



**HAL**  
open science

# Investigation of interface phenomena in formamide–water systems by computer simulation methods

Bálint Gábor Kiss

► **To cite this version:**

Bálint Gábor Kiss. Investigation of interface phenomena in formamide–water systems by computer simulation methods. Analytical chemistry. Université de Lille; Miskolci Egyetem (Hongrie), 2021. English. NNT: 2021LILUR005 . tel-03385092

**HAL Id: tel-03385092**

**<https://theses.hal.science/tel-03385092>**

Submitted on 19 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Université de Lille**

Faculté des Sciences et Technologies

Ecole Doctorale 104 - Sciences de la Matière, du Rayonnement, et de l'Environnement

et

# **Université de Miskolc**

Faculté des Sciences de la Matière et Ingénierie

## **Thèse de Doctorat en Cotutelle**

par

**Bálint Gábor KISS**

pour l'obtention du titre de

**Docteur de l'Université de Lille**

Discipline: Chimie théorique, physique, analytique

Analyse des phénomènes à l'interface dans les systèmes formamide-eau par  
simulations de dynamique moléculaire

Thèse présentée et soutenue à Miskolc, le 28/01/2021 devant le jury:

|                          |   |                              |
|--------------------------|---|------------------------------|
| <b>Péter MIZSEY</b>      |   | <b>Président</b>             |
|                          | Professeur, Université de Miskolc                   |                              |
| <b>Dezső BODA</b>        |   | <b>Rapporteur</b>            |
|                          | Professeur, Université de Pannonie                  |                              |
| <b>Michael KISELEV</b>   |   | <b>Rapporteur</b>            |
|                          | Professeur, Académie des sciences de Russie         |                              |
| <b>Szilárd FEJÉR</b>     |   | <b>Examineur</b>             |
|                          | Professeur, Universités de Pécs                     |                              |
| <b>Sophie FOURMENTIN</b> |   | <b>Examineur</b>             |
|                          | Professeure, Université du Littoral Côte d'Opale    |                              |
| <b>Sylvain PICAUD</b>    |   | <b>Examineur</b>             |
|                          | Directeur de recherche, Université de Franche-Comté |                              |
| <b>Abdenacer IDRISSI</b> |   | <b>Directeur de thèse</b>    |
|                          | Professeur, Université de Lille                     |                              |
| <b>Milán SZŐRI</b>       |   | <b>Co-directeur de thèse</b> |
|                          | Professeur associé, Université de Miskolc           |                              |



# **University of Lille**

Faculty of Sciences and Technologies

Doctoral School 104 – Science of the Matter, of the Radiation, and of the Environment  
and

# **University of Miskolc**

Faculty of Materials Science and Engineering

## **Co-tutorial Doctoral Thesis**

by

**Bálint Gábor KISS**

to obtain the degree of

**Doctor of the University of Lille**

Discipline: Theoretical, physical, analytical chemistry

Investigation of Interface Phenomena in Formamide–Water Systems by  
Computer Simulation Methods

Thesis presented and defended in Miskolc, on 28/01/2021 before the jury:

|                          |  |                      |
|--------------------------|--|----------------------|
| <b>Péter MIZSEY</b>      |  | <b>President</b>     |
|                          | Professor, University of Miskolc                 |                      |
| <b>Dezső BODA</b>        |  | <b>Referee</b>       |
|                          | Professor, University of Pannonia                |                      |
| <b>Michael KISELEV</b>   |  | <b>Referee</b>       |
|                          | Professor, Russian Academy of Sciences           |                      |
| <b>Szilárd FEJÉR</b>     |  | <b>Examiner</b>      |
|                          | Professor, University of Pécs                    |                      |
| <b>Sophie FOURMENTIN</b> |  | <b>Examiner</b>      |
|                          | Professor, University of the Littoral Opal Coast |                      |
| <b>Sylvain PICAUD</b>    |  | <b>Examiner</b>      |
|                          | Director of Research, University of Burgundy     |                      |
| <b>Abdenacer IDRISSE</b> |  | <b>Supervisor</b>    |
|                          | Professor, University of Lille                   |                      |
| <b>Milán SZŐRI</b>       |  | <b>Co-supervisor</b> |
|                          | Associate Professor, University of Miskolc       |                      |



## ***Acknowledgements***

*The creation of this thesis would not have been possible without the kind support of my supervisors. Therefore, I would like to express my sincere gratitude to Dr. Milán Szőri, Prof. Dr. Abdenacer Idrissi and Prof. Dr. Pál Jedlovszky for the continuous guidance during my research.*

*Beside my advisors, I also would like to thank to Dr. Balázs Fábrián, who has always taken the time to answer my questions and to Prof. Dr. Béla Viskolcz, who provided me the solid background needed for my research at the Institute of Chemistry.*

*I thank all of my colleagues and friends in Hungary and in France who stood by me in the most difficult times and even helped me through my first pandemic during my research in Lille.*

*Most importantly, I would like to thank my family: my parents, my brother, my sister and my uncle for their unconditional love and support.*

# Contents

|   |    |
|---|----|
| 1. Introduction.....  | 1  |
| 1.1. Aim and Objectives .....   | 11 |
| 2. Applied Methods.....   | 13 |
| 2.1 Thermodynamics of Mixing .....  | 13 |
| 2.2 Thermodynamic Integration .....   | 16 |
| 2.3 Basics of Computer Simulation Methods .....   | 18 |
| 2.4 Monte Carlo Simulation.....   | 25 |
| 2.4.1 Canonical (N,V,T) Monte Carlo .....   | 26 |
| 2.4.2 Grand-Canonical ( $\mu$ ,V,T) Monte Carlo .....   | 27 |
| 2.5 Molecular Dynamics Simulation .....   | 28 |
| 2.6 The Identification of Truly Interfacial Molecules (ITIM) .....                            | 30 |
| 2.7 Voronoi Analyses .....  | 33 |
| 3. Computational Details .....  | 35 |
| 3.1. The Mixing of Different Models of Formamide and Water .....                              | 35 |
| 3.2 The Liquid–Vapor Interface of Formamide–Water Mixtures.....                               | 40 |
| 3.3 Adsorption of Formamide at the Surface of Amorphous and Crystalline Ices.....             | 42 |
| 4. Results and Discussion .....   | 48 |
| 4.1. Miscibility and Thermodynamics of Mixing of Different Models of Formamide and Water..... | 48 |
| 4.1.1 Thermodynamic properties of mixing .....  | 48 |
| 4.1.2 The Miscibility of Different Formamide–Water Model Pairs.....                           | 56 |
| 4.2 Liquid–Vapor Interface of Formamide–Water Mixtures .....                                  | 58 |
| 4.2.1 Density Profiles.....   | 58 |
| 4.2.2 Lateral Distribution of the Surface Molecules.....                                      | 63 |

|  |     |
|--|-----|
| 4.2.3 Hydrogen Bonding at the Liquid Surface .....   | 65  |
| 4.2.4 Surface Orientation .....  | 67  |
| 4.2.5 Dynamics of the Surface Molecules .....  | 74  |
| 4.3 Adsorption of Formamide at the Surface of Amorphous and Crystalline Ices under<br>Interstellar and Tropospheric Conditions. .... | 78  |
| 4.3.1. Adsorption Isotherms.....   | 78  |
| 4.3.2 Density Profiles.....  | 84  |
| 4.3.3 Orientation of the First Layer Formamide Molecules .....   | 87  |
| 4.3.4 Energetics of the Adsorption .....   | 90  |
| 5. Summary.....  | 97  |
| References.....  | 107 |

# 1. Introduction

Mankind has always been fascinated with the mystery of the origin of life. Countless theories are born to explain this mystery, both scientific supported by facts, and some with less scientific explanations. Scientists have been trying for centuries to solve and understand the process of the formation of life as known on our Earth today. In 1859, Charles Darwin published his work "On the Origin of Species" and so the theory of biological evolution was born, which was one of the most significant steps to get closer to understand how life became so diverse on Earth. But the mystery of life was not solved at all. After the development of the evolution theory, the questions still remain: how and where was life born, and are we alone or does life exist outside of Earth?

