groups), are the driving forces of the resulting system;
- At acidic-to-neutral pH (physiological), the GA partitions inside the $[N_{11114}]Cl$ and the $[N_{44414}]Cl$ micelles (and the IL-phase in the latter when applicable), demonstrating the ability of these ILs for drug delivery and extraction purposes; This pH interval is the most prevalent in biological media; At this pH, the aggregation of $[N_{4444}]Cl$ around the GA is also highly efficient, showing the improved hydrotrope action; The lack of a defined IL structure hinders, however, the use in drug delivery, and in extraction to a certain extent;
- At basic pH, the triple deprotonated GA was generally partitioned in the aqueous phase, which is not applicable for GA extraction or delivery the GA in aqueous solutions;
- A novel IL was formed when the GA was negatively charged and the SAILs were used, particularly at higher GA concentrations; Ion-exchange occurred between the anionic GA and the chloride; Such process was possible without resorting other coadjuvants such as salts; Although the pH was the main parameter affecting this, concentrations above the solubility limit of GA trigger phase transitions;
- The production of API-ILs could enhance the beneficial properties of the API while actually promoting the solubility in aqueous media;[165] the ammonium-based ILs were used to produce API-IL-polymer vessels for drug delivery applications controlling the release of the API with the interaction between the API and the IL;[166] the sodium salicylate was also used for this purpose in dermal drug delivery;[167]
- The pH-dependent behaviour of GA partition and the consequent effect on the IL phase behaviour was present in all ILs with a larger impact in SAILs ($[N_{11114}]Cl$ and $[N_{44414}]Cl$); for the hydrotrope IL ($[N_{4444}]Cl$), this factor had its weight shared with

the GA concentration, since it does not have a well-defined supramolecular structure that optimizes certain interactions between the IL and the GA;

- $[N_{44414}]$Cl demonstrated an intermediate behaviour with both surfactant and hydrotrope capabilities, since the pH and the concentration were balanced factors in the phase behaviour and the GA partition; the pH is the dominant factor in the GA partition while concentrations apply the effects of different GA protonations states towards the IL;

These results render the micellar solubilisation mechanism by SAILs with the potential application in drug delivery and extraction, while the hydrotrope mechanism could be used for extraction purposes and as a solubility increase. The $[N_{11114}]$Cl could be the best SAIL for drug delivery due to its phase stability and the GA partition at physiological pH. The $[N_{44414}]$Cl shares the best of both, with micellar arrangements for drug delivery at low GA concentrations but more useful in extraction due to the presence of the bicontinuous phase or the IL separation at higher concentrations and pH. The $[N_{4444}]$Cl could be used for either extraction, especially at basic pH values, but generally as an efficient solubility enhancer. The phase behavior of this IL with GA indicates that the mechanism of action is closer to the one described by the existence of a MHC. This is explained by the hydrotrope clusters formation around the biomolecule with a particular orientation of both the IL and the GA to establish the electrostatic interactions. Even though, the cluster resembles a micellar shape, but not constitute one. However, our results do not prove the existence of a MHC as the driving hydrotropic mechanism, but rather the need of its combination with a particular pH and also the biomolecule concentration. It is only when the biomolecule has one or more charged regions and at a high concentration, that this mechanism prevails. Nonetheless, at lower GA concentrations initial signs of this hydrotrope effect are noticeable.

Chapter IV: Conclusions and future work

In this work the main goal was to rationalize the partition of GA in IL-based aqueous solutions to assess the effect of ILs with different solubilisation mechanisms. The ultimate purpose is the development of an easy way, transferrable platform to rapidly screen APIs for drug delivery applications, and to optimize extraction processes. Bearing this in mind, a CG framework for MD simulations for GA + IL aqueous solutions was developed. The [N_{11114}]Cl was used as SAIL with a micellar solubilisation mechanism, while [N_{4444}]Cl was used as an IL with hydrotropic action. The [N_{44414}]Cl was approached as an intermediate option, with a behaviour in between the SAILs and the hydrotropes. The CG models of these ILs had already been developed and validated. Novel models for GA and IBP were created and validated both with experimental and computational data. Trials regarding the SAILs and salt were performed to confirm their phase behaviour as it is well established in various publications. These captured the differences in both polar heads of the IL cations, regarding the phase behaviour of the IL before and after salt addition. It was denoted a phase separation for the larger IL, proposing it as the most efficient for extraction and the smaller IL for drug delivery. Overall, the GA partition was pH-dependent for the systems containing GA and IL, being intensified by the increase in the GA concentration. While for the [N_{11114}]Cl this effect was dominant, for the hydrotrope GA concentration shared strength with pH, both in terms of influence in IL phase behaviour. In the [N_{44414}]Cl, both parameters influenced strongly the IL phase behaviour. The SAILs, especially [N_{11114}]Cl, have the potential to be used in drug delivery applications. The [N_{44414}]Cl could be appropriate in this case at lower GA concentrations, but the phase is more unstable, moving towards a phase separation which is beneficial for extraction processes rather than in drug delivery. Finally, the hydrotrope IL would be useful in general to increase the solubility of GA in aqueous solutions, but beneficial for extraction purposes at higher GA concentrations. A possible novel GA-IL was also formed, particularly with SAILs, when GA was simulated at basic pH values.

The GA molecule was the main focus of this work, paving the way for future assays using other drugs such as the IBP. Preliminary results using the IBP and the [N_{11114}]Cl predict a partition behaviour similar to that verified in this project. In a future work, it is intended to use this API and eventually others with different ILs to further complement the data for each possible combination. The addition of salt, used for SAILs in water but not in systems containing GA, is also an interesting effect to study to assess the weight of using or not these conventional coadjuvants. In this work, this was not explored as the GA which demonstrated the possibility of being able to affect ILs without the addition of salt, contributing to the simplicity and cost-efficiency of these systems. Another interesting attempt would be to use recent MD methods to better depict the presence of the various protonation states of the molecules used, at the correct concentration of each, such as the constant-pH method.[168]

Chapter V: Bibliography

Chapter V: Bibliography

1. United Nations, 2015. Transforming our world: The 2030 agenda for sustainable development. General Assembly resolution 70/1, A/RES/70/1.
2. Ventura, S.P.M., e Silva, F.A., Quental, M. V., Mondal, D., Freire, M.G., Coutinho, J.A.P., 2017. Ionic-Liquid-Mediated Extraction and Separation Processes for Bioactive Compounds: Past, Present, and Future Trends. Chem. Rev. 117, 6984–7052.
3. Brodland, G.W., 2015. How computational models can help unlock biological systems. Semin. Cell Dev. Biol. 47–48, 62–73.
4. Savjani, K.T., Gajjar, A.K., Savjani, J.K., 2012. Drug Solubility: Importance and Enhancement Techniques. ISRN Pharm. 2012, 195727.
5. Cláudio, A.F.M., Neves, M.C., Shimizu, K., Canongia Lopes, J.N., Freire, M.G., Coutinho, J.A.P., 2015. The magic of aqueous solutions of ionic liquids: ionic liquids as a powerful class of catanionic hydrotropes. Green Chem. 17, 3948–3963.
6. Drioli, E., Giorno, L., 2016. Encyclopedia of Membranes, 1st ed. Springer, Berlin, Heidelberg.
7. Raffa, P., Wever, D.A.Z., Picchioni, F., Broekhuis, A.A., 2015. Polymeric surfactants: Synthesis, properties, and links to applications. Chem. Rev. 115, 8504–8563.
8. Lombardo, D., Kiselev, M.A., Magazù, S., Calandra, P., 2015. Amphiphiles Self-Assembly: Basic Concepts and Future Perspectives of Supramolecular Approaches. Adv. Cond. Matter Phys. 2015, 1–22.
9. Gadre, S.R., Yeole, S.D., Sahu, N., 2014. Quantum Chemical Investigations on Molecular Clusters. Chem. Rev. 114, 12132–12173.
10. Israelachvili, J.N., 2011. Intermolecular and Surface Forces, 3rd ed. Academic Press, Elsevier, United States.
11. Santos, D., Rufino, R., Luna, J., Santos, V., Sarubbo, L., Santos, D.K.F., Rufino, R.D., Luna, J.M., Santos, V.A., Sarubbo, L.A., 2016. Biosurfactants: Multifunctional Biomolecules of the 21st Century. Int. J. Mol. Sci. 17, 401.
12. Kumar, S., Bhargava, P., Sreekanth, V., Bajaj, A., 2015. Design, synthesis, and physico-chemical interactions of bile acid derived dimeric phospholipid amphiphiles with model membranes. J. Colloid Interf. Sci. 448, 398–406.
13. Yu, S., Tan, H., Wang, J., Liu, X., Zhou, K., 2015. High Porosity Supermacroporous Polystyrene Materials with Excellent Oil–Water Separation and Gas Permeability Properties. ACS Appl. Mater. Inter. 7, 6745–6753.
14. Zhang, P., Liu, Y., Fan, M., Jiang, P., 2016. Catalytic performance of a novel amphiphilic alkaline ionic liquid for biodiesel production: Influence of basicity and conductivity. Renew. Energ. 86, 99–105.
15. López-Martínez, A., Rocha-Uribe, A., 2018. Antioxidant Hydrophobicity and Emulsifier Type Influences the Partitioning of Antioxidants in the Interface Improving Oxidative Stability in O/W Emulsions Rich in n-3 Fatty Acids. Eur. J. Lipid Sci. Tech. 120, 1700277.
16. Fan, X., Wang, T., Miao, W., 2018. The preparation of pH-sensitive hydrogel based on host-guest and electrostatic interactions and its drug release studies in vitro. J. Polym. Res. 25, 215.

17. Vinarov, Z., Katev, V., Radeva, D., Tcholakova, S., Denkov, N.D., 2018. Micellar solubilization of poorly water-soluble drugs: effect of surfactant and solubilizate molecular structure. Drug Dev. Ind. Pharm. 44, 677–686.
18. Qamar, S., Brown, P., Ferguson, S., Khan, R.A., Ismail, B., Khan, A.R., Sayed, M., Khan, A.M., 2016. The interaction of a model active pharmaceutical with cationic surfactant and the subsequent design of drug based ionic liquid surfactants. J. Colloid Interf. Sci. 481, 117–124.
19. Liu, Y.-J., Lo, S.-L., Liou, Y.-H., Hu, C.-Y., 2015. Removal of nonsteroidal anti-inflammatory drugs (NSAIDs) by electrocoagulation–flotation with a cationic surfactant. Sep. Purif. Technol. 152, 148–154.
20. Hu, S., Lee, E., Wang, C., Wang, J., Zhou, Z., Li, Y., Li, X., Tang, J., Lee, D.H., Liu, X., Shen, Y., 2015. Amphiphilic drugs as surfactants to fabricate excipient-free stable nanodispersions of hydrophobic drugs for cancer chemotherapy. J. Control. Release 220, 175–179.
21. Pasquino, R., De Gennaro, B., Gaudino, D., Grizzuti, N., 2017. On the Use of Nonsteroidal Anti-Inflammatory Drugs as Rheology Modifiers for Surfactant Solutions. J. Pharm. Sci. 106, 3410–3412.
22. Chatzidaki, M.D., Papavasileiou, K.D., Papadopoulos, M.G., Xenakis, A., 2017. Reverse Micelles As Antioxidant Carriers: An Experimental and Molecular Dynamics Study. Langmuir 33, 5077–5085.
23. Dey, A., Sandre, V., Marangoni, D.G., Ghosh, S., 2018. Interaction between a Nonsteroidal Anti-inflammatory Drug (Ibuprofen) and an Anionic Surfactant (AOT) and Effects of Salt (NaI) and Hydrotrope (4-4-4). J. Phys. Chem. B 122, 3974–3987.
24. Campos, V.E.B. De, Cerqueira-Coutinho, C.S., Capella, F.N.C., Soares, B.G., Holandino, C., Mansur, C.R.E., 2017. Development and In Vitro Assessment of Nanoemulsion for Delivery of Ketoconazole Against Candida albicans. J. Nanosci. Nanotechno. 17, 4623–4630.
25. Pacheco-Fernández, I., Pino, V., Ayala, J.H., Afonso, A.M., 2018. Guanidinium ionic liquid-based surfactants as low cytotoxic extractants: Analytical performance in an in-situ dispersive liquid–liquid microextraction method for determining personal care products. J. Chromatogr. A 1559, 102–111.
26. Gras, M., Papaiconomou, N., Schaeffer, N., Chainet, E., Tedjar, F., Coutinho, J.A.P., Billard, I., 2018. Ionic-Liquid-Based Acidic Aqueous Biphasic Systems for Simultaneous Leaching and Extraction of Metallic Ions. Angew. Chem. Int. Ed. 57, 1563–1566.
27. Kandasamy, S., Moniruzzaman, M., Sivapragasam, M., Shamsuddin, M.R., Mutalib, M.I.A., 2018. Formulation and characterization of acetate based ionic liquid in oil microemulsion as a carrier for acyclovir and methotrexate. Sep. Purif. Technol. 196, 149–156.
28. Pal, A., Yadav, A., 2019. Mixed micellization of a trisubstituted surface active ionic liquid 1-dodecyl-2,3-dimethylimidazolium chloride [C12bmim][Cl] with an amphiphilic drug amitriptyline hydrochloride AMT: A detailed insights from conductance and surface tension measurements. J. Mol. Liq. 279, 43–50.

29. Vashishat, R., Chabba, S., Aswal, V.K., Mahajan, R.K., 2017. Probing molecular interactions of tetracaine with surface active ionic liquid and subsequent formation of vesicle in aqueous medium. J. Mol. Liq. 243, 503–512.
30. Pyne, A., Kundu, S., Banerjee, P., Sarkar, N., 2018. Unveiling the Aggregation Behavior of Doxorubicin Hydrochloride in Aqueous Solution of 1-Octyl-3-methylimidazolium Chloride and the Effect of Bile Salt on These Aggregates: A Microscopic Study. Langmuir 34, 3296–3306.
31. Bodratti, A.M., Sarkar, B., 2017. Adsorption of poly(ethylene oxide)-containing amphiphilic polymers on solid-liquid interfaces: Fundamentals and applications. Adv. Colloid Interfac. 244, 132–163.
32. Sousa, R. de C.S., Pereira, M.M., Freire, M.G., Coutinho, J.A.P., 2018. Evaluation of the effect of ionic liquids as adjuvants in polymer-based aqueous biphasic systems using biomolecules as molecular probes. Sep. Purif. Technol. 196, 244–253.
33. Korenman, Y. I., Makarova, T. V., 1974. Salting out and salting in of 2-Halophenols. J. phys. Chem, 48 (2), 385- 387.
34. Neuberg, C., 1916. Hydrotropy. Biochem. Z. 76, 107-176.
35. Lévêque, J-M., Cravotto, G., Delattre, F., Cintas, P., 2018. Efficient Organic Synthesis: What Ultrasound Makes Easier in Organic Sonochemistry. Springer, Switzerland, pp. 17-39.
36. Nidhi, K., Indrajeet, S., Khushboo, M., Gauri, K., Sen, D., 2011. Hydrotropy: A Promising tool for solubility Enhancement. Int. J. Drug Dev. & Res., 3(2), 26-33.
37. Dhapte, V., Mehta, P., 2015. Advances in hydrotropic solutions: An updated review. St. Petersbg. Polytech. Univ. J. Phys. Math. 1, 424–435.
38. Sintra, T.E., Shimizu, K., Ventura, S.P.M., Shimizu, S., Canongia Lopes, J.N., Coutinho, J.A.P., 2018. Enhanced dissolution of ibuprofen using ionic liquids as catanionic hydrotropes. Phys. Chem. Chem. Phys. 20, 2094–2103.
39. Das, R.P., Gandhi, V. V., Singh, B.G., Kunwar, A., Kumar, N.N., Priyadarsini, K.I., 2019. Preparation of albumin nanoparticles: Optimum size for cellular uptake of entrapped drug (Curcumin). Colloids Surfaces A Physicochem. Eng. Asp. 567, 86–95.
40. Takahashi, K., Komai, M., Kinoshita, N., Nakamura, E., Hou, X.-L., Takatani-Nakase, T., Kawase, M., 2011. Application of hydrotropy to transdermal formulations: hydrotropic solubilization of polyol fatty acid monoesters in water and enhancement effect on skin permeation of 5-FU. J. Pharm. Pharmacol. 63, 1008–1014.
41. Das, S., Paul, S., 2016. Computer Simulation Studies of the Mechanism of Hydrotrope-Assisted Solubilization of a Sparingly Soluble Drug Molecule. J. Phys. Chem. B 120, 3540–3550.
42. Singh, O., Kaur, R., Aswal, V.K., Mahajan, R.K., 2016. Composition and Concentration Gradient Induced Structural Transition from Micelles to Vesicles in the Mixed System of Ionic Liquid–Diclofenac Sodium. Langmuir 32, 6638–6647.
43. Wang, J., Liu, K., Xing, R., Yan, X., 2016. Peptide self-assembly: thermodynamics and kinetics. Chem. Soc. Rev. 45, 5589–5604.
44. Lee, Y.S., 2008. Self-assembly and nanotechnology: a force balance approach. John Wiley & Sons, New Jersey, United States.

45. Mahadevi, A.S., Sastry, G.N., 2016. Cooperativity in Noncovalent Interactions. Chem. Rev. 116, 2775–2825.
46. Djikaev, Y.S., Ruckenstein, E., 2016. Recent developments in the theoretical, simulational, and experimental studies of the role of water hydrogen bonding in hydrophobic phenomena. Adv. Colloid Interfac. 235, 23–45.
47. Grabowski, S.J., 2006. Challenges and Advances in Computational Chemistry and Physics: Hydrogen bonding - New insights, 3rd volume. Springer, Dordrecht, The Netherlands.
48. Hunt, P.A., Ashworth, C.R., Matthews, R.P., 2015. Hydrogen bonding in ionic liquids. Chem. Soc. Rev. 44, 1257–1288.
49. Kim, S.A., Jeong, K.-J., Yethiraj, A., Mahanthappa, M.K., 2017. Low-symmetry sphere packings of simple surfactant micelles induced by ionic sphericity. P. Natl. Acad. Sci. USA. 114, 4072–4077.
50. Li, Y., Yang, L., 2015. Driving forces for drug loading in drug carriers. J. Microencapsul. 32, 255–272.
51. Ryzhakov, A., Do Thi, T., Stappaerts, J., Bertoletti, L., Kimpe, K., Sá Couto, A.R., Saokham, P., Van den Mooter, G., Augustijns, P., Somsen, G.W., Kurkov, S., Inghelbrecht, S., Arien, A., Jimidar, M.I., Schrijnemakers, K., Loftsson, T., 2016. Self-Assembly of Cyclodextrins and Their Complexes in Aqueous Solutions. J. Pharm. Sci. 105, 2556–2569.
52. Ding, Y., Shen, S.Z., Sun, H., Sun, K., Liu, F., Qi, Y., Yan, J., 2015. Design and construction of polymerized-chitosan coated Fe3O4 magnetic nanoparticles and its application for hydrophobic drug delivery. Mater. Sci. Eng. C 48, 487–498.
53. Reilly, A.M., Tkatchenko, A., 2015. van der Waals dispersion interactions in molecular materials: beyond pairwise additivity. Chem. Sci. 6, 3289–3301.
54. Marina, P.F., Cheng, C., Sedev, R., Stocco, A., Binks, B.P., Wang, D., 2018. Van der Waals Emulsions: Emulsions Stabilized by Surface-Inactive, Hydrophilic Particles via van der Waals Attraction. Angew. Chem. Int. Edit. 130, 9654–9658.
55. Singh, P., Brar, S.K., Bajaj, M., Narang, N., Mithu, V.S., Katare, O.P., Wangoo, N., Sharma, R.K., 2017. Self-assembly of aromatic α-amino acids into amyloid inspired nano/micro scaled architects. Mater. Sci. Eng. C 72, 590–600.
56. Santos, A.P., Panagiotopoulos, A.Z., 2016. Determination of the critical micelle concentration in simulations of surfactant systems. J. Chem. Phys. 144, 044709.
57. Bergström, L. M., 2016. Second CMC in surfactant micellar systems. Curr. Opin. Colloid In. 22, 46–50.
58. Mao, X., Jiang, R., Xiao, W., Yu, J., 2015. Use of surfactants for the remediation of contaminated soils: A review. J. Hazard. Mater. 285, 419–435.
59. Singh, G., Singh, G., Kang, T.S., 2016. Micellization Behavior of Surface Active Ionic Liquids Having Aromatic Counterions in Aqueous Media. J. Phys. Chem. B 120, 1092–1105.
60. Šarac, B., Cerkovnik, J., Ancian, B., Mériguet, G., Roger, G.M., Durand-Vidal, S., Bešter-Rogač, M., 2011. Thermodynamic and NMR study of aggregation of dodecyltrimethylammonium chloride in aqueous sodium salicylate solution. Colloid Polym. Sci. 289, 1597–1607.

61. Alargova, R. G., Kochijashky, I. I., Sierra, M. L., Zana, R., 1998. Micelle Aggregation Numbers of Surfactants in Aqueous Solutions: A Comparison between the Results from Steady-State and Time-Resolved Fluorescence Quenching. Langmuir. 14, 5412-5418.
62. Alothman, Z.A., Siddiqui, M.R., Khan, M.A., 2013. Enciclopedia of Membranes, Asian Journal of Chemistry.
63. Feizi, N., Yamini, Y., Moradi, M., Karimi, M., Salamat, Q., Amanzadeh, H., 2017. A new generation of nano-structured supramolecular solvents based on propanol/gemini surfactant for liquid phase microextraction. Anal. Chim. Acta 953, 1–9.
64. Gargallo, L., DadiĆ, D., Martínez-Piña, F., 1997. Phase separation behaviour of polymers in water. Temperature and surfactant effect. Eur. Polym. J. 33, 1767–1769.
65. Teo, A., Goh, K.K.T., Wen, J., Oey, I., Ko, S., Kwak, H.-S., Lee, S.J., 2016. Physicochemical properties of whey protein, lactoferrin and Tween 20 stabilised nanoemulsions: Effect of temperature, pH and salt. Food Chem. 197, 297–306.
66. Niskanen, J., Tenhu, H., 2017. How to manipulate the upper critical solution temperature (UCST)? Polym. Chem. 8, 220–232.
67. Naqvi, A.Z., Kabir-ud-Din, 2018. Clouding Phenomenon in Amphiphilic Systems: A Review of Five Decades. Colloid. Surface. B. 165, 325–344.
68. Kotsuchibashi, Y., Ebara, M., Aoyagi, T., Narain, R., 2016. Recent Advances in Dual Temperature Responsive Block Copolymers and Their Potential as Biomedical Applications. Polymers (Basel). 8, 380.
69. Schaeffer, N., Pérez-Sánchez, G., Passos, H., Gomes, J.R.B., Papaiconomou, N., Coutinho, J.A.P, 2019. Mechanisms of phase separation in temperature-responsive acidic aqueous biphasic systems. Phys. Chem. Chem. Phys. 21, 7462–7473.
70. Saita, S., Kohno, Y., Ohno, H., 2013. Detection of small differences in the hydrophilicity of ions using the LCST-type phase transition of an ionic liquid–water mixture. Chem. Commun. 49, 93–95.
71. Zhong, Y., Feng, X., Chen, W., Wang, X., Huang, K.-W., Gnanou, Y., Lai, Z., 2016. Using UCST Ionic Liquid as a Draw Solute in Forward Osmosis to Treat High-Salinity Water. Environ. Sci. Technol. 50, 1039–1045.
72. Walden, P., 1914. Molecular weights and electrical conductivity of several fused salts. Bull. Acad. Imper. Sci. 405-422.
73. Earle, M.J., Seddon, K.R., 2000. Ionic liquids. Green solvents for the future. Pure Appl. Chem. 72, 1391–1398.
74. Wilkes, J.S., 2004. Properties of ionic liquid solvents for catalysis. J. Mol. Catal. A-Chem. 214, 11–17.
75. Handy, S.T., 2011. Ionic Liquids - Classes and Properties. InTech, Croatia.
76. Łuczak, J., Hupka, J., Thöming, J., Jungnickel, C., 2008. Self-organization of imidazolium ionic liquids in aqueous solution. Colloid. Surface. A. 329, 125–133.
77. Pei, Y., Hao, L., Ru, J., Zhao, Y., Wang, H., Bai, G., Wang, J., 2018. The self-assembly of ionic liquids surfactants in ethanolammonium nitrate ionic liquid. J. Mol. Liq. 254, 130–136.
78. Hayes, R., Warr, G.G., Atkin, R., 2015. Structure and Nanostructure in Ionic Liquids. Chem. Rev. 115, 6357–6426.

79. Nishimura, N., Ohno, H., 2014. 15th anniversary of polymerised ionic liquids. Polymer. 55, 3289–3297.
80. Passos, H., Luís, A., Coutinho, J.A.P., Freire, M.G., 2016. Thermoreversible (Ionic-Liquid-Based) Aqueous Biphasic Systems. Sci. Rep. 6, 20276.
81. Jensen, F., 2016. Introduction to Computational Chemistry, 3rd ed. John Wiley & Sons, England.
82. Borgia, A., Borgia, M.B., Bugge, K., Kissling, V.M., Heidarsson, P.O., Fernandes, C.B., Sottini, A., Soranno, A., Buholzer, K.J., Nettels, D., Kragelund, B.B., Best, R.B., Schuler, B., 2018. Extreme disorder in an ultrahigh-affinity protein complex. Nature 555, 61–66.
83. Bottaro, S., Bussi, G., Kennedy, S.D., Turner, D.H., Lindorff-Larsen, K., 2018. Conformational ensembles of RNA oligonucleotides from integrating NMR and molecular simulations. Sci. Adv. 4, eaar8521.
84. Li, Y., Ogorzalek, T.L., Wei, S., Zhang, X., Yang, P., Jasensky, J., Brooks, C.L., Marsh, E.N.G., Chen, Z., 2018. Effect of immobilization site on the orientation and activity of surface-tethered enzymes. Phys. Chem. Chem. Phys. 20, 1021–1029.
85. Moradi, S., Taran, M., Mohajeri, P., Sadrjavadi, K., Sarrami, F., Karton, A., Shahlaei, M., 2018. Study of dual encapsulation possibility of hydrophobic and hydrophilic drugs into a nanocarrier based on bio-polymer coated graphene oxide using density functional theory, molecular dynamics simulation and experimental methods. J. Mol. Liq. 262, 204–217.
86. Khatun, S., Riyazuddeen, Yasmeen, S., Kumar, A., Subbarao, N., 2018. Calorimetric, spectroscopic and molecular modelling insight into the interaction of gallic acid with bovine serum albumin. J. Chem. Thermodyn. 122, 85–94.
87. Dabir, T.O., Gaikar, V.G., Jayaraman, S., Mukherjee, S., 2018. Thermodynamic modeling studies of aqueous solubility of caffeine, gallic acid and their cocrystal in the temperature range of 303 K–363 K. Fluid Phase Equilib. 456, 65–76.
88. Kacar, G., 2019. Molecular understanding of interactions, structure, and drug encapsulation efficiency of Pluronic micelles from dissipative particle dynamics simulations. Colloid Polym. Sci. 1–15.
89. Sun, T., Mirzoev, A., Minhas, V., Korolev, N., Lyubartsev, A.P., Nordenskiöld, L., 2019. A multiscale analysis of DNA phase separation: from atomistic to mesoscale level. Nucleic Acids Res. 47, 5550–5562.
90. Dignon, G.L., Zheng, W., Best, R.B., Kim, Y.C., Mittal, J., 2018. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. Proc. Natl. Acad. Sci. U. S. A. 115, 9929–9934.
91. Borhani, D.W., Shaw, D.E., 2012. The future of molecular dynamics simulations in drug discovery. J. Comput. Aided. Mol. Des. 26, 15–26.
92. Pérez-Sánchez, G., Chien, S.-C., Gomes, J.R.B., D. S. Cordeiro, M.N., Auerbach, S.M., Monson, P.A., Jorge, M., 2016. Multiscale Model for the Templated Synthesis of Mesoporous Silica: The Essential Role of Silica Oligomers. Chem. Mater. 28, 2715–2727.
93. Perez-Sanchez, G., Vicente, F.A., Schaeffer, N., Cardoso, I.S., Ventura, S.P.M., Jorge, M., Coutinho, J.A.P., 2019. Rationalizing the Phase Behavior of Triblock-

Copolymers Through Experiments and Molecular Simulations. J. Phys. Chem. C (Publication pending).