In 1938, a new theory in the "The Origin of Life" was published by Oparin,<sup>1</sup> which was later supported by Urey<sup>2</sup>. Based on this work, the prebiotic chemical evolution theory was developed. According to the ever-expanding theory of prebiotic evolution, from simple precursors, such as H<sub>2</sub>, H<sub>2</sub>O, N<sub>2</sub>, NH<sub>3</sub>, CO, and CO<sub>2</sub>, through a complex network of prebiotic chemical reactions under plausible conditions occurred the formation of several building blocks of biomacromolecules, such as nucleotides and amino acids.<sup>3,4,13-16,5-12</sup> When the basis of life was formed, Earth had primitive atmosphere containing mostly methane, ammonia, water, and hydrogen. To confirm this theory, Miller<sup>17</sup> built an apparatus to model the atmosphere of the early Earth, in which CH<sub>4</sub>, NH<sub>3</sub>, H<sub>2</sub>O and H<sub>2</sub> were circulated. During the experiment, electric discharge was generated, which could have an important role in the process of the formation of more complex molecules in the primordial atmosphere of Earth (Figure 1.1). His experiment was a success, as he was able to identify glycine,  $\alpha$ -alanine,  $\beta$ -alanine, aspartic acid and  $\alpha$ -amino-n-butyric acid and suggested the presence of  $\beta$ - and  $\gamma$ -amino acids. During the last decades there has been an increasing amount of both theoretical and experimental evidences suggesting that hydrogen cyanide (HCN) and its hydration product formamide (HCONH<sub>2</sub>, FA) might have played a key role in the prebiotic evolution.<sup>3,4,13-16,18,5-12</sup> Matthews et al. sketched in their works<sup>19,20</sup> a HCN based world, where polypeptides and polynucleotides, formed from HCN polymers, were the precursors of both proteins and nucleic acids. Laboratory and extraterrestrial studies suggest that HCN polymerization is a universal process that

proceeds in our days as well as in the past in the interstellar medium (ISM) on planetary bodies or even satellites. Hence, we arrived to another significant question: was life born on Earth or somewhere else? The above-mentioned facts suggest the existence of protein-based life in an Earth-like environment also elsewhere in the Universe.

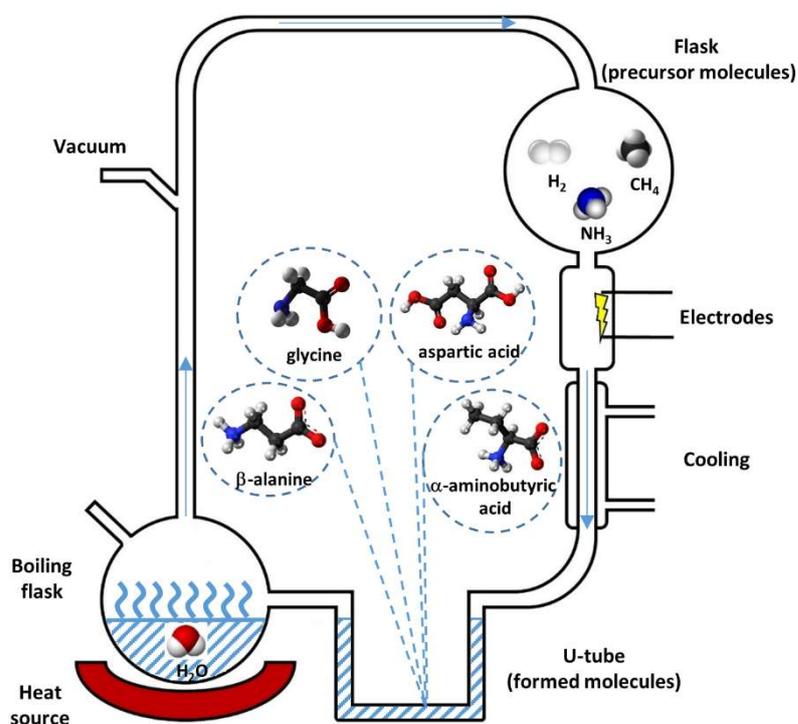


Figure 1.1: Illustration of the Miller experiment.

The Swedish chemist Svante Arrhenius proposed the radiopanspermia hypothesis in 1903.<sup>21</sup> According to this, seeds of life (between the size 200-300 nm) travelled by solar radiation to Earth. This assumption was followed by the lithopanspermia (interstellar panspermia) and ballistic panspermia (interplanetary panspermia) hypotheses about the transfer of organisms in rocks from one planet to another. Oró published in 1961 his hypothesis, according to which comets impacting the primitive Earth could be the source of terrestrial materials, including some important precursors for prebiotic synthesis.<sup>22,23</sup> In 1978, Fred Hoyle suggested viruses travelled to Earth by interstellar comets.<sup>24</sup> Current astrochemical research address the question how prebiotic molecules happened to reach the early Earth, creating the opportunity for the formation of life as we know it, and whether this process could happen elsewhere in the Universe. Many molecules of interest for exobiology have been detected in comets: in 1996 Comet Hyakutake was found to contain ammonia, methane, acetylene, acetonitrile and hydrogen isocyanide.<sup>25</sup> Comet Hale-Bopp showed hydrogen cyanide, formaldehyde methane, acetylene, formic acid, acetonitrile, hydrogen isocyanide, isocyanic acid, cyanoacetylene,

and thioformaldehyde content.<sup>25</sup> These findings indicate that molecules in cometary grains could have played an important role of organic materials reaching the primitive Earth.<sup>22,23,26,27</sup> Formamide has also been detected several times in astronomical sources, as it is shown on Table 1.1, based on the work of López-Sepulcre et al.<sup>18</sup>

*Table 1.1: Detection of Formamide in space.<sup>18</sup>*

| Reference                            | Type of            | Telescope      | Astronomical Sources                                  |
|--------------------------------------|--------------------|----------------|---|
| <b><i>Star-Formation Regions</i></b> |                    |                |   |
| Rubin et al. <sup>28</sup>           | high-mass          | 140 ft NRAO    | Sgr B2; Sgr A   |
| Turner <sup>29</sup>                 | high-mass          | 11 m NRAO      | OMC-1/Ori KL  |
| Gibb et al. <sup>30</sup>            | high-mass          | SEST           | G327.3–0.6  |
| Bisschop et al. <sup>31</sup>        | high-mass          | JCMT           | G24.78; G75.78; NGC 6334 IRS1;<br>NGC 7538 IRS1       |
| Feng et al. <sup>32</sup>            | high-mass          | IRAM PdBI      | NGC 7538 S  |
| Bisschop et al. <sup>31</sup>        | high-mass          | JCMT           | W 3(H <sub>2</sub> O); W 33A                          |
| Isokoski et al. <sup>33</sup>        | high-mass          | JCMT           | IRAS 20126 + 4104;<br>IRAS 18089-1732; G31.41 + 0.31  |
| Suzuki et al. <sup>34</sup>          | high-mass          | NRO 45 m       | G10.47 + 0.03; G19.61–0.23;<br>G34.3 + 0.2; W51 e1/e2 |
| Xu and Wang <sup>35</sup>            | high-mass          | SMA            | G20.08–0.14N  |
| Allen et al. <sup>36</sup>           | high-mass          | ALMA           | G35.20; G35.03  |
| Sakai et al. <sup>37</sup>           | high-mass          | ALMA           | G34.43 + 00.24 MM3                                    |
| Lopez-Sepulcre et al. <sup>38</sup>  | IM hot corino      | IRAM 30 m      | Cep E; OMC-2 FIR 4                                    |
| Lopez-Sepulcre et al. <sup>38</sup>  | protocluster       | IRAM 30 m      | OMC-2 FIR 4   |
| Marcelino et al. <sup>39</sup>       | young hot corino   | ALMA           | Barnard 1   |
| Kahane et al. <sup>40</sup>          | hot corino         | IRAM 30 m      | IRAS 16293-2422                                       |
| Lopez-Sepulcre et al. <sup>38</sup>  | hot corino         | IRAM 30 m      | NGC 1333 IRAS 4A                                      |
| Taquet et al. <sup>41</sup>          | hot corino         | IRAM PdBI      | NGC 1333 IRAS 2A                                      |
| Imai et al. <sup>42</sup>            | hot corino         | ALMA           | B335  |
| Oya et al. <sup>43</sup>             | hot corino         | ALMA           | L483  |
| Lee et al. <sup>44</sup>             | hot corino         | ALMA           | HH 212  |
| Lopez-Sepulcre et al. <sup>38</sup>  | evolved hot corino | IRAM 30 m      | NGC 1333 SVS13A                                       |
| Yamaguchi et al. <sup>45</sup>       | protostellar shock | NRO 45 m       | L1157-B1  |
| Mendoza et al. <sup>46</sup>         | protostellar shock | IRAM 30 m      | L1157-B2  |
| Ceccarelli et al. <sup>47</sup>      | protostellar shock | NOEMA          | IRAS 4 As   |
| <b><i>Solar System</i></b>           |                    |                |   |
| Bockelee-Morvan et al. <sup>48</sup> | comet              | CSO            | Hale-Bopp   |
| Biver et al. <sup>49</sup>           | comet              | IRAM 30 m      | C/2012 F6; C/2013 R1                                  |
| Goesmann et al. <sup>50</sup>        | comet              | COSAC(Rosetta) | 67P/Churyumov-Gerasimenko                             |
| <b><i>Other Environments</i></b>     |                    |                |   |
| Müller et al. <sup>51</sup>          | spiral galaxy      | ATCA           | PKS 1830-211  |
| Corby et al. <sup>52</sup>           | translucent cloud  | ATCA           | l.o.s. Sgr B2   |
| Eyres et al. <sup>53</sup>           | BD-WD merger       | ALMA           | CK Vulpeculae   |

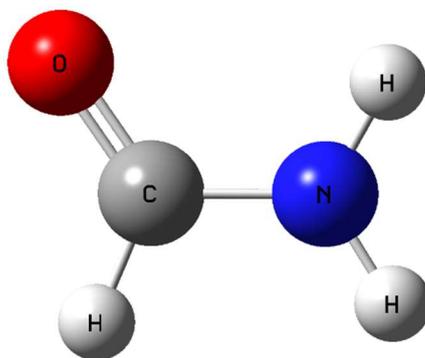
Formamide is present in galactic centers,<sup>28,54</sup> star-forming regions of dense molecular clouds,<sup>55</sup> high-mass young stellar objects,<sup>56</sup> the interstellar medium,<sup>57</sup> comets,<sup>48,58,59</sup> and satellites.<sup>60</sup> Although its formation, whether on the surfaces of the interstellar grains or in the gas phase, is currently under debate, a recent quantum chemical calculation of two competitive reaction channels (i.e.,  $\text{NH}_2 + \bullet\text{HCO}$  and  $\bullet\text{CN} + \text{H}_2\text{O}$ ) suggested plausible formamide formation routes in the presence of amorphous ice.<sup>61</sup> Since the temperature of the interstellar medium is supposed to be below 110 K, eutectic aqueous mixture of formamide cannot play an important role at the interstellar conditions.<sup>62</sup> Of all the molecules that can be found in space, the frequently detected formamide is one of the most significant one in regard to the prebiotic evolution. High energy particles (e.g., cosmic ray, solar wind), heat, electromagnetic radiation, and radioactive decay continuously interact with simple chemical precursors to yield new complex derivatives. If the activated molecule can interact with the proper reactant, then amino acids, nucleobases, sugars, lipids, and carboxylic acids can emerge very easily as synthesizable molecules.<sup>63</sup> Formamide is a key precursor molecule, as it has an amide functional group to form chains of amino acids and build up proteins. It is a precursor of carboxylic acids, amino acids, and sugars, furthermore in the presence of an energy source, it promotes the synthesis of adenine, guanine, cytosine, and uracil, which are the four nucleobases of ribonucleic acid.<sup>8,10,18,64,65</sup> However, under the extremely low densities present in the interstellar medium, sufficient formamide chemistry requires strong local enhancement of the formamide concentration. Such areas of locally high formamide concentration can again be expected at the surfaces of icy objects, such as cometary<sup>66</sup> or interstellar<sup>67</sup> dust grains that are frequently covered by low density amorphous ice (LDA). It is important to note that if formamide is formed on solid surfaces, its gas-phase detection and the subsequent (photo)chemistry processes would thus strongly depend on the formamide interaction with water molecules (WAT).<sup>68,69</sup>

In 2019, Vichiatti and his colleagues further confirmed that more complex molecules like formamide could also be synthesized in Earth's primordial atmosphere.<sup>70</sup> Ferus et al. investigated the formation of nucleobases from formamide during a high energy density event (i.e. the impact of an extraterrestrial body into the planetary atmosphere) by irradiation of "formamide ice" created with liquid nitrogen cooled to 77 K and liquid samples with a high-power laser in the presence of potential catalysts.<sup>71</sup> They suggested kinetic and thermodynamic models, as well as reaction routes leading from the

dissociation products (HCN, H<sub>2</sub>O, HNCO, H<sub>2</sub>, CO, NH<sub>3</sub>, HNC, •CN, and •NH) to the nucleobases. They also studied isocyanic acid (HNCO) based synthesis of formamide in planetary atmospheres and they proved that the direct hydrogenation of HNCO (HNCO+H<sub>2</sub> → HCONH<sub>2</sub>) is thermodynamically favored.<sup>72</sup> Based on our knowledge, the synthesis of formamide could have taken place on the early Earth as well as in the interstellar medium.

Formamide is considered as an alternative prebiotic solvent to water which can be explained by their similar properties.<sup>73</sup> So far, there is no proof of the existence of highly purified formamide on the primitive Earth, but there could have been formamide–water mixtures. The processes of producing more complex biomolecules from these sources are less effective than from pure formamide. There could have been change in environmental conditions (e.g. temperature) that may have led to the process to form sufficient quantity of pure formamide for the prebiotic synthesis. On other celestial bodies as planets or moons, formamide–water eutectic or binary solutions are existing at temperatures not below 223 K.<sup>74</sup>

Water is the most common polar chemical compound on Earth and it also bears a lot of unusual properties both in itself and as a solvent. Therefore, water is frequently in the center of scientific attention, and became the subject of countless scientific works,<sup>75–77</sup> which is mainly due to that fact the water molecules can both donate and accept 2-2 hydrogen-bonds, which enables them to form extensive hydrogen-bond network.



*Figure 1.2: The sematic representation of the formamide molecule structure (optimized by GAUSSIAN software with b3lyp/6-31g method)*

Similarly to water, formamide and its aqueous mixtures are also important systems, both from the scientific point of view and from the approach of industrial application. Formamide (FA) was synthesized for the first time in 1863 by W. A. Hoffmann from ethyl formate and ammonia.<sup>78</sup> It is the first member of the primary amide series, colorless and hygroscopic liquid at room temperature, with a faint odor of ammonia. Formamide is the smallest molecule that contains a peptide bond. It is also the simplest molecule that is able to form N-H $\cdots$ O type hydrogen bonds, thus, its aqueous mixtures are often considered as model systems in studying various properties of hydrated proteins and polypeptides, e.g., how hydrogen bonds<sup>79,80</sup> and peptide bonds<sup>81</sup> are formed and broken, or how water exercises kinetic and thermodynamic control over the chemical activities of polypeptides.<sup>82</sup> These mixtures can also be used as reference systems to study hydrophobic and hydrophilic interactions.<sup>83</sup> Some results of earlier works concerning the structure of liquid formamide, based on the summary of Bakó et al.,<sup>84</sup> supplemented by those of more recent investigations are collected in Table 1.2.

*Table 1.2: Conclusions of earlier investigations of the liquid formamide structure*

| Reference                             | Methods     | Conclusion                                   |
|---------------------------------------|-------------|--|
| De Sando and Brown <sup>85</sup>      | XD          | Ring dimer, resembles crystal                |
| Sieghbarn et al. <sup>86</sup>        | ESCA        | H-bonded chain                               |
| Nielsen <sup>87</sup>                 | Raman       | Chain structure                              |
| Kálmán et al. <sup>88</sup>           | XD,ND,ED    | Four H-bonds per molecule                    |
| Ohtaki et al. <sup>89</sup>           | XD          | Cyclic dimer and chain                       |
| Radnai et al. <sup>90</sup>           | XD          | Ring and chain mixed structure               |
| Miyake et al. <sup>91</sup>           | XD          | Distorted ring dimer, similar to crystal     |
| Bellissent-Funel et al. <sup>92</sup> | ND,XD       | Ring and open chain structure                |
| Wiesmann et al. <sup>93</sup>         | ND          | Antiparallel orientation, ring structure     |
| Essex and Jorgensen <sup>94</sup>     | MC          | Ring and chain mixed                         |
| Puhovski and Rode <sup>83</sup>       | MD          | Continuous H-bonded network                  |
| Puhovski et al. <sup>95</sup>         | MD          | Continuous H-bonded network                  |
| Tsuchida <sup>96</sup>                | AbMD        | H-bonded network                             |
| Megyes and co-workers <sup>97</sup>   | MD,XD       | H-bonded network                             |
| Cordeiro, J. M. M. <sup>98</sup>      | MC          | Flexible chains with immersed cyclic dimers  |
| Richardi et al. <sup>99</sup>         | MOZ         | Networks with parallel dipole in first shell |
| Eloia and Ladányi <sup>100</sup>      | MD          | H-bonded network, very few cycles            |
| Barthel et al. <sup>101</sup>         | Dielectrsp. | Six member ring, chain                       |
| Jadzyn et Swiergiel <sup>102</sup>    | MD          | H-bonded network, linear and cycles          |
| Lima et al. <sup>103</sup>            | IR,MD       | Extent of the H-bond network                 |
| Ludwig et al. <sup>104</sup>          | NMR         | Six member ring                              |
| Macchiagodena et al <sup>105</sup>    | MD          | Strong H-bonded network                      |