94. Peter, C., Kremer, K., 2009. Multiscale simulation of soft matter systems – from the atomistic to the coarse-grained level and back. Soft Matter 5, 4357.
95. Malavasi, G., Menziani, M.C., Pedone, A., Civalleri, B., Corno, M., Ugliengo, P., 2007. A computational multiscale strategy to the study of amorphous materials. Theor. Chem. Acc. 117, 933–942.
96. Jones, R.O., 2015. Density functional theory: Its origins, rise to prominence, and future. Rev. Mod. Phys. 87, 897–923.
97. Cottenier, S., 2002-2013. Density Functional Theory and the family of (L)APW-methods: a step-by-step introduction, 2nd ed. Freely available at http://www.wien2k.at/reg_user/textbooks.
98. Morgado, C., Vincent, M.A., Hillier, I.H., Shan, X., 2007. Can the DFT-D method describe the full range of noncovalent interactions found in large biomolecules? Phys. Chem. Chem. Phys. 9, 448–451.
99. Li, C., Strachan, A., 2015. Molecular scale simulations on thermoset polymers: A review. J. Polym. Sci. Part B Polym. Phys. 53, 103–122.
100. Wu, R., Deng, M., Kong, B., Yang, X., 2009. Coarse-Grained Molecular Dynamics Simulation of Ammonium Surfactant Self-Assemblies: Micelles and Vesicles. J. Phys. Chem. B 113, 15010–15016.
101. Dong, K., Liu, X., Dong, H., Zhang, X., Zhang, S., 2017. Multiscale Studies on Ionic Liquids. Chem. Rev. 117, 6636–6695.
102. Pérez-Sánchez, G., Gomes, J.R.B., Jorge, M., 2013. Modeling Self-Assembly of Silica/Surfactant Mesostructures in the Templated Synthesis of Nanoporous Solids. Langmuir 29, 2387–2396.
103. Chien, S.-C., Pérez-Sánchez, G., Gomes, J.R.B., Cordeiro, M.N.D.S., Jorge, M., Auerbach, S.M., Monson, P.A., 2017. Molecular Simulations of the Synthesis of Periodic Mesoporous Silica Phases at High Surfactant Concentrations. J. Phys. Chem. C. 121, 4564–4575.
104. Aldeghi, M., Heifetz, A., Bodkin, M.J., Knapp, S., Biggin, P.C., 2017. Predictions of Ligand Selectivity from Absolute Binding Free Energy Calculations. J. Am. Chem. Soc. 139, 946–957.
105. Vchirawongkwin, V., Kritayakornupong, C., Tongraar, A., Rode, B.M., 2012. Characterization of structure and dynamics of an aqueous scandium(iii) ion by an extended ab initio QM/MM molecular dynamics simulation. Dalton Trans. 41, 11889.
106. Ponder, J. W., Case, D. A., 2003. Force fields for protein simulations. Adv. Prot. Chem. 66, 27-85.
107. Vanommeslaeghe, K. Hatcher, E. Acharya, C. Kundu, S. Zhong, S. Shim, J. E. Darian, E. Guvench, O. Lopes, P. Vorobyov, I., MacKerell, Jr. A.D., 2010. CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, J. Comput. Chem. 31, 671-690.
108. Jorgensen, W.L., Maxwell, D.S., Tirado-Rives, J., 1996. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. J. Am. Chem. Soc. 118, 11225–11236.

109. Van Gusteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P., Mark A. E., Scott, W. R. P., Tironi, I. G., 1996. Biomolecular Simulation: The GROMOS96 Manual and User Guide, Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland, 1-1042.
110. Abraham, M. J., van der Spoel, D., Lindahl, E., Hess, B., and the GROMACS development team, *GROMACS User Manual version 5.1.5*, www.gromacs.org (2017).
111. Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P., De Vries, A.H., 2007. The MARTINI force field: Coarse grained model for biomolecular simulations. J. Phys. Chem. B 111, 7812–7824.
112. Rapaport, D.C., 2004. The Art of Molecular Dynamics Simulation, 2nd ed. Cambridge University Press, Cambridge, United Kingdom.
113. Riniker, S., 2018. Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview. J. Chem. Inf. Model. 58, 565–578.
114. Karkute, S.G., Easwaran, M., Gujjar, R.S., Piramanayagam, S., Singh, M., 2015. Protein modeling and molecular dynamics simulation of SlWRKY4 protein cloned from drought tolerant tomato (Solanum habrochaites) line EC520061. J. Mol. Model. 21, 255.
115. Khan, I., Ansari, I.A., 2017. Prediction of a highly deleterious mutation E17K in AKT-1 gene: An in silico approach. Biochem. Biophys. Reports 10, 260–266.
116. Frenkel, D., Smit, B., 1996. Understanding Molecular Simulation: From Algorithms to Applications, 2nd ed. Academic Press, San Diego.
117. Hansen, J-P., McDonald, I.R., 2013. Theory of Simple Liquids, 4th ed. Academic Press
118. Lorenz, C., Doltsinis, N.L., 2017. Molecular Dynamics Simulation: From “Ab Initio” to “Coarse Grained” in Leszczynski, J. *et al* Handbook of Computation Chemistry, 2nd ed. Springer, Switzerland, pp. 337-396.
119. Hockney, R., Goel, S., Eastwood, J., 1974. Quiet High-Resolution Computer Models of a Plasma. J. Comput. Phys. 14 (2), 148–158
120. Parrinello, M., Rahman, A., 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. J. Appl. Phys. 52, 7182-7190.
121. Bussi, G., Donadio, D., Parrinello, M., 2007. Canonical sampling through velocity rescaling. J. Chem. Phys. 126, 014101.
122. Nosé, S., 1984. A molecular dynamics method for simulations in the canonical ensemble. Mol. Phys. 52, 255-268.
123. Darden, T., York, D., Pedersen, L., 1993. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. J. Chem. Phys. 98, 10089–10092.
124. Hess, B., Bekker, H., Berendsen, H. J. C., Fraaije, J. G. E. M., 1997. LINCS: a linear constraint solver for molecular simulations. J. Comput. Chem. 18, 1463-1472.
125. Bekker, H., Dijkstra, E.J., Renardus, M.K.R., Berendsen, H.J.C., 1995. An Efficient, Box Shape Independent Non-Bonded Force and Virial Algorithm for Molecular Dynamics. Mol. Simul. 14, 137–151.
126. Breneman, C.M., Wiberg, K.B., 1990. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. J. Comput. Chem. 11, 361–373.

127. Vueba, M.L., Pina, M.E., Batista de Carvalho, L.A.E., 2008. Conformational Stability of Ibuprofen: Assessed by DFT Calculations and Optical Vibrational Spectroscopy. J. Pharm. Sci. 97, 845–859.
128. Lopes, J.N.C., Dechamps, J., Pádua, A.A.H., 2004. Modeling Ionic Liquids Using a Systematic All-Atom Force Field. J. Phys. Chem. B. 108 (6), 2038-2047.
129. Lopes, J.N.C., Pádua, A.A.H., 2006. Molecular Force Field for Ionic Liquids III: Imidazolium, Pyridinium, and Phosphonium Cations; Chloride, Bromide, and Dicyanamide Anions. J. Phys. Chem. B. 110 (39), 19586-19592.
130. Åqvist, J., 1990. Ion-water interaction potentials derived from free energy perturbation simulations. J. Phys. Chem. 94 (21), 8021–8024.
131. Berendsen, H.J.C., Grigera, J.R., Straatsma, T.P., 1987. The missing term in effective pair potentials. J. Phys. Chem. 91, 6269–6271.
132. Socher, E., Sticht, H., 2016. Mimicking titration experiments with MD simulations: A protocol for the investigation of pH-dependent effects on proteins. Sci. Rep. 6, 22523.
133. Swails, J.M., York, D.M., Roitberg, A.E., 2014. Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation.
134. Tam, K.Y., Takács-Novák, K., 2001. Multi-wavelength spectrophotometric determination of acid dissociation constants: a validation study. Anal. Chim. Acta 434, 157–167.
135. Beltrán, J.L., Sanli, N., Fonrodona, G., Barrón, D., Özkan, G., Barbosa, J., 2003. Spectrophotometric, potentiometric and chromatographic pKa values of polyphenolic acids in water and acetonitrile–water media. Anal. Chim. Acta 484, 253–264.
136. Stoyanova, K., Vinarov, Z., Tcholakova, S., 2016. Improving Ibuprofen solubility by surfactant-facilitated self-assembly into mixed micelles. J. Drug Deliv. Sci. Technol. 36, 208–215.
137. Dabir, T.O., Gaikar, V.G., Jayaraman, S., Mukherjee, S., 2018. Thermodynamic modeling studies of aqueous solubility of caffeine, gallic acid and their cocrystal in the temperature range of 303 K–363 K. Fluid Phase Equilib. 456, 65–76.
138. Yazdanian, M., Briggs, K., Jankovsky, C., Hawi, A., 2004. The “High Solubility” Definition of the Current FDA Guidance on Biopharmaceutical Classification System May Be Too Strict for Acidic Drugs. Pharm. Res. 21, 293–299.
139. Kawabata, Y., Wada, K., Nakatani, M., Yamada, S., Onoue, S., 2011. Formulation design for poorly water-soluble drugs based on biopharmaceutics classification system: Basic approaches and practical applications. Int. J. Pharm. 420, 1–10.
140. PubChem, Open Chemistry database, PubChem Substance and Compound database, 2019. Available at https://pubchem.ncbi.nlm.nih.gov/.
141. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The Protein Data Bank. Nucleic Acids Res. 28, 235–242.
142. Kraft, J.F., Vestergaard, M., Schiøtt, B., Thøgersen, L., 2012. Modeling the Self-Assembly and Stability of DHPC Micelles Using Atomic Resolution and Coarse Grained MD Simulations. J. Chem. Theory Comput. 8, 1556–1569.

143. Yesylevskyy, S.O., Schäfer, L. V., Sengupta, D., Marrink, S.J., 2010. Polarizable Water Model for the Coarse-Grained MARTINI Force Field. PLoS Comput. Biol. 6, e1000810.

144. Humphrey, W., Dalke, A., Schulten, K., 1996. VMD – Visual Molecular Dynamics. J. Molec. Graphics. 14, 33-38.

145. Brehm, M., Kirchner, B., 2011. TRAVIS – A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories. J. Chem. Inf. Model. 51, 2007-2023.

146. Eisenhaber, F., Lijnzaad, P., Argos, P. *et al*, 1995. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. J. Comput. Chem. 16, 273-284.

147. Hoshen, J., Kopelman, R., 1976. Percolation and Cluster Distribution. I. Cluster Multiple Labelling Technique and Critical Concentration Algorithm. Phys. Rev. B. 14 (8), 3438-3445.

148. Badhani, B., Sharma, N., Kakkar, R., 2015. Gallic acid: a versatile antioxidant with promising therapeutic and industrial applications. RSC Adv. 5, 27540–27557.

149. Ding, T., Yang, M., Zhang, J., Yang, B., Lin, K., Li, J., Gan, J., 2017. Toxicity, degradation and metabolic fate of ibuprofen on freshwater diatom Navicula sp. J. Hazard. Mater. 330, 127–134.

150. Nayeem, N., SMB, A., Salem, H., AHEI-Alfqy, S., 2016. Gallic acid: A Promising Lead Molecule for Drug Development. J. App. Pharm. 8(2), 1000213.

151. Nazeruddin, G.M., Osman, H.A., Shaikh, Y.I., Suryawanshi, S.B., 2014. QSAR/QSPR: Designing of new non-steroidal anti-inflammatory drugs (NSAIDs) considering ibuprofen as a lead compound followed by their synthesis and evaluation. Der. Pharm. Lett., 6(6), 135-141.

152. Boggara, M.B., Krishnamoorti, R., 2010. Partitioning of Nonsteroidal Antiinflammatory Drugs in Lipid Membranes: A Molecular Dynamics Simulation Study. Biophys. J. 98, 586–595.

153. Deruiter, J., 2005. Carboxylic Acid Structure and Chemistry: Part 2. Principles of Drug Action. 1, 1-10.

154. Cappelli, C., Mennucci, B., Monti, S., 2005. Environmental Effects on the Spectroscopic Properties of Gallic Acid: A Combined Classical and Quantum Mechanical Study. J. Phys. Chem. 109, 1933-1943.

155. Wheeler, S.E., 2013. Understanding Substituent Effects in Noncovalent Interactions Involving Aromatic Rings. Acc. Chem. Res. 46, 1029–1038.

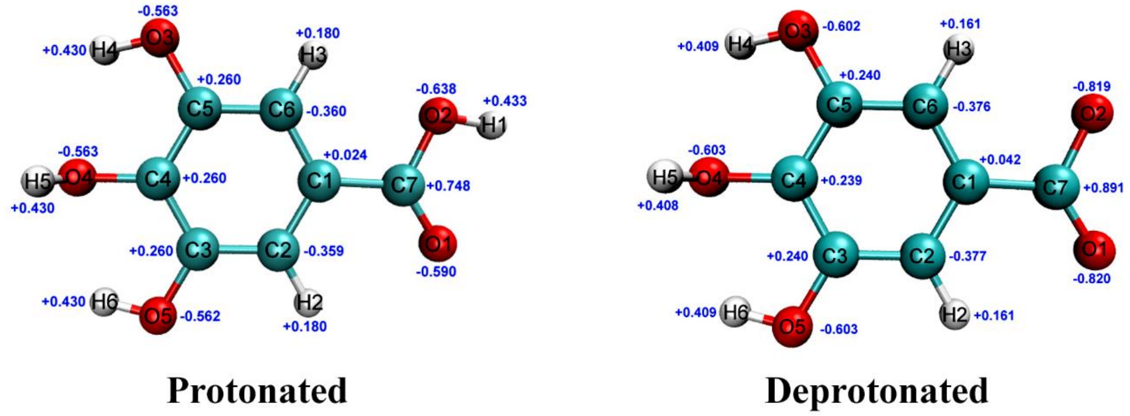
156. Thota, N., Hu, Z., Jiang, J., 2016. Ibuprofen loading and release in amphiphilic peptide FA32 and its derivatives: a coarse-grained molecular dynamics simulation study. Mol. Simul. 42, 679–687.Saxena, H.O., Faridi, U., Srivastava, S., Kumar, J.K.K., Darokar, M.P.P., Luqman, S., Chanotiya, C.S.S., Krishna, V., Negi, A.S., Khanuja, S.P.S.P.S., 2008. Gallic acid-based indanone derivatives as anticancer agents. Biorgan. Med. Chem. Lett. 18, 3914–3918.

157. Ye, Y.S., Wang, H., Bi, S.G., Xue, Y., Xue, Z.G., Liao, Y.G., Zhou, X.P., Xie, X.L., Mai, Y.W., 2015. Enhanced ion transport in polymer–ionic liquid electrolytes containing ionic liquid-functionalized nanostructured carbon materials. Carbon N. Y. 86, 86–97.