Formamide is a highly polar molecule, and its liquid has an unusually high dielectric constant. This high dielectric constant of about  $109^{106}$  cannot be simply explained by the large molecular permanent dipole moment of 3.7 D.<sup>107</sup> Thus, for instance, the dielectric constant of its double methylated derivative, N,N-dimethylformamide, is only 37,<sup>106</sup> despite the fact that the magnitude of dipole moment of this molecule is as large as 3.9 D.<sup>107</sup> Instead, the large dielectric constant of liquid formamide is related to the fact that, similarly to water, its molecules also exhibit extensive hydrogen bonding by donating both hydrogen atoms of the nitrogen atom and accept up to two hydrogen atoms with the oxygen atom for hydrogen bonding. As a consequence, similarly to water,<sup>103,108,109</sup> formamide is also a network-forming liquid,<sup>96,84</sup> they both form continuous and extensive hydrogen-bonding network. Hence, these molecules can locally substitute each other in the hydrogen bond network,<sup>100</sup> although such substitution may change the topology of the network.<sup>84,102</sup> Due to these similar properties of the two neat liquids, mixtures of formamide and water are often thought to be examples of the ideal mixture.<sup>110,111</sup> Indeed, these two compounds are fully miscible with each other,<sup>112</sup> and the thermodynamic properties, such as the energy, entropy and free energy of their mixture depend almost linearly on the composition. The change of the energy and enthalpy occurring upon mixing the two compounds in any proportion turned out to be very close to zero both in experiments<sup>113,114</sup> and in computer simulations.<sup>115</sup>

Neat liquid formamide<sup>81,83,84,94,96,98,102,103,105,116</sup> and its aqueous mixtures<sup>80,100,110,111,115,117–120</sup> have been studied by computer simulation methods on several occasions, due to their aforementioned great importance. However, in simulating liquid mixtures, one has to encounter the problem that potential models that work perfectly for the respective neat liquids might be incompatible with each other, resulting in a poor reproduction of the properties of the mixtures. If the thermodynamic driving force behind the miscibility of the two compounds is small, then limited miscibility of certain model pairs of the two fully miscible compounds may occur.<sup>121,122</sup> Since the mixing of water and formamide is close to the ideal mixing,<sup>110,111,113–115</sup> the thermodynamic changes occurring upon their mixing are expected to be rather small (as it was already shown concerning their energy of mixing<sup>113–115</sup>). Consequently, the thermodynamic driving force of this mixing must be small, which makes the above issue – concerning the compatibility of the potential models with each other and their ability to

reproduce full miscibility – particularly important for these mixtures. Puhovski and Rode investigated the structure and dynamics of formamide-water mixtures by molecular dynamics simulations and found that a formamide-like order dominates in the spatial arrangement of the system.<sup>95,117</sup> Cordeiro and his colleagues studied water-amide liquid mixtures by Monte Carlo simulations, applying a six-site model of formamide together with the TIP4P model for water.<sup>118</sup> They observed the typical formation of three hydrogen-bonds between a formamide molecule and its neighboring waters. They also observed that the strength of the amide-water hydrogen-bond increases as the number of hydrogen bonds per amide decreases.

Formamide is also present in a non-negligible amount in the atmosphere, and their atmospheric fate is mainly due to its rainout.<sup>123</sup> Amides are emitted directly to Earth's atmosphere from biological sources as well as from industrial processes.<sup>124</sup> Large amount of amides can also form directly in the atmosphere via oxidation of available amines. In particular, methylamine, the smallest alkylamine, the oxidation of which may lead to formamide, has a global emission of  $24 \pm 15$  Gg N per year.<sup>125</sup> Atmospheric level of amides is in the range of pptv, and they have been detected in ambient particles, biomass burning aerosols, and fogwater, as well.<sup>126–128</sup> Carbon capture and storage (CCS) technology may represent a significant source of amides to ambient air in the future, since formamide was detected from an industrial scale carbon capture site.<sup>129</sup> However, up to now, only a few investigations have been carried out to understand the atmospheric fate of amides.<sup>130</sup> Photolysis can be a negligible sink for amides, as indicated by a study of cross-section measurement.<sup>131</sup> On the other hand, oxidation can be a relevant process concerning the atmospheric loss of formamide. Barnes et al. determined rate coefficients for the reaction of OH•, Cl, NO<sub>3</sub>, and O<sub>3</sub> with formamide at room temperature by means of in situ FTIR spectrometry.<sup>132</sup> The established rate constant of the formamide + OH• reaction was also confirmed using proton transfer reaction–mass spectrometry (PTR-MS) detection by two independent research groups.<sup>130,133</sup> The corresponding atmospheric lifetime of formamide was estimated to be a few days.<sup>132</sup> Due to the high affinity of formamide to water, uptake and subsequent deposition are proposed to be another important sink,<sup>134</sup> although, to the best of our knowledge, it was never investigated. In the troposphere, such an uptake might predominantly involve adsorption at the surface of ice grains.

To the best of our knowledge, no computer simulation study has targeted so far the liquid-vapor interface of formamide-water mixtures as well as ice surfaces containing adsorbed formamide molecules, although the above questions concerning the mixing of these molecules and the fate of formamide in the modern atmosphere as atmospheric pollutant<sup>123</sup> as well as the biomimetic modelling of polypeptides are closely related also to the behavior of formamide at the surface of dense aqueous phases. To contribute to the prebiotic evolution, concentration enrichment of formamide is also essential. These phenomena may take place at interfaces like water or ice in the atmosphere of Earth or in the interstellar medium, as in the case of HCN.<sup>135,136</sup> The detailed understanding of the adsorption of formamide at the surface of crystalline (I<sub>h</sub>) ice at tropospheric temperatures (i.e., around 200 K) is important from the point of view of atmospheric chemistry. On the other hand, its adsorption on amorphous ice (LDA) under interstellar conditions, i.e., at temperatures below 110 K, is a process of relevance in astrochemistry, with a special importance also in the field of the above discussed prebiotic evolution. However, due to the extreme conditions, experimental investigation of these problems is far from being straightforward.

Binary mixtures represent more complex systems, from the thermodynamical point of view, than neat systems due to the presence of the composition as an additional intensive variable. In other words, the mole fraction of the components determines the thermodynamic state of the system. Understanding the behavior of binary systems is, however, of great importance in the fields of chemistry and physical chemistry. A two-component system is the simplest model for the study of mixtures and this makes it easier to explain the processes and phenomena occurring in more complex systems. Thus, although industrial processes usually involve blends with much larger number of compounds than two, nonetheless, binary systems are still in the focus of researchers. Such studies can be used to approximate and model the chemical, physics, and thus also the thermodynamical behavior of real mixtures. Binary systems can be applied to develop complex models for mixtures (e.g., activity coefficient models or equations of state), in which the intermolecular interactions are usually limited to binary interactions.<sup>137</sup> Such binary mixtures can nowadays be examined in detail by methods of computation chemistry and computer simulation. Formamide and water create one of the most interesting binary systems, while its components are of great importance in their neat state, as well.

Computer technology is unstoppably and essentially spreading to all different branches of science. Our predecessors were only able to rely on their observations and experiments. Experimentalists in chemistry and physics have tried to reach better understanding over the systems they studied, for examples the structure and dynamics of their particles, while theoreticians have worked on models to explain the behavior of the systems in interest. In our days, we have another technique in our hands, namely computer simulation.<sup>138</sup> This method allows us to make models about such complicated systems that was never possible before, and get an insight into their properties at the atomistic level, even in cases when experiments are too dangerous, expensive, difficult, or impossible to perform, such as under the extreme conditions of the interstellar medium. This way, computer simulation methods provide us with the resource to determine macroscopically measurable properties and to interpret macroscopic phenomena based on the knowledge of the microscopic properties of the studied systems. In statistical mechanics, computer simulation is able to provide exact results for problems, while only approximate results are achieved if other, theoretical methods are utilized. Hence, computer simulations can act as numerical experiments in testing theories. Comparing the results of the simulation with experimental results, on the other hand, provides the opportunity of verifying the models used in the simulation.

The mixture and the miscibility of formamide and water can be investigated by the Monte Carlo variant of computer simulations as it was successfully used in other works to study different mixtures.<sup>139-144</sup> The vapor-liquid interface can be investigated by molecular dynamics simulation as in the case of similar systems.<sup>145</sup> The adsorption processes at the ice surface can easily be studied in detail by grand canonical Monte Carlo (GCMC) simulation,<sup>138,146</sup> in which the chemical potential rather than the number of adsorbate molecules is controlled. Thus, by systematically varying the chemical potential and determining the number of adsorbed molecules as its function, the adsorption isotherm can easily be calculated. Information gained from computer simulations can provide interesting support for the interpretation of experimental data. For example, thermal desorption experiments<sup>69,147</sup> as well as simplified chemical models<sup>68</sup> indicate that the binding energy of formamide on ice is larger than the interaction between the water molecules, which is easy to check by GCMC simulations. In addition, the detailed analysis of the molecular orientations in the simulations can provide important information on the orientation and surrounding of the C=O bond, which corresponds to

an intense spectral feature usually easily identifiable in infrared spectroscopic observations.<sup>147</sup>

## 1.1. Aim and Objectives

As presented in the introductory chapter, there are still knowledge gaps in the leading theories in connection with the formation of life, including the role of formamide. To reach an effective formamide enrichment which is necessary for playing a role in the biomolecule synthesis, accumulation of formamide molecules has occurred. This can be explained with the sometimes extreme and complex environments including interfaces. The aim of this doctoral dissertation is to interpret the phenomena occurring at interfaces of formamide-water binary systems using computer simulation methods at ambient and cold condition ( $T=200\text{ K}$ ,  $100\text{ K}$  and  $50\text{ K}$ ).