158. Ryu, H., Lee, H., Iwata, S., Choi, S., Ki Kim, M., Kim, Y.-R., Maruta, S., Min Kim, S., Jeon, T.-J., 2015. Investigation of Ion Channel Activities of Gramicidin A in the Presence of Ionic Liquids Using Model Cell Membranes. Sci. Rep. 5, 11935.
159. Fu, X., Sheng, X., Zhou, Y., Fu, Z., Zhao, S., Zhang, Z., Zhang, Y., 2016. One-step synthesis of hierarchical aluminosilicates using alkoxy-functionalized ionic liquid as a novel template. New J. Chem. 40, 6036–6045.
160. Spigno, G., Dermiki, M., Pastori, C., Casanova, F., Jauregi, P., 2010. Recovery of gallic acid with colloidal gas aphrons generated from a cationic surfactant. Sep. Purif. Technol. 71, 56–62.Puoci, F., Scoma, A., Cirillo, G., Bertin, L., Fava, F., Picci, N., 2012. Selective extraction and purification of gallic acid from actual site olive mill wastewaters by means of molecularly imprinted microparticles. Chem. Eng. J. 198–199, 529–535.
161. Neves, C.M.S.S., de Cássia S. Sousa, R., Pereira, M.M., Freire, M.G., Coutinho, J.A.P., 2019. Understanding the effect of ionic liquids as adjuvants in the partition of biomolecules in aqueous two-phase systems formed by polymers and weak salting-out agents. Biochem. Eng. J. 141, 239–246.
162. Hofmeister, F., 1888. Zur Lehre von der Wirkung der Salze [Translation: about the science of the effect of salts]. Arch. Exp. Pathol. Pharmakol., 24, 247-260.
163. Zakharova, L.Y., Pashirova, T.N., Fernandes, A.R., Doktorovova, S., Martins-Gomes, C., Silva, A.M., Souto, E.B., 2018. Self-assembled quaternary ammonium surfactants for pharmaceuticals and biotechnology. Org. Mater. as Smart Nanocarriers Drug Deliv. 601–618.
164. Melo, C.I., Bogel-Łukasik, R., Nunes da Ponte, M., Bogel-Łukasik, E., 2013. Ammonium ionic liquids as green solvents for drugs. Fluid Phase Equilib. 338, 209–216.
165. Egorova, K.S., Ananikov, V.P., 2018. Fundamental importance of ionic interactions in the liquid phase: A review of recent studies of ionic liquids in biomedical and pharmaceutical applications. J. Mol. Liq. 272, 271–300.
166. Halayqa, M., Zawadzki, M., Domańska, U., Plichta, A., 2019. Polymer – Ionic liquid – Pharmaceutical conjugates as drug delivery systems. J. Mol. Struct. 1180, 573–584.
167. Zavgorodnya, O., Shamshina, J.L., Mittenthal, M., McCrary, P.D., Rachiero, G.P., Titi, H.M., Rogers, R.D., 2017. Polyethylene glycol derivatization of the non-active ion in active pharmaceutical ingredient ionic liquids enhances transdermal delivery. New J. Chem. 41, 1499–1508.
168. Radak, B.K., Chipot, C., Suh, D., Jo, S., Jiang, W., Phillips, J.C., Schulten, K., Roux, B., 2017. Constant-pH Molecular Dynamics Simulations for Large Biomolecular Systems. J. Chem. Theory Comput. 13, 5933–5944.

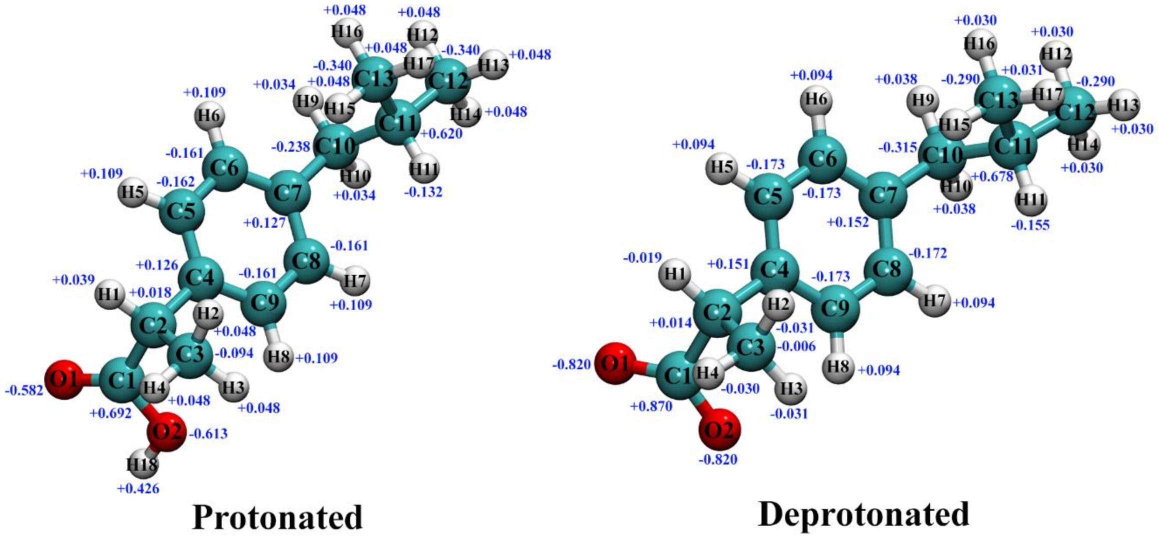
Supporting Information

Supporting Information


H4 +0.430
O3 -0.563
C5 +0.260
C6 -0.360
H3 +0.180
O2 -0.638
H1 +0.433
C1 +0.024
C7 +0.748
O1 -0.590
O4 -0.563
H5 +0.430
C4 +0.260
C3 +0.260
C2 -0.359
H6 +0.430
O5 -0.562
H2 +0.180
H4 +0.409
O3 -0.602
C5 +0.240
C6 -0.376
H3 +0.161
O2 -0.819
C1 +0.042
C7 +0.891
O1 -0.820
O4 -0.603
H5 +0.408
C4 +0.239
C3 +0.240
C2 -0.377
H6 +0.409
O5 -0.603
H2 +0.161

Protonated **Deprotonated**

Figure SI:1 – AA descriptions of the GA, with the protonated and deprotonated versions of the carboxylic acid group. The atom labelling is the one used in both the structure file and the topology information. Partial charges are also presented in blue, aside the respective atom.


H16 +0.048
H12 +0.048
C13 -0.340
H17 +0.048
C12 -0.340
H13 +0.048
H9 +0.034
H15 +0.048
C11 +0.620
H14 +0.048
C10 -0.238
H10 +0.034
H11 -0.132
H6 +0.109
C6 -0.161
H5 +0.109
C5 -0.162
C7 +0.127
C8 -0.161
H7 +0.109
C4 +0.126
C9 -0.161
H1 +0.039
C2 +0.018
H2 +0.048
C3 -0.094
H8 +0.109
O1 -0.582
C1 +0.692
H4 +0.048
H3 +0.048
O2 -0.613
H18 +0.426
H16 +0.030
H12 +0.030
C13 -0.290
H17 +0.031
C12 -0.290
H13 +0.030
H9 +0.038
H15
C10 -0.315
C11 +0.678
H14 +0.030
H10 +0.038
H11 -0.155
H6 +0.094
C6 -0.173
H5 +0.094
C5 -0.173
C7 +0.152
C8 -0.172
H7 +0.094
C4 +0.151
C9 -0.173
H1 -0.019
C2 +0.014
H2 -0.031
C3 -0.006
H8 +0.094
O1 -0.820
C1 +0.870
H4 -0.030
H3 -0.031
O2 -0.820

Protonated **Deprotonated**

Figure SI:2 – AA structures of the IBP, both the protonated and deprotonated version of the carboxylic acid group. The partial charges of each atom are indicated in blue. Atom labelling is the one used throughout the whole simulation and analysis procedures.

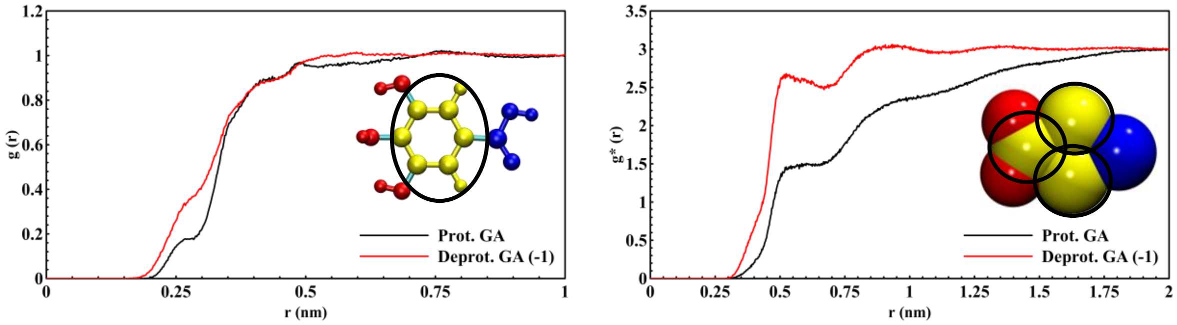


Figure SI:3 – RDF profiles of AA (left) and CG (right) GA (systems 1-4), using the aromatic ring as the reference group and water (PW) as the selection.

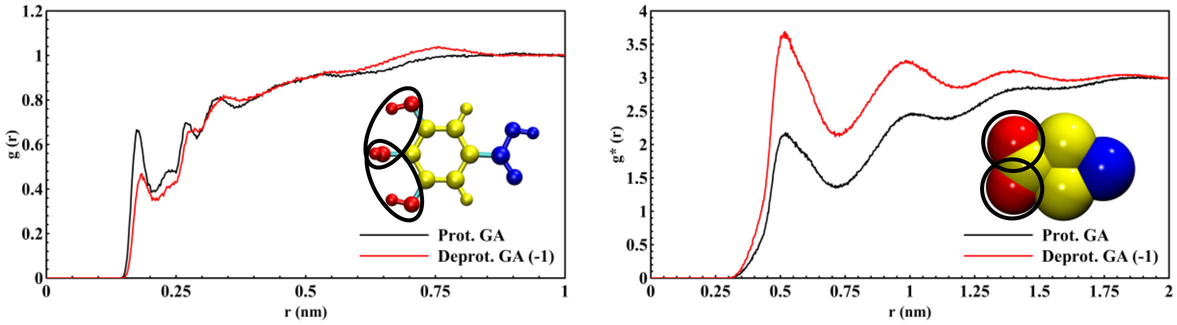


Figure SI:4 – RDF profiles of AA (left) and CG (right) GA (systems 1-4), using the hydroxylic groups as a reference and water (PW) as the selection.

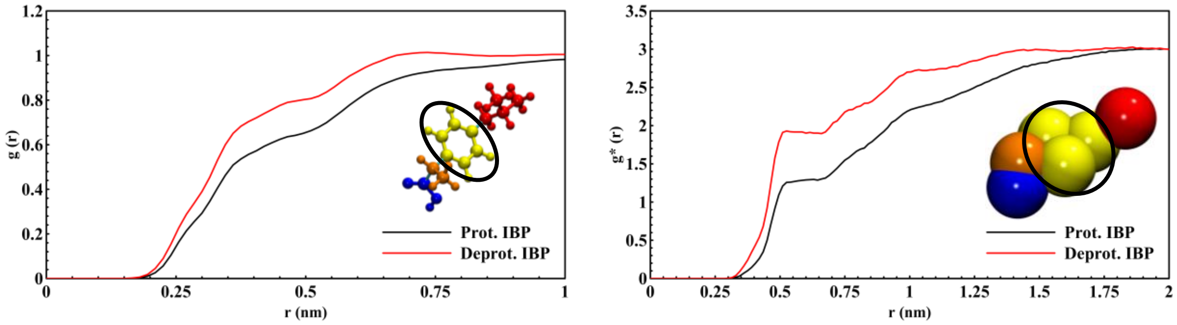


Figure SI:5 – RDF profiles of AA (left) and CG (right) IBP (systems 5-8), using the aromatic ring as the reference and water (PW) as the selection.

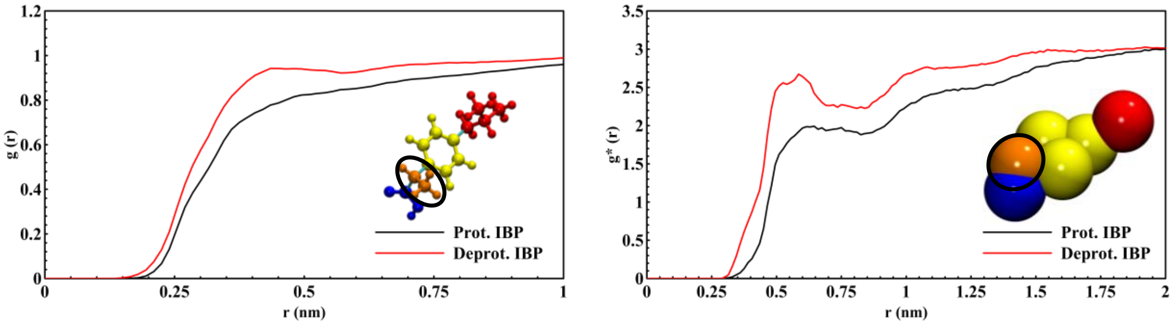


Figure SI:6 – RDF profiles of AA (left) and CG (right) IBP (systems 5-8), using the alkyl region between the carboxylic acid group and the aromatic ring as the reference and water/PW as selection.

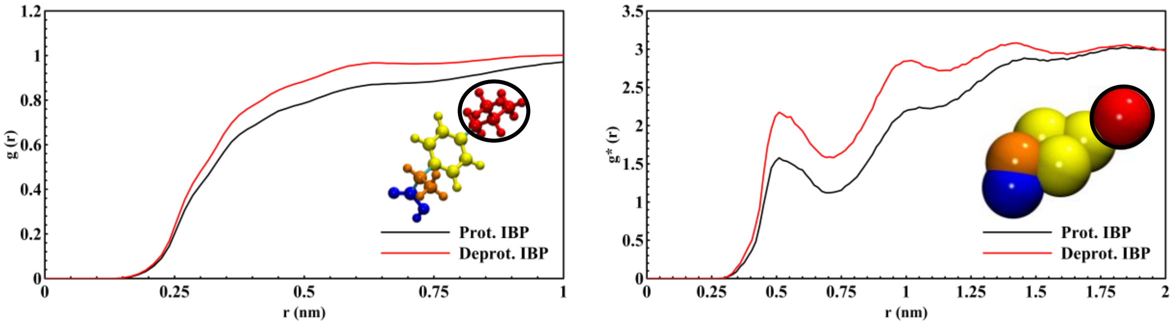


Figure SI:7 – RDF profiles of AA (left) and CG (right) IBP (systems 5-8), using the alkyl group farthest from the carboxylic acid group as a reference and water (PW) as the selection.