To achieve this purpose, first a formamide-water model pair has to be selected to reproduce the real thermodynamic properties of the binary mixtures. Therefore, the thermodynamics of mixing formamide and water are studied by computer simulation applying five widely-used formamide and three water models. The Helmholtz free energy ( $A$ ), energy ( $U$ ), and entropy ( $S$ ) of mixing of these model combinations in the entire composition range is needed to be calculated and analyzed. Thermodynamic changes occurring upon mixing will be calculated along an appropriately chosen thermodynamic cycle<sup>148</sup> using the method of thermodynamic integration<sup>149,150</sup> and Monte Carlo simulations on the canonical ensemble. The performance of the different model combinations considered is evaluated according to their ability to reproduce full miscibility, and by comparing the calculated energy of mixing to the existing experimental data.<sup>113,114</sup> The formamide-water model pair with the best performance will then be applied for simulations in further investigations of this thesis.

Since other prebiotic precursor molecules show enrichment at the interface of their aqueous mixtures<sup>145,151,152</sup>, the liquid-vapor interface of formamide-water mixtures is promising for accumulation. Hence, the properties of the interface of formamide aqueous mixtures covering the entire composition range from neat water to neat formamide are then investigated by molecular dynamics simulation and surface analysis. The adsorption of formamide in the surface layer, the density profiles, the lateral self-association,

orientation and the dynamics of the surface molecules, as well as their lifetime and diffusion at the liquid surface will be the subject of detailed analysis and discussion.

Alternative to the formamide enrichment at vapor/liquid interface of formamide-water mixtures, adsorption of formamide at cold surfaces ( $T= 200$  K,  $100$  K and  $50$  K) was investigated which may have importance in the prebiotic evolution, as well as might support panspermia theory. The adsorption of formamide both at the surface of crystalline ( $I_h$ ) ice (at  $200$  K) and at the surface of low density amorphous (LDA) ice (at  $200$ ,  $100$ , and  $50$  K) is simulated by the Grand Canonical Monte Carlo method. The adsorption isotherms, density profile of the adsorption layer, as well as binding energy and surface orientation of the adsorbed molecules that belong to the first molecular layer are investigated and discussed in detail.

## 2. Applied Methods

### 2.1 Thermodynamics of Mixing

Based on the IUPAC definition,<sup>153</sup> mixing is the process of combining components, particles or layers into a more homogeneous state. This process must not permit segregation of particles of different size or properties. Homogeneity may be considered to have been achieved in a practical sense, when the sampling error of the processed portion is negligible compared to the total error of the measurement system. The mixing of two chemical compounds (A=FA, B=WAT) results in a change in the entropy ( $S$ ), energy ( $U$ ) and volume ( $V$ ) of the solution. These changes in the case of isothermic mixing cause a variation of the Gibbs free energy as:<sup>154-156</sup>

$$\Delta G^{mix} = \Delta U^{mix} + p\Delta V^{mix} - T\Delta S^{mix} \quad (2.1)$$

where  $p$  is the pressure and  $T$  is the absolute temperature. The miscibility of two compounds is thermodynamically favored when  $\Delta G^{mix} < 0$ . Furthermore, based on the IUPAC definition of the miscibility<sup>157,158</sup> for a two-component mixture, a necessary and sufficient condition for stable or metastable equilibrium of a homogeneous single phase is for the entire composition:

$$\left( \frac{\partial^2 \Delta G^{mix}}{\partial x_A^2} \right)_{T,p} > 0 \quad (2.2)$$

where  $x_A$  is the mole fraction of component  $A$ . The borderline (spinodal curve) between metastable and unstable states is defined by the above second derivative equaling zero. Hence, if free energy of mixing as function of concentration has an inflection point, the system is unstable and phase separation occurs. The experiments are usually performed under constant pressure, hence, to describe the thermodynamics of mixing, the Gibbs free energy of mixing is the appropriate choice.<sup>154,159</sup> However, if the experiment requires constant volume as it is the case in computer simulations, the Helmholtz free energy of mixing ( $\Delta A^{mix}$ ) is the suitable alternative. The difference between the Gibbs and Helmholtz free energies is negligible for condensed systems at atmospheric pressure ( $\Delta G^{mix} = \Delta A^{mix} + p\Delta V^{mix}$ ).<sup>121</sup> The Helmholtz free energy of mixing can be written as:

$$A^{mix} = U^{mix} - TS^{mix} \quad (2.3)$$

For the case of ideal mixing (as in the case of the mixing of ideal gases) in a two-component system, it can be expressed as:

$$A^{\text{mix}} = RT (x_A \ln x_A + x_B \ln x_B) \quad (2.4)$$

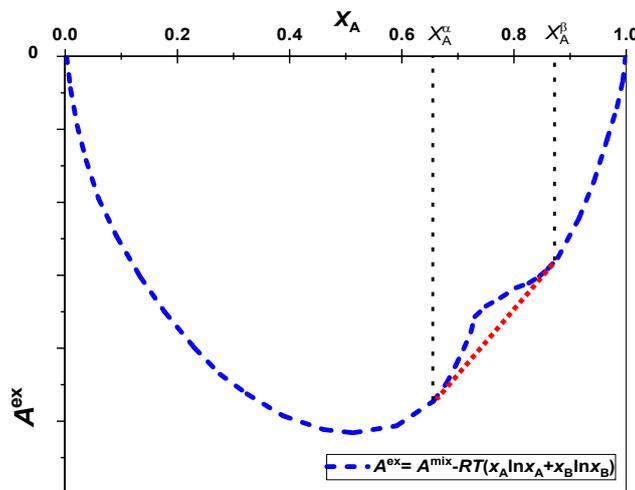
where  $R$  is the universal gas constant. The thermodynamic condition that two liquid compounds, A and B, are fully miscible with each other, applying the excess molar Helmholtz free energy ( $A^{\text{ex}}$ ), can be formulated as the following inequality:<sup>158</sup>

$$D = 1 + x_A x_B \frac{\partial^2 (A^{\text{ex}} / RT)}{\partial x_A^2} > 0 \quad (2.5)$$

This inequality has to be valid in the entire composition range of  $0 < x_A < 1$ . Here,  $x_A$  and  $x_B$  denote the mole fractions of the respective components ( $x_A + x_B = 1$ ), and the excess molar Helmholtz free energy with respect to the ideal mixing of the two components can be written as:

$$A^{\text{ex}} = A^{\text{mix}} - RT(x_A \ln x_A + x_B \ln x_B) \quad (2.6)$$

It should be emphasized that the inequality  $A^{\text{mix}} < 0$  is a necessary but not a sufficient condition of full miscibility, as it simply reflects that the mixture is thermodynamically more stable than the two separate neat components. However, miscibility also requires that the single-phase mixture is stable also with respect to any two phase systems consisting of mixed phases of different compositions. This condition is taken into account in eq. 2.5, and is illustrated in Figure 2.1.



*Figure 2.1: The schematic representation of the excess Helmholtz free energy in the case of phase separation in the composition range between  $x_A^\alpha$  and  $x_A^\beta$ . If the second derivative in the eq. 2.5 has an inflection point, the mixture is not stable. Red line shows full miscibility.*

The quantity  $D$  (eq. 2.5) is equal to zero at the critical solution points, i.e., where the phase separation occurs in a binary system. Further, in the case of ideal mixtures, its value is unity in the entire composition range. The  $D$  parameter can therefore be used to measure the “closeness” of the system both to phase separation and to the ideal mixture. In computer simulations,  $A^{\text{mix}}$  can be calculated by considering the following thermodynamic cycle (see Figure 2.2).<sup>148</sup>

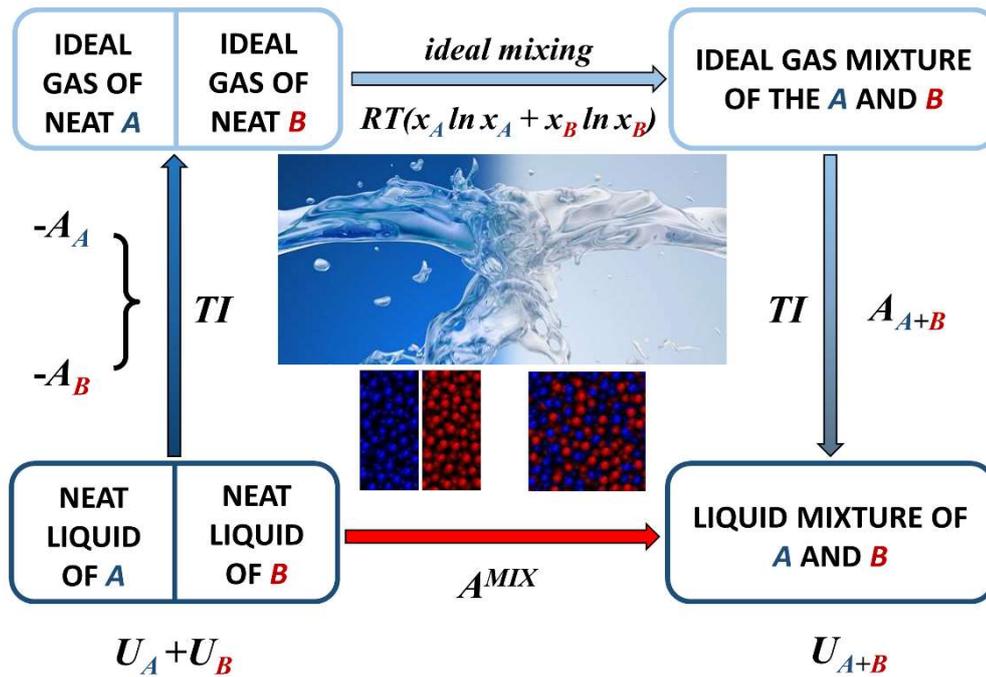


Figure 2.2: Schematic diagram showing the thermodynamic cycle along which the free energy of mixing of the two components is calculated.