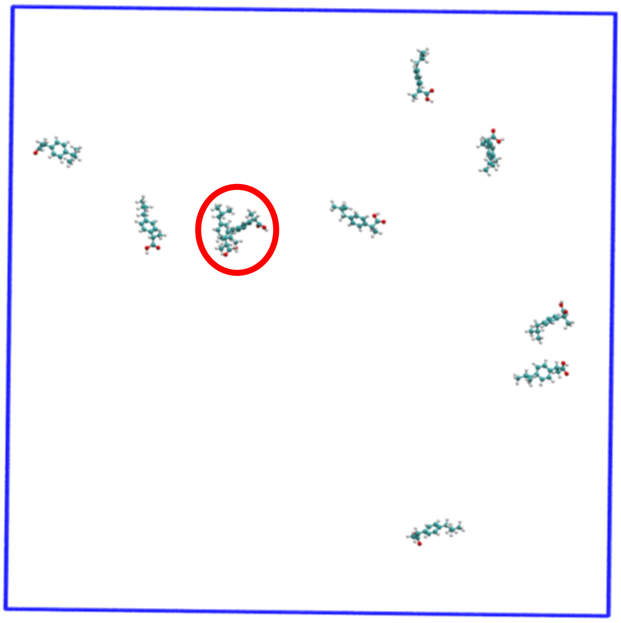


Figure SI:8 – Final snapshot of the protonated AA IBP in water (system 5). Water was removed for clarity. The two aggregated IBP molecules are highlighted in red.

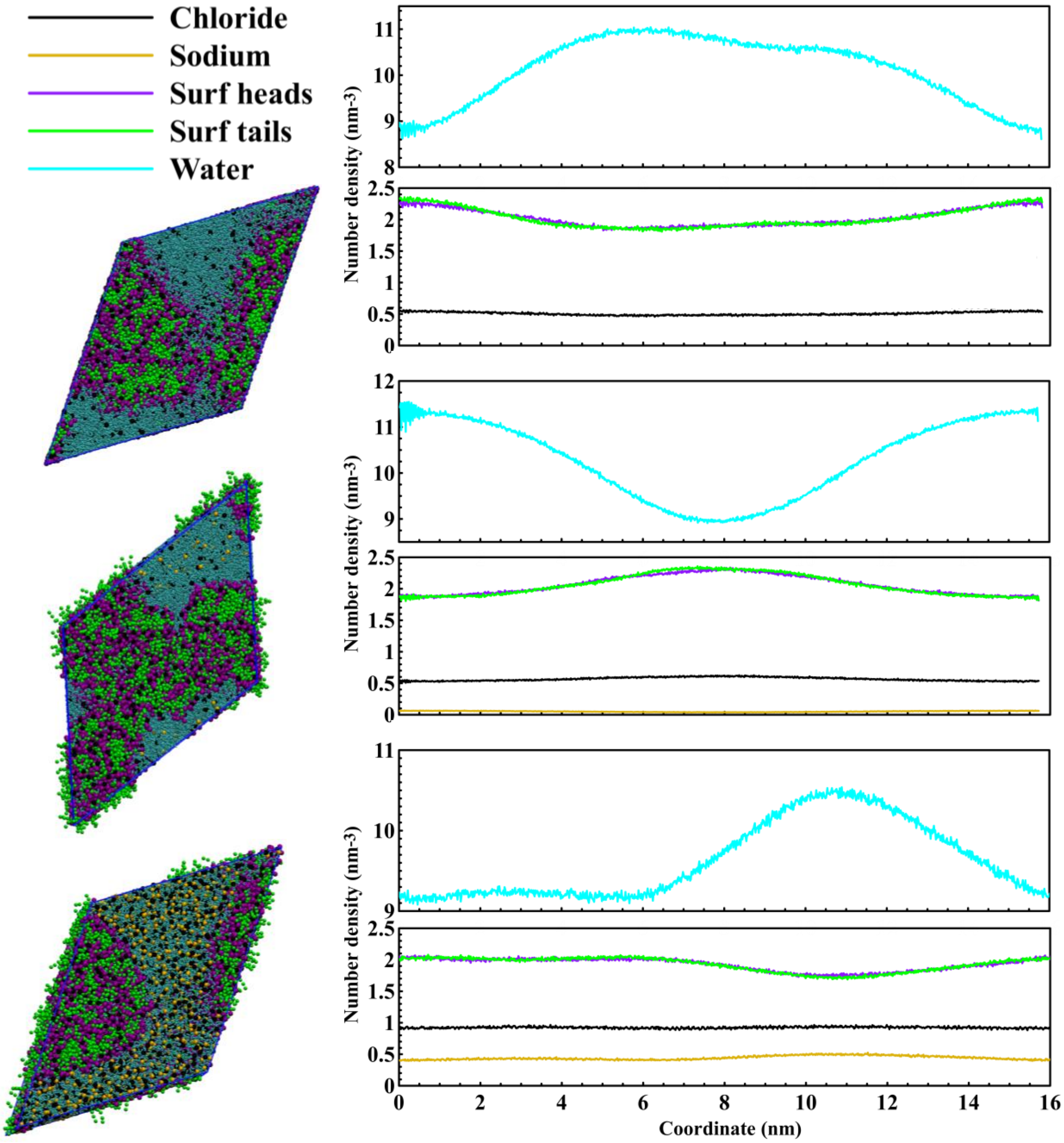


Figure SI:9 – Density profiles of the [N$_{44414}$]Cl system before the addition of salt (top), with NaCl at 0.64 wt% (center) and at 4.72 wt% (bottom) (systems 12-14). The SAIL polar heads are depicted in purple and the alkyl tails in green. Chloride of both the SAIL anion and NaCl is presented in black. Sodium is represented in dark yellow and water in cyan. In each block, water density is represented in the top segment while the rest of the components is depicted in the bottom one, for clarity. The profile was generated along the Z axis of the simulation box. Snapshots of a side view of the simulation box are presented on the left side of the figure, following the same order and progress along the axis as the density profiles.

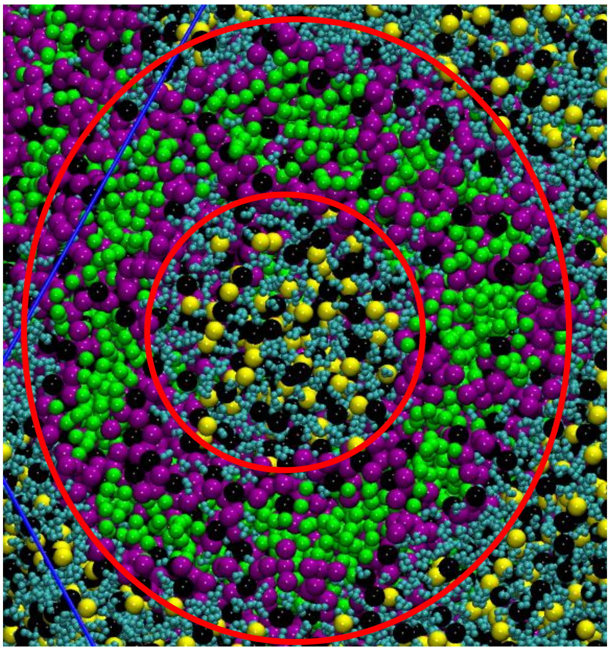


Figure SI:10 – Close-up of the ion channel formed in the system with [N$_{44414}$]Cl and 4.72 wt.% of NaCl (system 14). Cation polar heads are presented in purple, alkyl tails in green, chloride ions in black (both the IL anion and salt component), sodium in yellow and water in cyan. The PW bead size was decreased for increased clarity. The inner red circumference denotes the inner wall of the channel while the outer one highlights the external side of the channel.

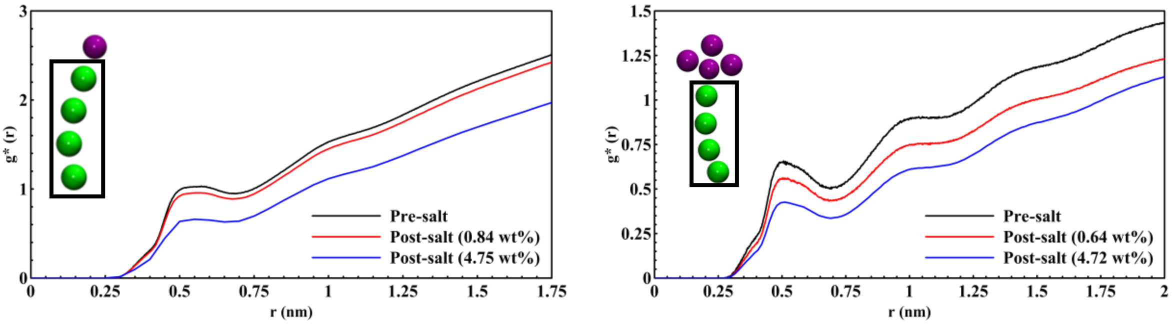


Figure SI:11 – RDF profiles of CG [N$_{11114}$]$^+$ (left) and [N$_{44414}$]$^+$ (right) (systems 9-14), using the alkyl tail of the cation as the reference and the PW as the selection.

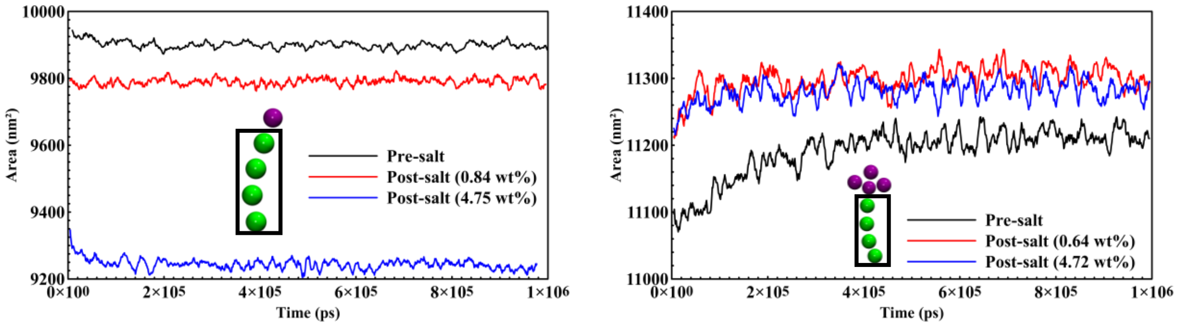


Figure SI:12 – SASA profiles for the alkyl tail of the CG $[N_{11114}]^+$ (left) and $[N_{44414}]^+$ (right) (systems 9-14). A moving average was applied to smooth the function.

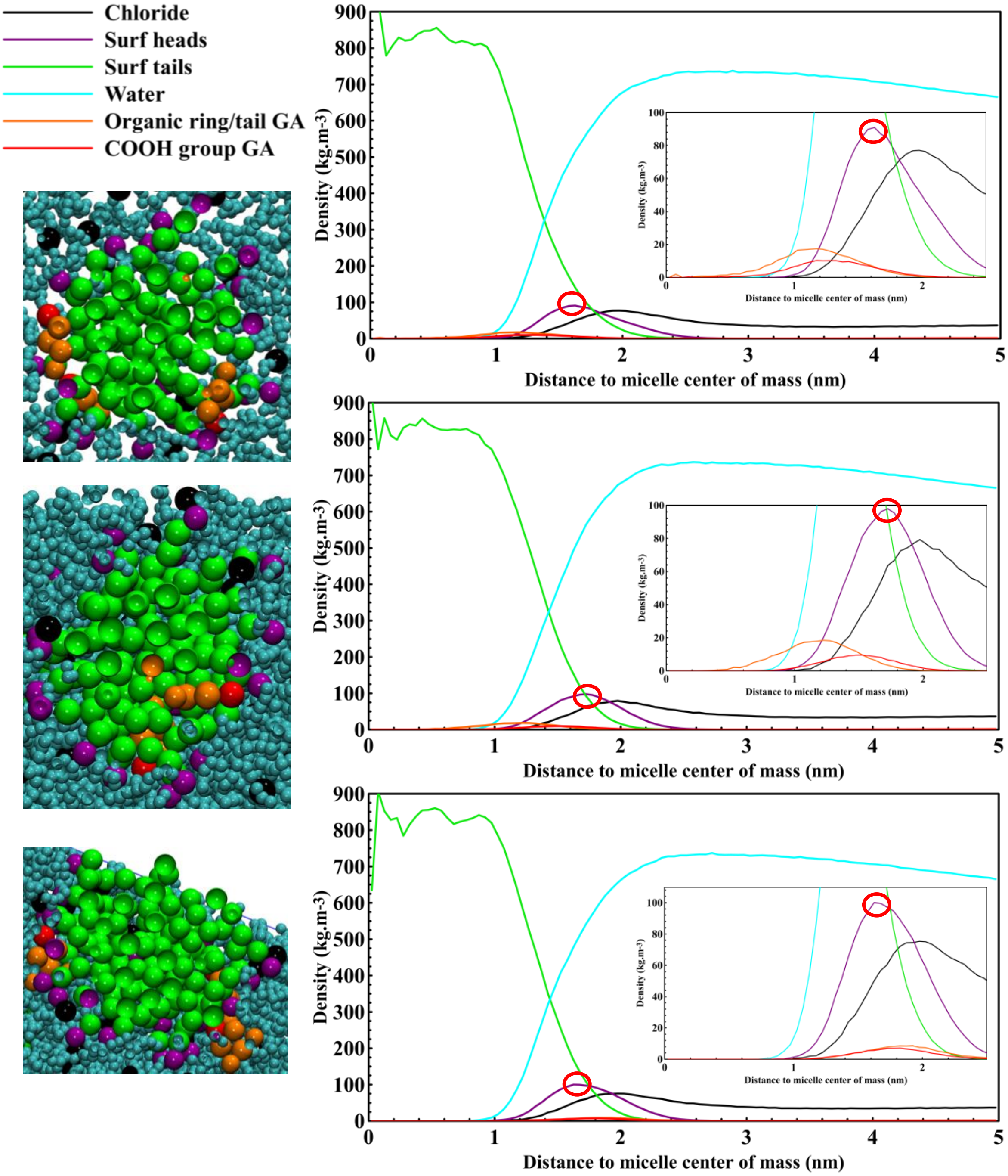


Figure SI:13 – Density profiles for systems 16-18 containing GA and [N_{11114}]Cl, using the COM of IL micelles as the origin and the density of each component of the system in relation to it. A close-up of the area corresponding to the micelle surface (red circle) and GA orientation in it is depicted inside each general profile. The aromatic ring and hydroxyl beads were appended into a single group due to bead overlapping interfering with the employed code.

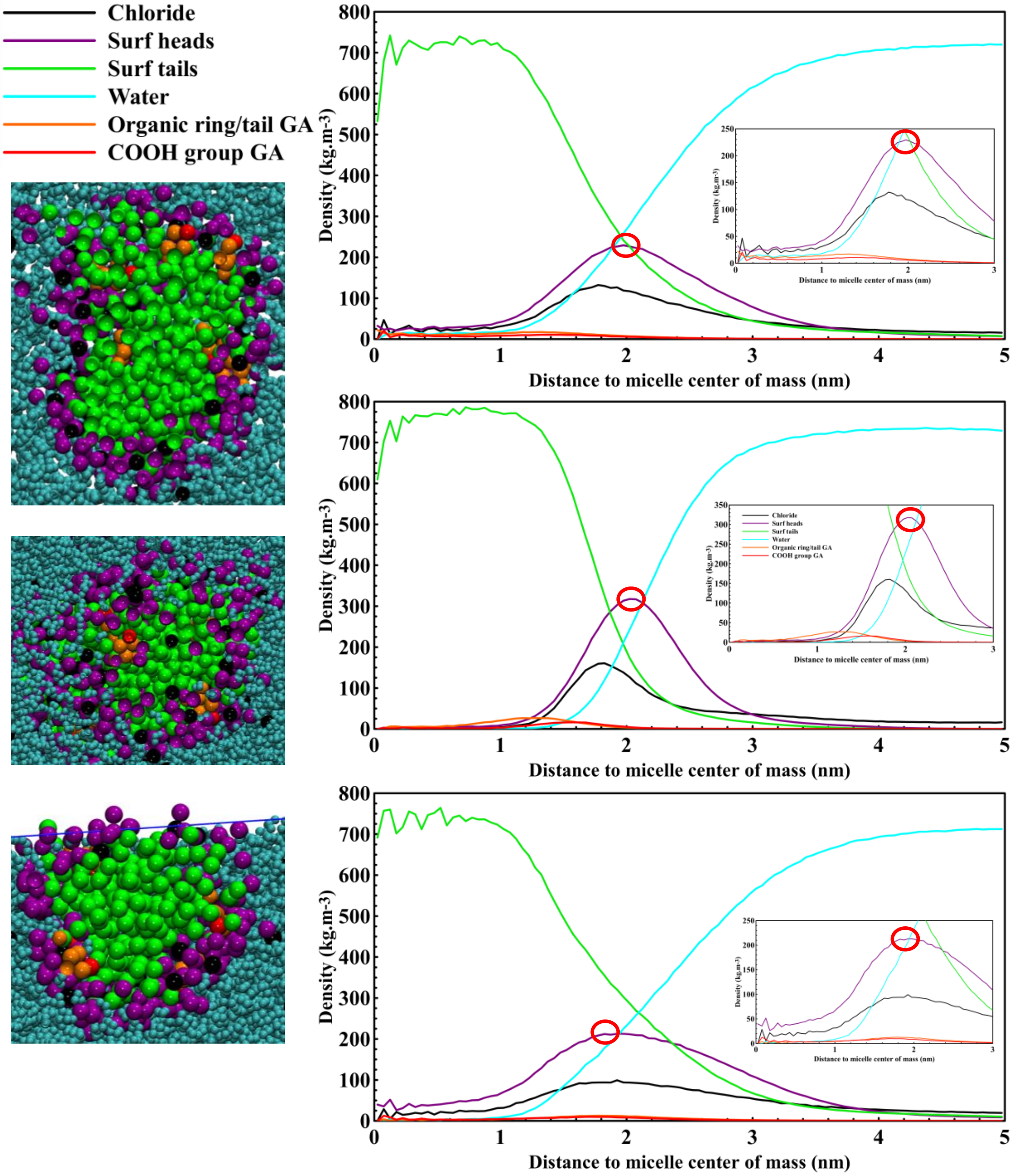


Figure SI:14 - Density profiles for systems 23-25 containing GA and [N$_{44414}$]Cl, using the COM of IL micelles as the origin and the density of each component of the system in relation to it. A close-up of the area corresponding to the micelle surface (red circle) and GA orientation in it is depicted inside each general profile. The aromatic ring and hydroxyl beads were appended into a single group due to bead overlapping interfering with the employed code.

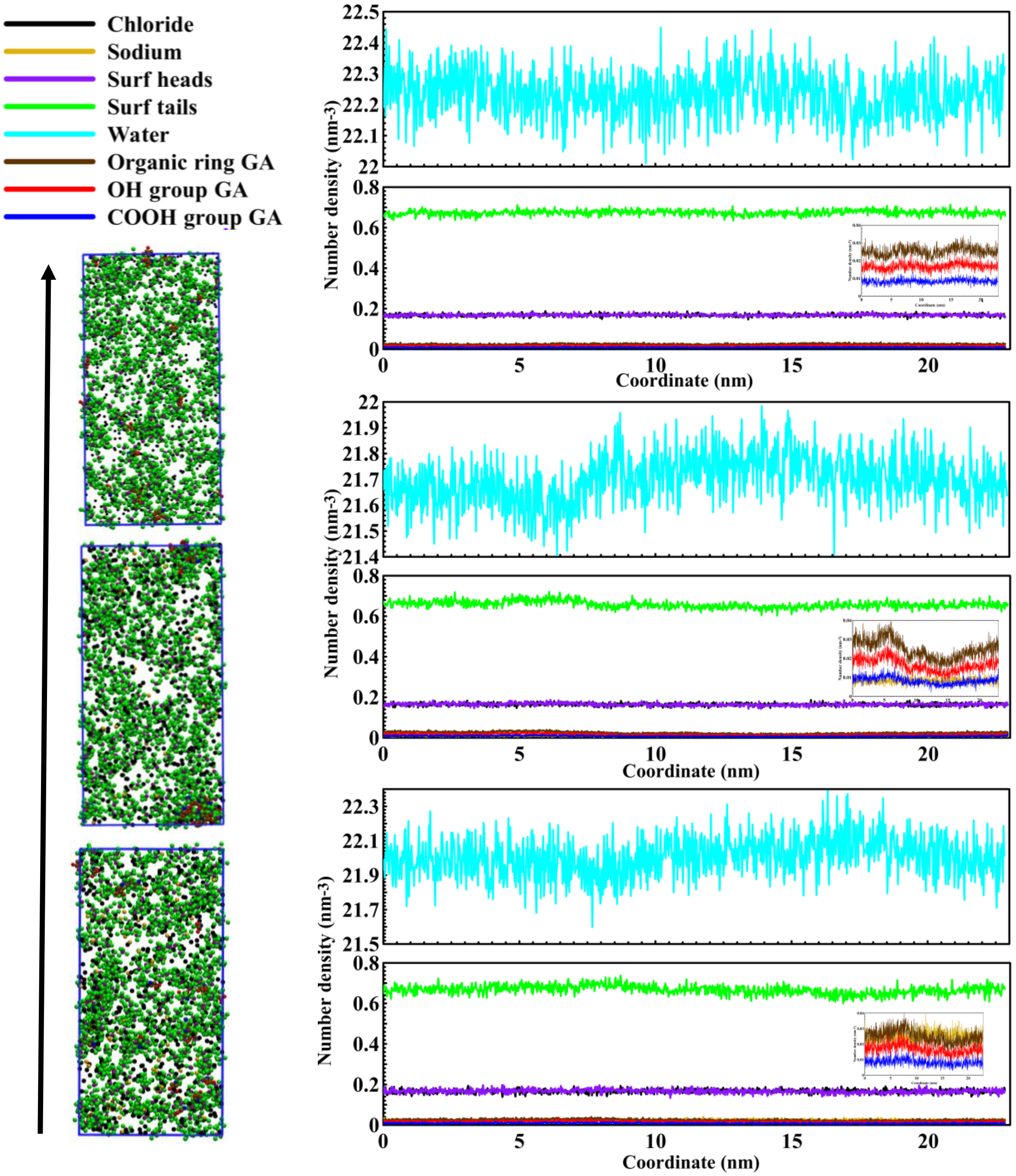


Figure SI:15 - Density profiles for systems 30-32 containing GA and [N_{4444}]Cl, using full simulation box (left side) through the *z* axis and presenting the density of each component at every point in the box. A close-up of the density corresponding to the different GA groups is depicted inside each general profile. Water was removed from the box snapshots on the left side for increased clarity.

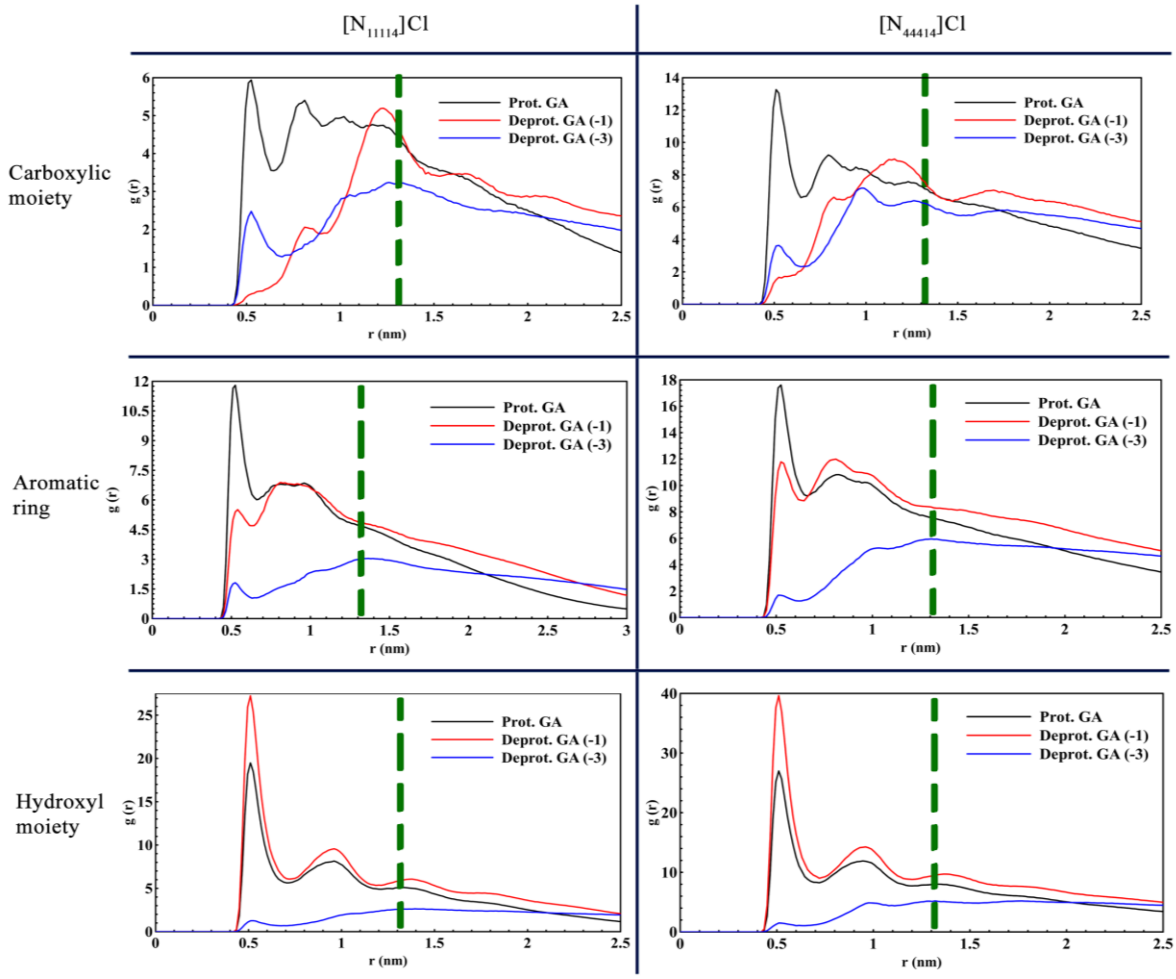


Figure SI:16 – RDF profiles using the Ct3 bead of the IL cation alkyl chain as the reference and each GA group as the selection (systems 19-21, 26-28, 33-35). A green dashed line indicates the distance *r* corresponding to the IL micelle surface.

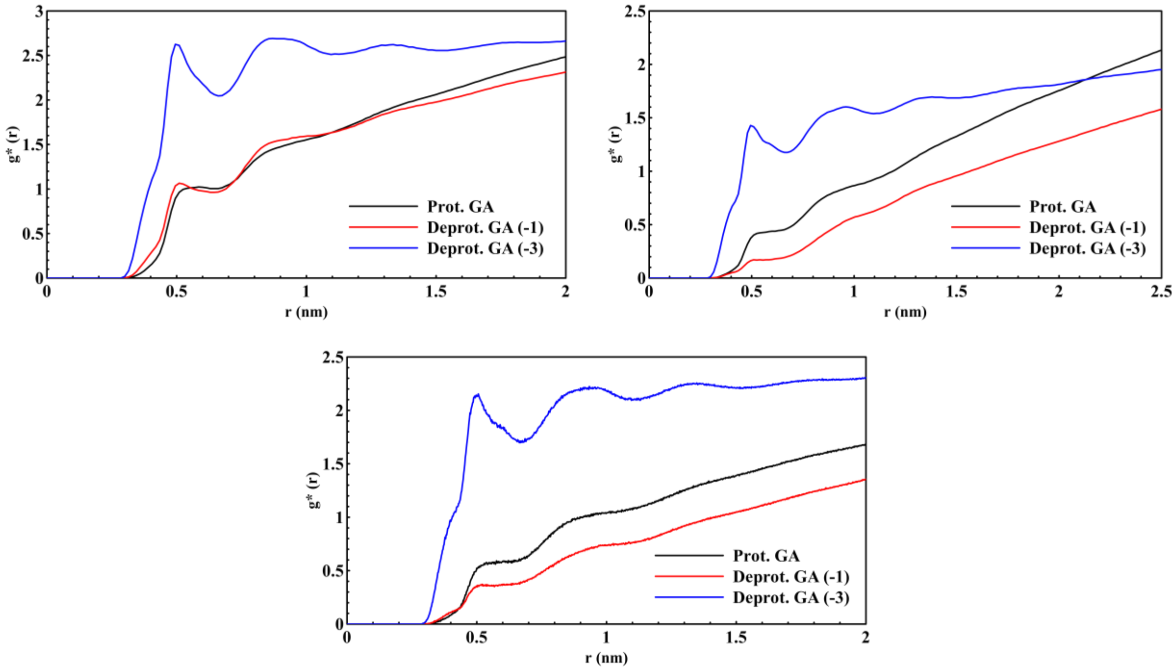


Figure SI:17 - RDF profiles using the aromatic ring of GA as the reference and PW as the selection, in systems with $[N_{11114}]$Cl (top left), $[N_{44414}]$Cl (top right) and $[N_{4444}]$Cl (bottom) (systems 19-21, 26-28, 33-35).