In the first step, the neat components are brought from the liquid to the ideal gas state under isochoric conditions. In the second step, the two components are mixed in the ideal gas state; this step is accompanied by the free energy change of the ideal mixing of  $RT(x_A \ln x_A + x_B \ln x_B)$ . Finally, in the third step, the mixture is brought back isochorically from the ideal gas to the liquid state. Thus, the free energy change of the net process,  $A^{\text{mix}}$ , can be written as

$$A^{\text{mix}} = A_{A+B} - x_A A_A - x_B A_B + RT(x_A \ln x_A + x_B \ln x_B) \quad (2.7)$$

where  $A_A$ ,  $A_B$ , and  $A_{A+B}$  are the Helmholtz free energies of the corresponding systems (i.e., neat liquid A, neat liquid B, and their mixture, respectively) with respect to their

ideal gas state. Using the same thermodynamic cycle as for the calculation of  $A^{\text{mix}}$ , and considering that mixing in the ideal gas state is accompanied by no energy change, the energy of mixing of the two components,  $U^{\text{mix}}$ , can be written as

$$U^{\text{mix}} = U_{A+B} - x_A U_A - x_B U_B \quad (2.8)$$

where  $U_A$ ,  $U_B$ , and  $U_{A+B}$  denote the internal energies of the corresponding systems. Finally, the entropy of mixing of the two components,  $S^{\text{mix}}$ , can simply be calculated from  $U^{\text{mix}}$  and  $A^{\text{mix}}$  as

$$S^{\text{mix}} = \frac{U^{\text{mix}} - A^{\text{mix}}}{T} \quad (2.9)$$

## 2.2 Thermodynamic Integration

While the terms  $U_A$ ,  $U_B$ , and  $U_{A+B}$  in eq. 2.8 can simply be calculated in computer simulations as the total potential energy of the corresponding systems, the determination of  $A_A$ ,  $A_B$ , and  $A_{A+B}$  in eq. 2.7 is not that straightforward. In general, unlike other thermodynamic quantities, such as the internal energy, the configurational part of the Helmholtz free energy, being proportional to the logarithm of the configurational integral (i.e., configurational part of the total partition function) on the canonical ensemble ( $N, V, T$ ), is practically impossible to be accessed in a computer simulation. This is because here the entire configurational space needs to be sampled instead of only its lowest energy domains. (The kinetic part of the Helmholtz free energy depends only on the velocities of the particles, and hence remains unchanged at constant temperature.) It can be computationally feasible, however, to calculate the free energy *difference* between two states, as in this case only those domains of the configurational space are needed to be sampled that are considerably different in the two states.

In the method of thermodynamic integration (TI),<sup>149,150</sup> the difference of the Helmholtz free energies between states X and Y is calculated as an integral along an arbitrarily chosen path connecting the two states:

$$\Delta A = A_Y - A_X = \int_0^1 \left( \frac{\partial A(\lambda)}{\partial \lambda} \right) d\lambda \quad (2.10)$$

where  $\lambda$  is the coupling parameter that describes this path, its value being 0 in state X and 1 in state Y. Considering the fundamental relations of statistical mechanics of  $A = -k_B T \ln Q$  and  $Q = \int \exp(-\underline{U}/k_B T) d\mathbf{q}^N$ , the integrand of eq. 2.10 can be written as

$$\frac{\partial A(\lambda)}{\partial \lambda} = -k_B T \frac{1}{Q(\lambda)} \frac{\partial Q(\lambda)}{\partial \lambda} = \frac{\int \frac{\partial U(\lambda)}{\partial \lambda} \exp(-\beta U(\lambda)) d\mathbf{q}^N}{\int \exp(-\beta U(\lambda)) d\mathbf{q}^N} = \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \quad (2.11)$$

where  $Q$  is the configurational integral,  $\mathbf{q}^N$ , which denotes the coordinates of all  $N$  particles in the system, represents a given point of the configurational space,  $k_B$  is the Boltzmann constant,  $\beta=1/k_B T$ , and the brackets  $\langle \dots \rangle_{\lambda}$  denote ensemble averaging at the state corresponding to the given value of  $\lambda$ .

When connecting states X and Y, a polynomial path is usually chosen. Considering also that in systems where the energy is pairwise additive and the leading term of the pair potential (i.e., steric repulsion) decays with  $r^{-12}$ , this polynomial has to be at least of fourth order to avoid divergence of  $U(\lambda)$  at  $\lambda = 0$ ,<sup>149</sup> this path is conventionally chosen in computer simulations as

$$U(\lambda) = \lambda^4 U_Y + (1 - \lambda)^4 U_X \quad (2.12)$$

where  $U_X$  and  $U_Y$  are the energies of the corresponding states. If the Helmholtz free energy difference between the (isochoric) liquid and ideal gas states of a system needs to be calculated,<sup>160-164</sup> as in the present case, eq. 2.12 can be simplified to  $U(\lambda) = \lambda^4 U_Y$ , considering that the potential energy in the ideal gas state, X, is zero. In this case, using eq. 2.11, eq. 2.10 can be simplified as

$$\Delta A = \int_0^1 \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda = \int_0^1 4\lambda^3 \langle U_Y \rangle_{\lambda} d\lambda \quad (2.13)$$

Further, the Boltzmann factor in eq. 2.11, evaluated in each step of a Monte Carlo simulation, can be rewritten as

$$\begin{aligned} \exp(-U(\lambda)/k_B T) &= \exp(-\lambda^4 U_Y / k_B T) \\ &= \exp(-U_Y / k_B (T / \lambda^4)) = \exp(-U_Y / k_B T^*) \end{aligned} \quad (2.14)$$

where  $T^* = T/\lambda^4$ . Therefore, a simulation performed at a given  $\lambda$  point along the (fictitious) path connecting states X and Y with the potential function  $U(\lambda)$ , needed to evaluate the

ensemble average in eq. 2.13, can be substituted with a simulation performed in state Y (i.e., at  $\lambda = 1$ , in the liquid state of interest) with the full potential function  $U(1) = U_Y$ , but at the virtual temperature  $T^* = T/\lambda^4$ . In other words, the path connecting the liquid and ideal gas states along  $\lambda$  by gradually decreasing the potential energy to zero can be substituted by a path along which the virtual temperature of the system is increased gradually to infinity. This way, the ideal gas state is simply defined as a state of infinite kinetic energy rather than that of zero potential energy. It should be recalled that the path connecting states X and Y is fictitious, therefore, the points along which it passes (i.e., the systems simulated at various virtual temperatures  $T^*$ ) have no physical relevance by themselves. These fictitious states only serve to connect the two endpoints of the path, i.e., the two states that are of real physical relevance, the free energy difference of which is to be calculated. Having the ensemble average of eq. 2.11 thus evaluated at several values of  $\lambda$  (or  $T^*$ ), the integral of eq. 2.13 can be performed, and  $\Delta A$  can be calculated.

### 2.3 Basics of Computer Simulation Methods

Computer simulation plays a key role in the research of the structure of disordered condensed phase systems on a statistical mechanical basis. In computer simulation, we strive to model the properties of a real system as accurately as possible. Hence, we generate systems of different microstates and calculate the macroscopic properties of the system as an average over all of its microstates. The thermodynamic properties can be determined as the functions of the positions ( $\mathbf{q}$ ) and the momenta ( $\mathbf{p}$ ) of the particles. The macroscopic properties can be calculated as the long-term time average. In the case of an ergodic system, all accessible microstates are equally probable over a long period of time. Therefore, if we are interested only in momenta of the independent properties, this time average will be equal to the average over independent microstates (ensemble average). From the approach of classical mechanics, the Hamiltonian ( $H$ ) of the system can be defined as the sum of the kinetic ( $K$ ) and potential energy ( $U$ ):

$$H(\mathbf{q}, \mathbf{p}) = K(\mathbf{p}) + U(\mathbf{q}) \quad (2.15)$$

The kinetic energy of a system containing  $N$  particles can be given as:

$$K = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} \quad (2.16)$$

where  $m_i$  is the mass of the  $i$ -th particle. The calculation of the potential energy ( $U$ ) is a more complex task, it can be done through the interactions between the particles of the system as the sum of the single molecule terms, pair interaction energies, triplets, and so on. For an  $N$  particles system, the potential energy can be given as:

$$U = \sum_i u_1(\mathbf{q}_i) + \sum_i \sum_{j>i} u_2(\mathbf{q}_i, \mathbf{q}_j) + \sum_i \sum_{j>i} \sum_{k>j>i} u_3(\mathbf{q}_i, \mathbf{q}_j, \mathbf{q}_k) + \dots + \sum_{k>j>i} u_N(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N) \quad (2.17)$$

where the first term is the effect of an external potential. The second term is the sum of the energy coming from the pair interactions of all possible particle pairs, which only depends on the distances between the pairs. The third term contains the additional energy of particle triplets with respect to the sum of the three pair interaction energy terms acting between any pairs of these three particles. Similarly, multiparticle energy terms represent the additional energy with respect to all terms defined between the particles of any of its subset. To reduce the computational and time requirements of the simulation, the higher order energy terms are truncated. However, the contribution of these members to the total potential energy is not negligible. Therefore, an effective pair potential ( $u_2^{\text{eff}}$ ) is used to take into account also the contributions of the higher energy terms in an average way:

$$U \approx \sum_i u_1(\mathbf{q}_i) + \sum_i \sum_{j>i} u_2^{\text{eff}}(\mathbf{r}_{ij}) \quad (2.18)$$

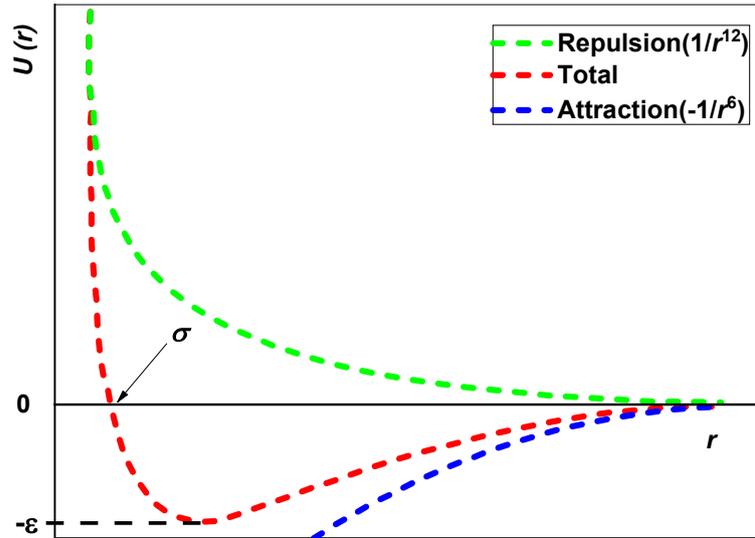
where  $r_{ij}$  is the distance between the  $i$  and  $j$  particles.

One of the simplest effective pair potentials is the hard-sphere potential (stepwise model) with a given  $d$  diameter. If the distance between two particles is less than  $d$ , the value of the pair potential is infinite. Otherwise, if the distance is greater than  $d$ , the pair potential becomes 0. To describe dispersion interactions in real systems, we need a finer or more realistic model. Continuous pair potentials, such as the Buckingham-, Kihara-, or Mie potential take into account longer-term attraction in addition to short-term

repulsion between particles.<sup>165,166</sup> The most commonly used version of the Mie potential is the Lennard-Jones potential:

$$U_{LJ}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] \quad (2.19)$$

where  $r$  is the distance between two particles, while the  $\sigma$  and  $\epsilon$  Lennard-Jones parameters are the finite distance at which the inter-particle potential is zero and the depth of the potential well, respectively. The first term in the square brackets stands for the repulsion, which decreases as a function of the distance by  $r^{-12}$  and the second term is the attraction decreasing as a function of  $r^{-6}$  (see Figure 2.3). At the short-range interaction, the nuclear and electronic repulsions and the electronic kinetic energy are the determining components, while the attractive dispersion interaction dominates at longer distances.



*Figure 2.3: Representation of the Lennard-Jones pair potential with the repulsive and attractive interactions;  $\sigma$  is the finite distance at which the inter-particle potential is zero and  $\epsilon$  is the depth of the potential well.*

In the case of molecular systems, we cannot treat the molecules as point-like particles during the potential energy calculation; we need to consider their shape and orientation as well. To overcome this problem, we build up the molecules from interacting sites, which can be atoms, groups of atoms, or virtual sites. Different interaction sites may have different potential parameters, in this case the Lennard-Jones parameters can be calculated with using the Lorentz-Berthelot rule:<sup>138</sup>

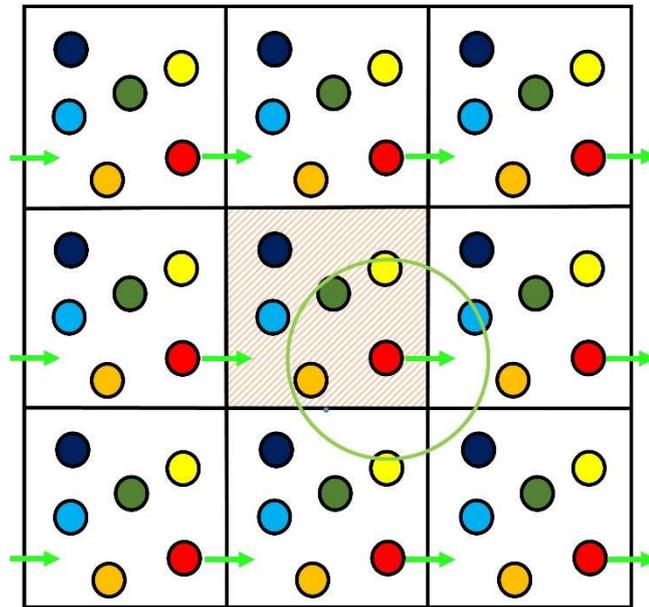
$$\epsilon_{AB} = \sqrt{\epsilon_A \epsilon_B} \quad ; \quad \sigma_{AB} = \frac{\sigma_A + \sigma_B}{2} \quad (2.20)$$

The interaction between charged particles can be described with the Coulomb potential:

$$U_c = \frac{1}{4\pi\epsilon_0} \frac{q_A q_B}{r} \quad (2.21)$$

where  $\epsilon_0$  is the vacuum permittivity and  $q_A$  and  $q_B$  are the point charges of the two interaction sites which are in distance  $r$ .

In a computer simulation of a condensed phase, we obtain information about the structural and thermodynamical properties of the macroscopic system from the simulation of a few thousand particles. This number is several orders of magnitude smaller than the size of a macroscopic system. As a consequence, the fraction of the particles that are at the surface of the simulation box is orders of magnitude higher than that of the surface particles in a macroscopic system. Thus, in computer simulations, the occurrence of artificial physical phase boundaries have to be avoided, which can be done by applying Periodic Boundary Conditions (PBC).<sup>167</sup> The particles building up the system are placed into a space-filling unit cell that is repeated infinitely in every direction. Thus, any molecule that leaves the cell at one side enters from the opposite side (see Figure 2.4).



*Figure 2.4: Illustration of the Periodic Boundary Conditions (PBC) in the case of cubic basic simulation box. The original cell takes place in the middle of the figure, the copies are the periodic boundary images. The green arrows show the movement of the periodic images of molecules traversing the box edge and the green circle represents the maximum interaction cutoff length ( $L/2$ ).*