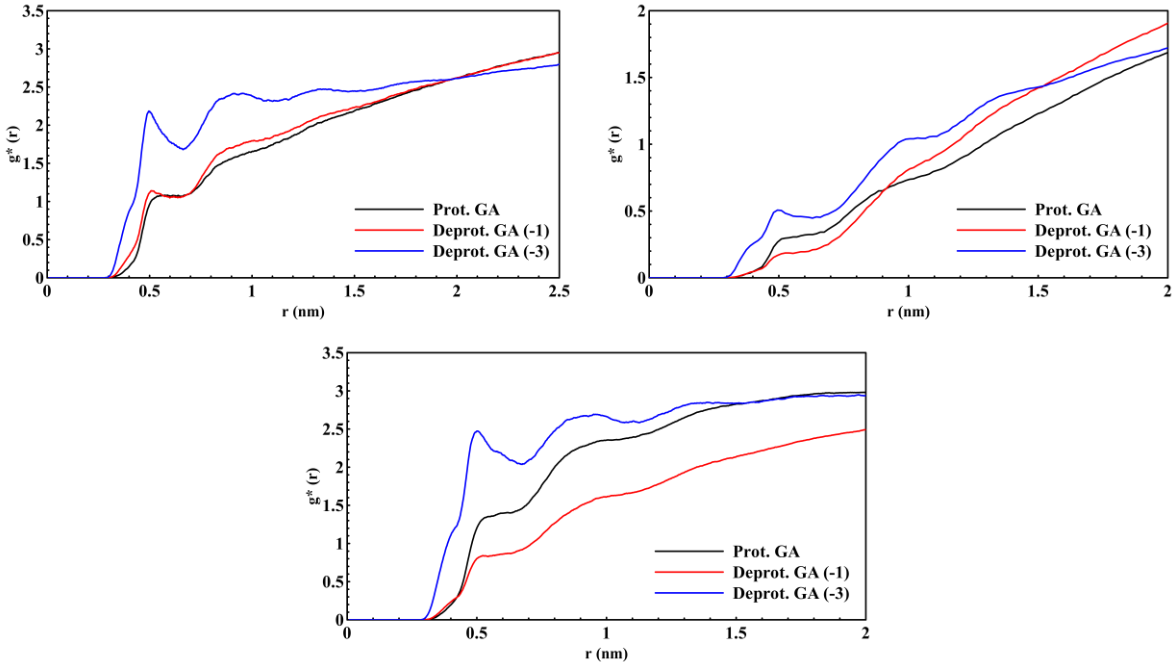


Figure SI:18 – RDF profile of the aromatic ring of GA as the reference and PW as the selection, in systems containing $[N_{11114}]$Cl (top left), $[N_{44414}]$Cl (top right) and $[N_{4444}]$Cl (bottom) (systems 16-18, 23-25, 30-32).

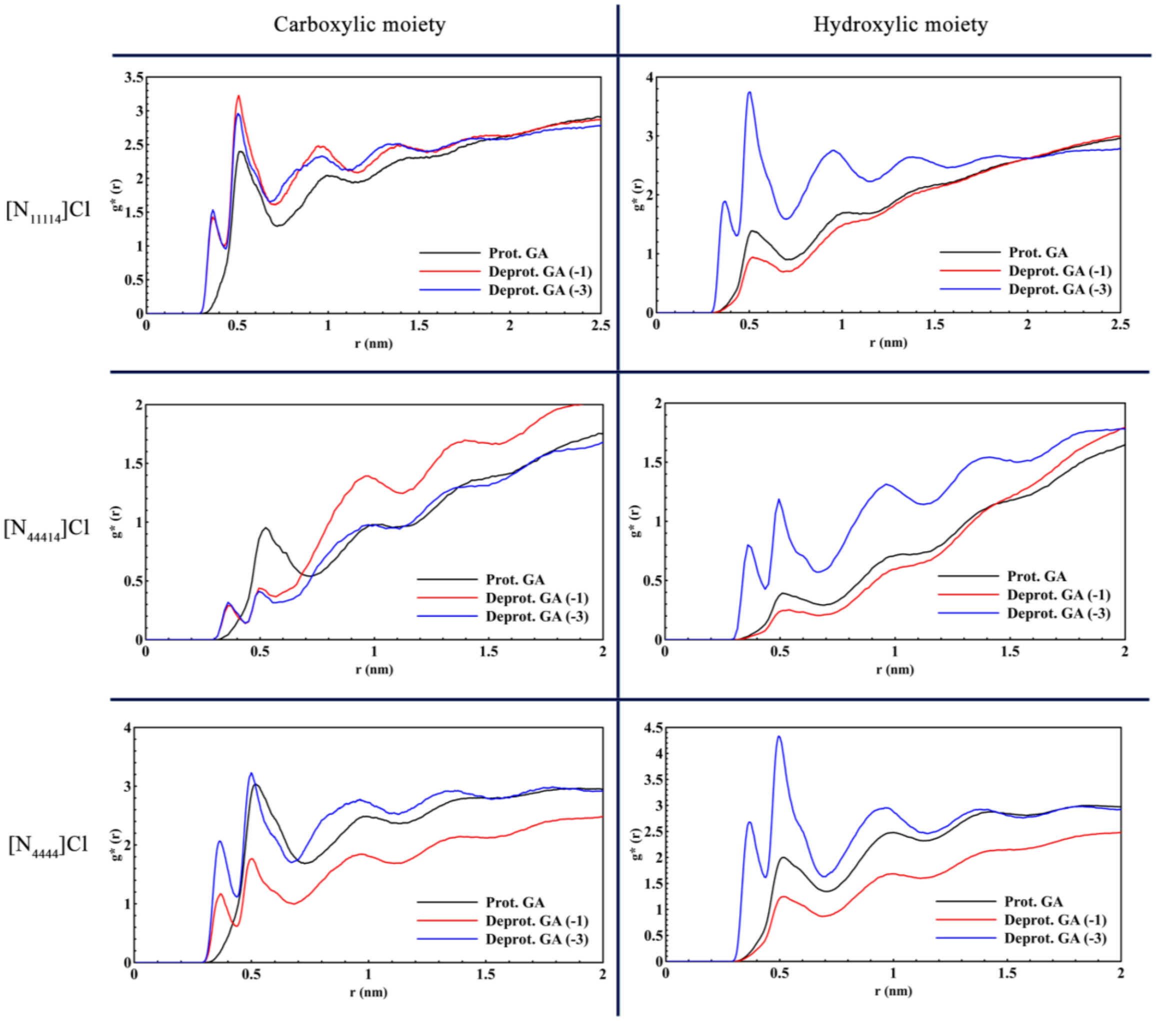


Figure SI:19 – RDF profiles using either the carboxylic or the hydroxylic moieties of GA as the reference and PW as the selection (systems 16-18, 23-25, 30-32).

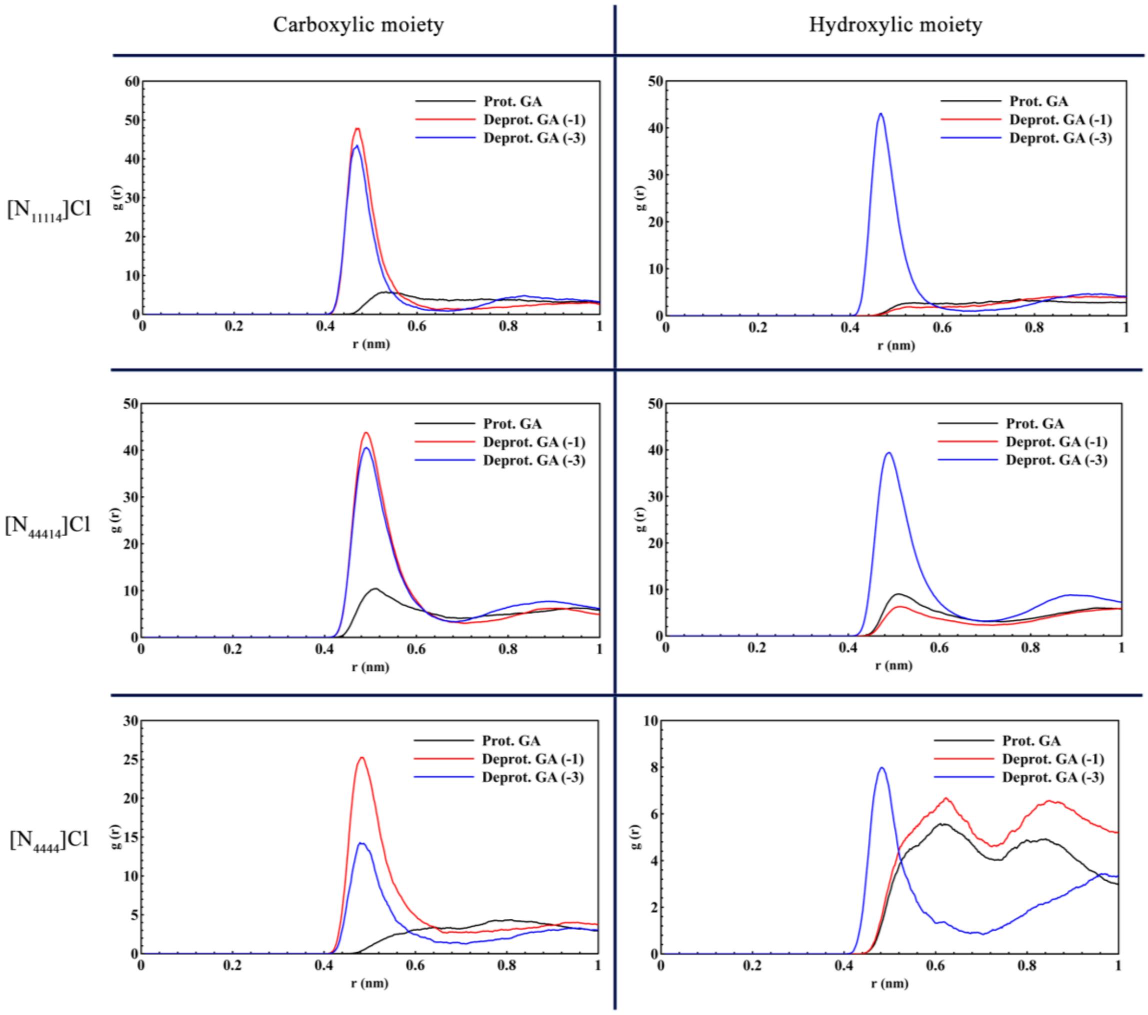


Figure SI:20 – RDF profiles using the carboxylic or hydroxylic moieties of GA as the reference and the polar heads of the IL cation as the selection (systems 16-18, 23-25, 30-32).

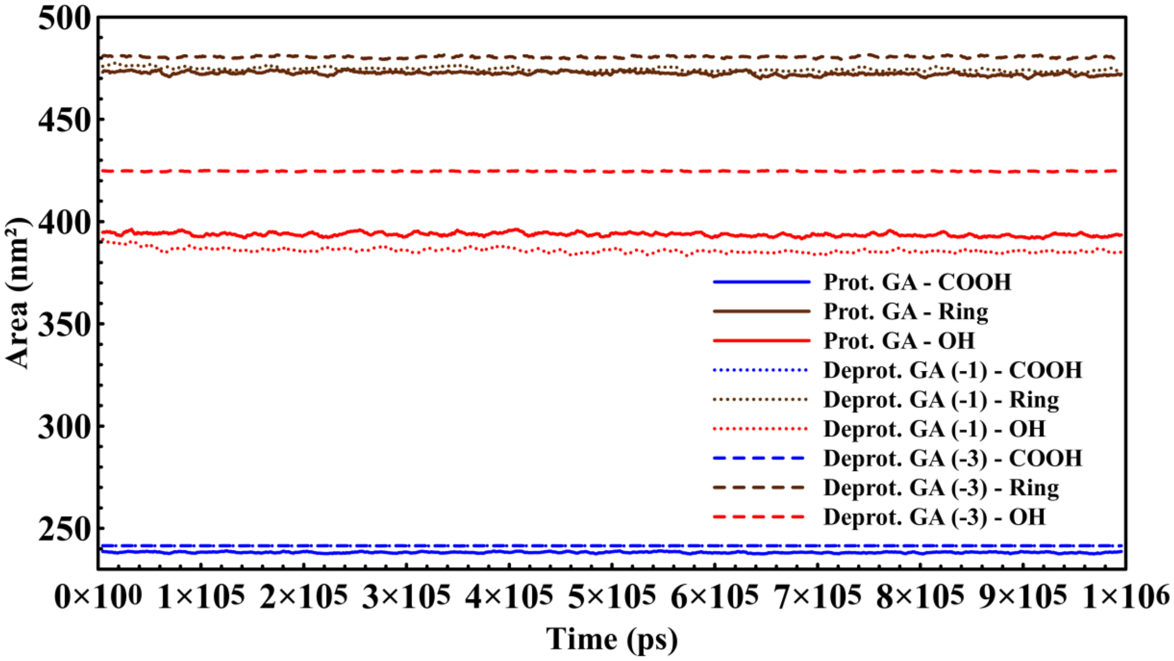


Figure SI:21 – SASA profile summarizing the area available for each GA group and protonation state, in systems with [N_{11114}]Cl (systems 19-21).

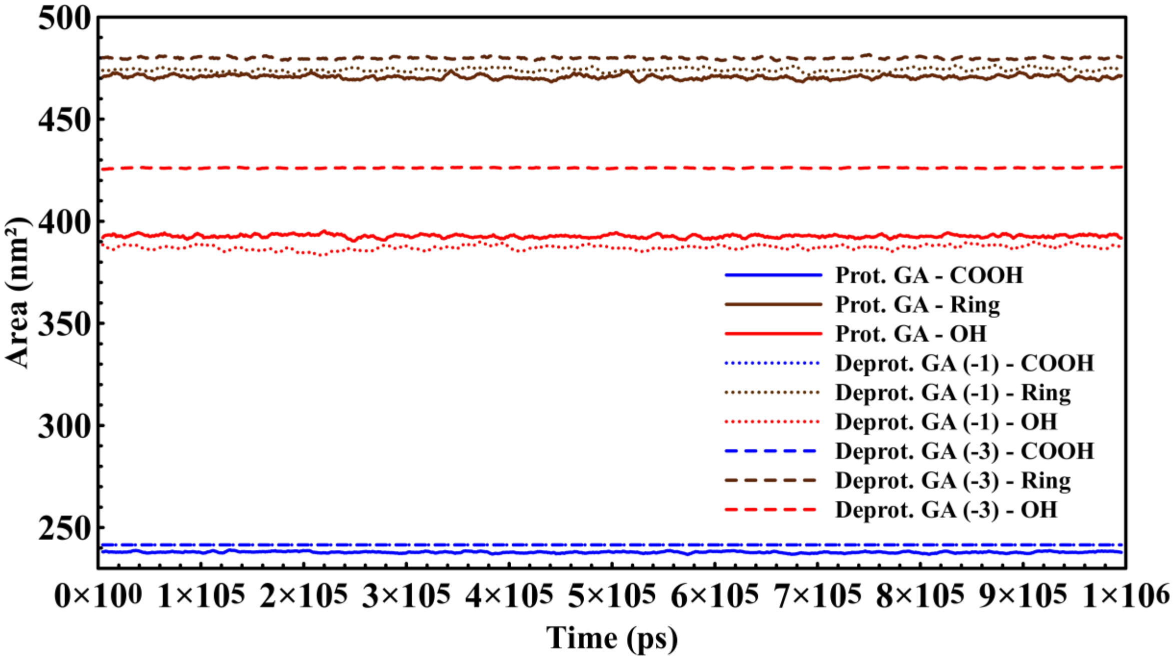


Figure SI:22 – SASA profile summarizing the area available for each GA group and protonation state, in systems with [N_{44414}]Cl (systems 26-28).

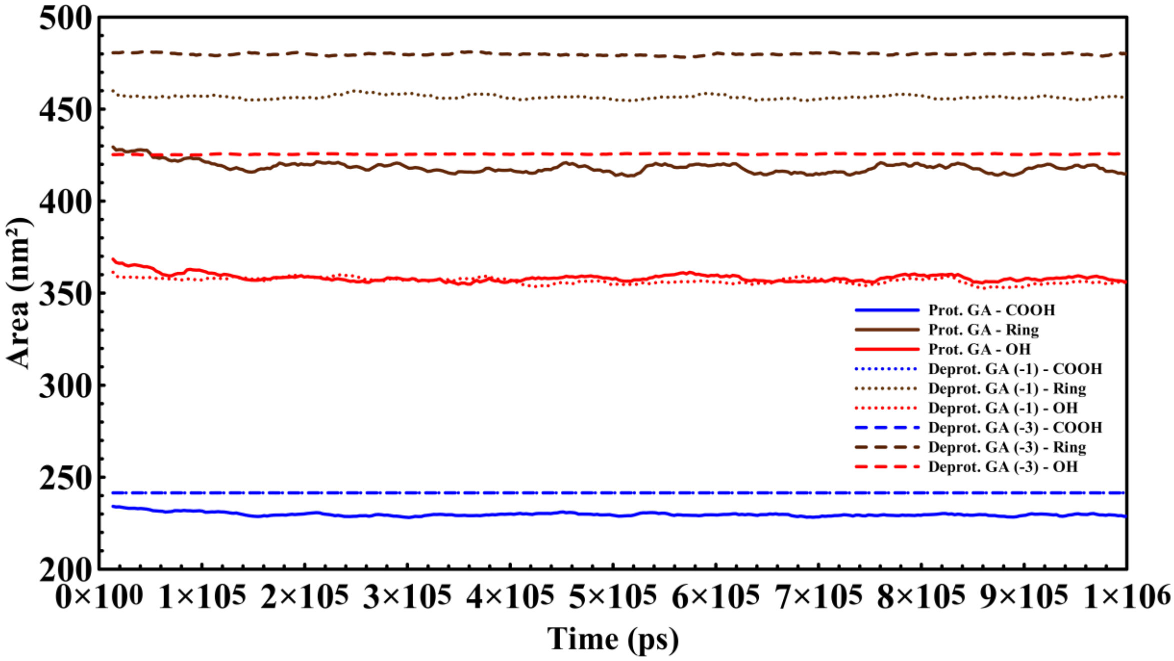


Figure SI:23 – SASA profile summarizing the area available for each GA group and protonation state, in systems with [N_{4444}]Cl (systems 33-35).

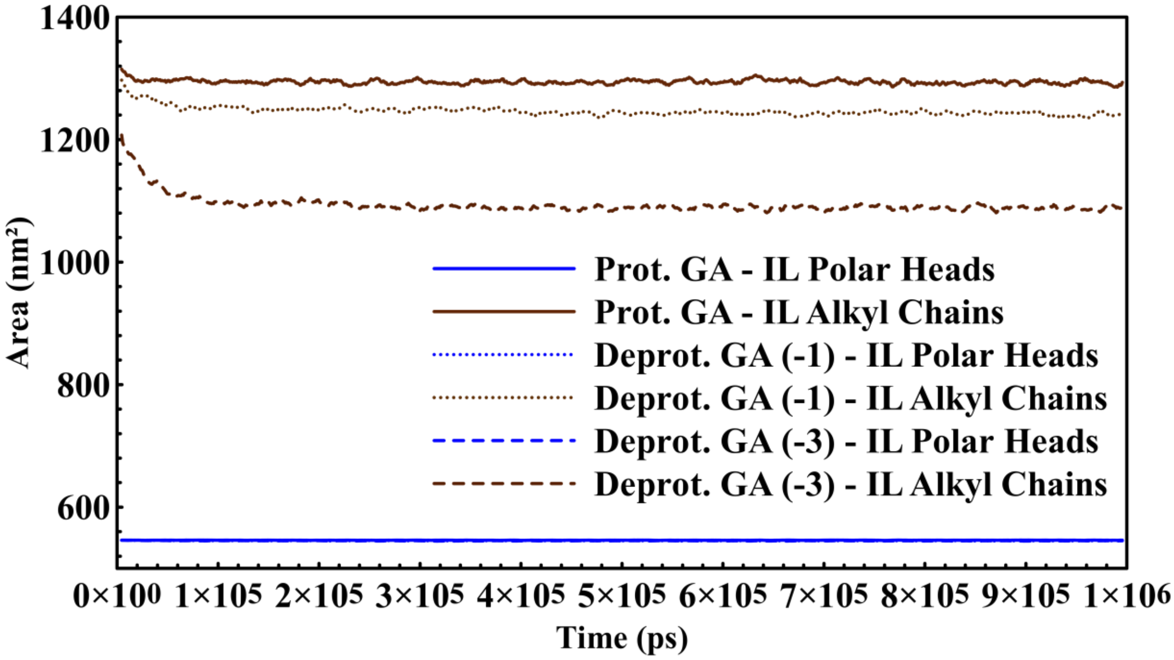


Figure SI:24 – SASA profile summarizing the area available for each [N_{11114}]$^+$ group, in the various systems with GA (systems 19-21).

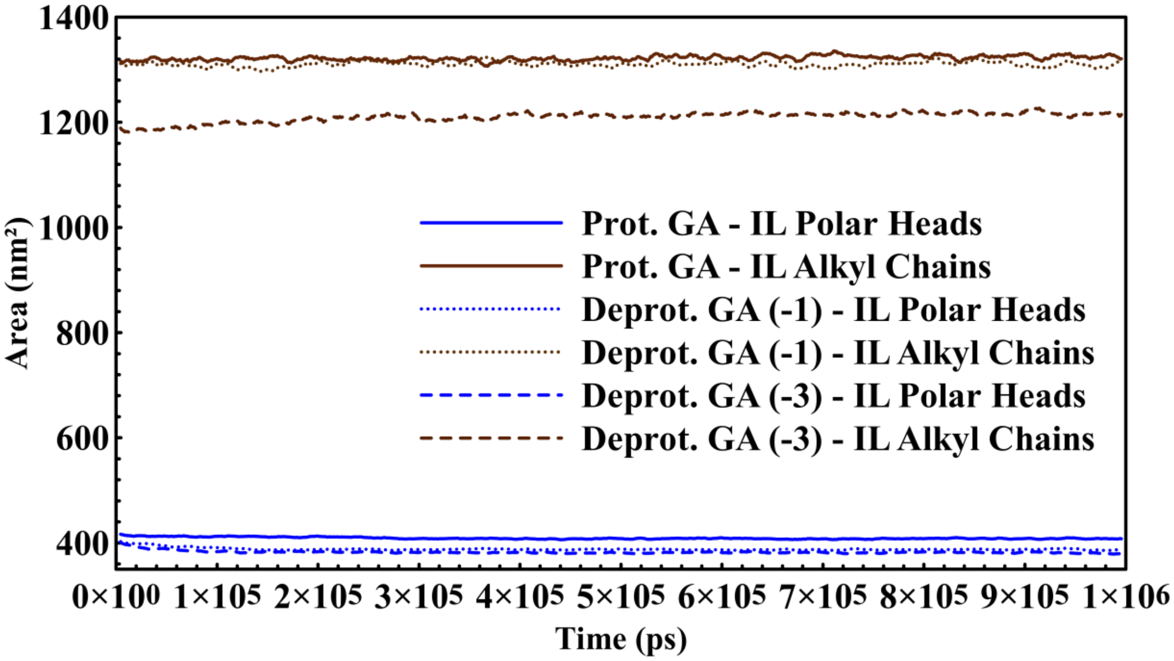


Figure SI:25 – SASA profile summarizing the area available for each $[N_{44414}]^+$ group, in the various systems with GA (systems 26-28).

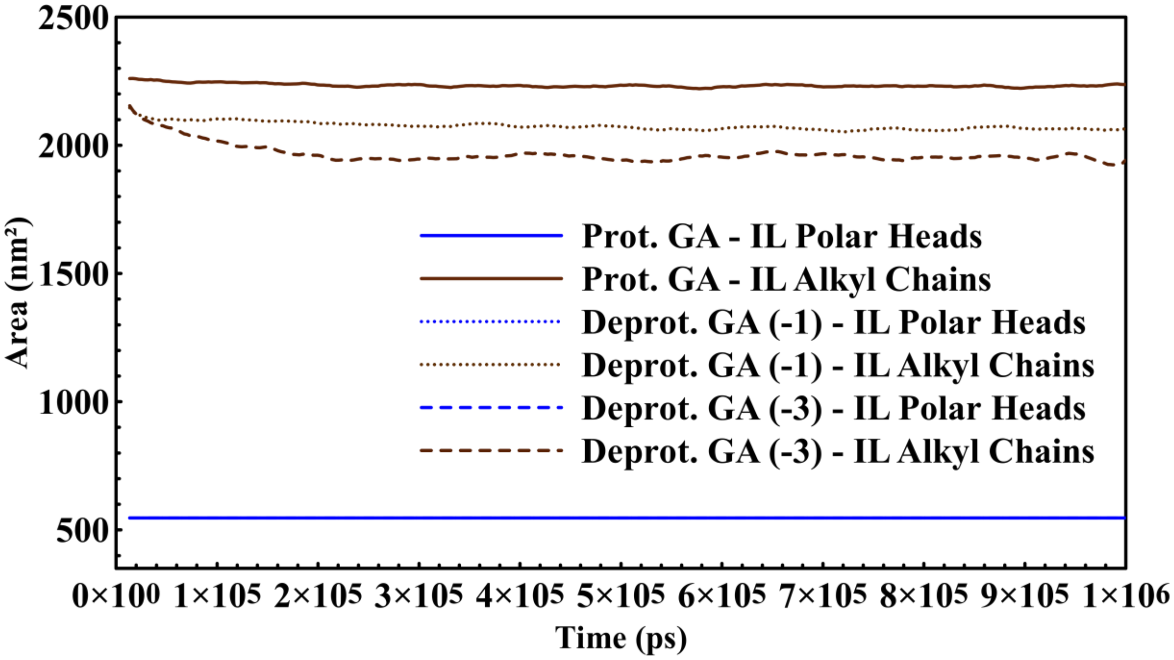


Figure SI:26 – SASA profile summarizing the area available for each $[N_{4444}]^+$ group, in the various systems with GA (systems 33-35).

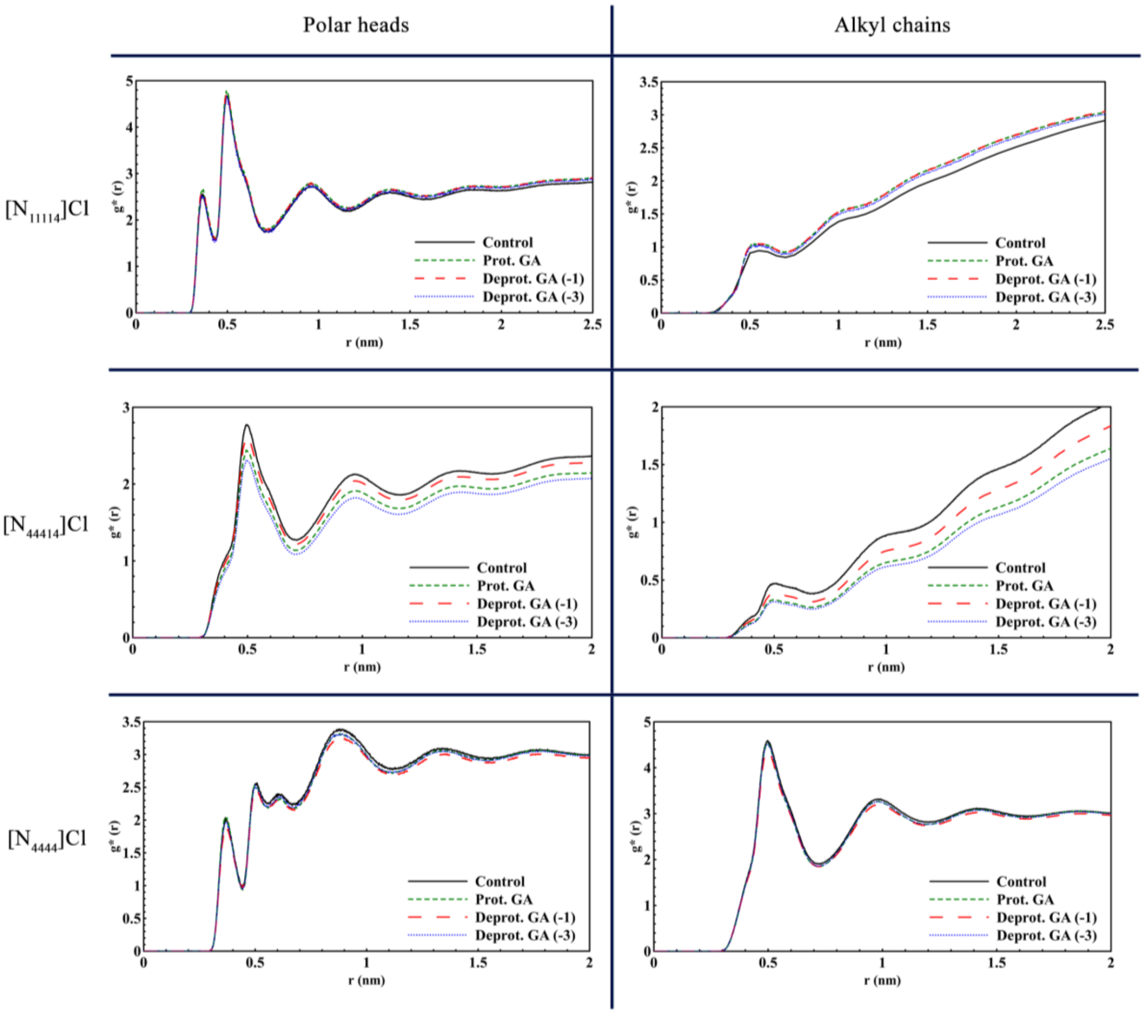


Figure SI:27 – RDF profiles using the polar heads or the alkyl chains of each IL cation as the reference and PW as the selection (systems 16-18, 23-25, 30-32).