The use of the PBC introduces an artificial periodicity into the system. Thus, to avoid treating more than one replicas of any particles, distances within the basic box should not

be considered beyond the radius of the largest inscribed sphere (in the case of a cubic box with the edge length  $L$  this is  $L/2$ ). This restriction affects also the calculation of the pair interactions. Since the dispersion interaction approaches zero rather quickly with increasing distance, this restriction has a minor effect in this case. On the other hand, electrostatic interactions extend to a considerably longer range of distances, therefore the missing long-range term of the electrostatic interaction has to be taken into account by appropriate correction methods, such as reaction field correction or Ewald summation.<sup>168</sup> In the reaction field correction method, the Coulomb interactions are treated explicitly within the cut-off sphere, while the system beyond this sphere is treated as a dielectric continuum bearing the uniform dielectric constant  $\epsilon_{RF}$ . The molecules inside the sphere polarize this continuum, which generates a reaction field (Onsager reaction field).<sup>169–171</sup> In the case of pairwise additive interaction and neutral particles, the interaction of the central particle with this continuum can be taken into account by modifying its Coulomb interaction with the particles inside the sphere as:<sup>172</sup>

$$U_C = \frac{1}{4\pi\epsilon_0} \frac{q_A q_B}{r_{ij}} \left[ 1 + \frac{\epsilon_{RF} - 1}{2\epsilon_{RF} + 1} \left( \frac{r_{ij}}{R_C} \right)^3 \right] \quad (2.22)$$

where  $R_c$  is the radius of the cut-off sphere. The other possible way of taking the effect of the long-range electrostatic interactions into account is to calculate the interaction of the central particle with all possible images of the other particles up to infinity. The number of calculations required to do this can be substantially reduced by using the Ewald summation.<sup>138</sup> The Ewald summation is a special case of the Poisson summation formula, replacing the summation of the interaction energies in real space with an equivalent summation in the reciprocal space (Fourier space). The potential interactions are separated into two parts, which represent the short-range interactions and the long-range interactions. The short-range part can be summed easily, and it quickly converges in real space, while the summation of the long-range terms is more problematic. The advantage of this approach is the rapid convergence of the Fourier-space summation compared to its real-space equivalent when the real-space interactions are long-range. The summation can be further accelerated using the particle mesh Ewald (PME) method.<sup>168</sup> As in the case of the Ewald summation, the direct summation of the interaction potential between point charges is divided into two parts: short-range (in real space) and long-range (in Fourier space) interactions. To evaluate the Fourier transform of the charge density field, the Fast Fourier Transformation (FFT) method<sup>173</sup> can be

applied, which requires the evaluation of the density field on a mesh as a discrete lattice in space. Charge is interpolated to grid points, using a weighting function.

Average properties of the system can be determined by stochastic Monte Carlo and molecular dynamics simulation. Monte Carlo calculates the average over an ensemble of microstates of the entire configurational space, while the molecular dynamics simulation uses the time average of equilibrium configurations, generated in a suitably long trajectory, to determine the macroscopic properties of interest. Sampling the entire phase space is not possible for practical reasons, therefore in a computer simulation a statistically relevant ensemble of sample configurations is generated to calculate the time or ensemble average over the microstates. The ergodic hypothesis states the equality of the time and ensemble averages over a set of microstates for a system in equilibrium.

$$\langle Q \rangle_t = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau Q(\mathbf{q}(t), \mathbf{p}(t)) dt = \int Q(\mathbf{q}, \mathbf{p}) f(\mathbf{q}, \mathbf{p}) d\mathbf{q} d\mathbf{p} = \langle Q \rangle_\Gamma \quad (2.23)$$

where  $\mathbf{p}(t)$  and  $\mathbf{q}(t)$  are the momenta and the position of the particles,  $t$  denotes the time and  $\tau$  is the duration of the simulation.  $Q(\mathbf{p}, \mathbf{q})$  is the value of the quantity at the  $\Gamma = (\mathbf{p}(t), \mathbf{q}(t))$  microstate.  $\langle Q \rangle_t$  is the time average and  $\langle Q \rangle_\Gamma$  is the ensemble average of  $Q$  over the microstates, while  $f(\mathbf{p}, \mathbf{q})$  is the probability density function of the microstates.

As a consequence, the Monte Carlo and molecular dynamics methods are equivalent in calculating the momenta-independent equilibrium properties of a system (which significantly reduces the computational cost of the simulation). In Monte Carlo simulations sample configurations are generated stochastically in the phase space to produce an ensemble of the system in different microstates.<sup>138</sup> Following random changes made in the system, the resulting configurations are accepted or rejected according to appropriately chosen acceptance criteria. Since the different sample configurations do not belong to a given trajectory of the system, in other words, the momenta of the particles are not considered in a Monte Carlo simulation, this method cannot be used for the simulation of properties depending on the momenta of the particles, time dependent processes or non-equilibrium systems, only the static characteristics of equilibrium systems can be calculated this way. On the other hand, in molecular dynamics simulations, the motion of the particles is followed by solving the Newtonian equations of motion in every step of the simulation for all of the particles building up the system. Thus, new configurations are calculated based on the movement of the particles and the forces acting on them. Due to the multiparticle nature of such simulations, the equations

of motion are solved numerically. The trajectory is calculated in small time steps ( $\Delta t \sim 2$  fs), and the macroscopic property of interest is obtained as a time average. Since, in principle, molecular dynamics is a deterministic method, the initial state of the system uniquely defines the entire set of microstates through which the system goes in the simulation.

In computer simulations for example liquid/vapor equilibrium or solid surface adsorption, results are highly dependent on the different set up conditions, e.g., size of simulation cell, number of particles, cut-off radius, time of simulations, etc. Investigating the long range part of the electrostatic interaction in an inhomogeneous and anisotropic system at the surface adsorption on ice phase is not a simple task. By utilizing the standard Ewald summation method<sup>138</sup>, the simulation would produce an infinite stack of ice and vapor layers. However, if the reaction field correction technique<sup>138,172,174</sup> is applied the system is consisted of two phases of significantly different dielectric constants. Hantal et al.<sup>175</sup> investigated this problem by using reaction field correction and setting the dielectric constant of the continuum beyond the cut-off sphere ( $\epsilon_{RF}$ ) to infinity (conducting boundary conditions). In the other set of simulation no long range correction has been used, which corresponds to the reaction field correction method with  $\epsilon_{RF} = 1$ . These two parameter sets represent the lower and upper estimation of the long range electrostatics effect. During the analysis of the density profiles and surface orientations of the adsorbed molecules on the surface, no significant differences were found. These findings are particularly important for the simulation of formamide adsorption at the ice surface.

Due to the finite size of the simulated interface, the mass and electron density profiles can show oscillations in some cases at liquid-liquid systems' aqueous phase.<sup>176</sup> Toxvaerd and Stecki<sup>177</sup> work demonstrated that systems with different size (number of particles 1020 and 9216) show no significant deviation in their density profiles. Chapela et al.<sup>178</sup> investigated the liquid-vapor surface dependence on the system size. They made computer simulations with systems containing 255, 1020 and 4080 molecules. They found that the surface thickness increases with the size of the simulation. The 1020 molecule system was larger by 19 % than the smallest, and the 4080 molecule system thickness was 4.3 % greater than the middle one. The source of this deviation was the repetition of the coordinates of the molecules in the x-y directions, rather than being statistical error due to the different numbers of molecules employed. The restriction in surface area results in the damping of all surface waves of wavelength greater than the

width of the box, and this will lead to the differences in the density profiles. They found the limit beyond, which this effect will become unimportant, is the region of the applied largest system. Holcomb et al.<sup>179</sup> continued Chapela's work and investigated the size effects, cut-off effects of the simulation of the liquid-vapor interface. They came to the following conclusions. Size effects are related to the number of particles, the area of the interface and the thickness of the slab. To overcome errors emerging from the effect of the system size, they recommend that systems should contain at least 1000 particles. The interfacial area has to be at least four times the square of the cut-off to avoid periodic boundary effects. Furthermore, the cut-off has to be large enough, at least  $4.4 \sigma$ , so that the bulk densities are no longer changing significantly. Chen<sup>180</sup> results on surface tension, show that the finite-size effect is significant in small surface areas only. For the systems of  $L > 10\sigma$  show essentially no difference from an infinite system. This means, the contribution of the thermal fluctuations of wavelengths longer than  $10\sigma$  to the surface tension becomes negligible in the systems of  $L > 10\sigma$ . Based on these results, with the selection of the recommended system size for the investigation of the liquid-vapor interface at formamide-water mixtures, the effects of finite system size and periodicity are become negligible.

## 2.4 Monte Carlo Simulation

In each step of the Monte Carlo simulation, a randomly selected particle is tried to be moved in a random direction and rotated with a random angle. The displacement of the randomly chosen molecule within a given maximum distance can be described with the following equation:

$$q'_{ij} = q_{ji} + \Delta q^{\max} \xi_i \quad (2.24)$$

where  $q'_{ij}$  is the coordinate of the  $i$ -th particle after displacement and  $q_{ij}$  is the coordinate of the  $j$ -th particle before displacement,  $\Delta q^{\max}$  is the maximum permitted distance of the displacement, and  $\xi_i$  is a random number generated between 0 and 1. The change of the orientation of a randomly selected molecule can be done in a similar way, by rotating it with a random angle less than a given  $\Delta\alpha^{\max}$  maximum value about a randomly chosen space-fixed axis (parallel to an edge of the simulation box) that goes through a suitably selected point (e.g., center) of this molecule. In case the system is built up from flexible

molecules, the geometry of each molecule also has to be changed randomly during the simulation.<sup>138,170,181</sup>

### 2.4.1 Canonical ( $N, V, T$ ) Monte Carlo

In the case of the canonical ensemble, the number of the particles as well as the volume and temperature of the system are constant, and, in equilibrium, the energy of the system follows the Boltzmann distribution. In other words, the canonical ensemble can be approached as the available states of a system are in contact with a heat bath, and hence the energy is no longer of a constant value. The partition function, which describes the statistical properties of a system in thermodynamic equilibrium can be written as<sup>181</sup>:

$$Q(N, V, T) = \frac{1}{N!} \frac{1}{h^{3N}} \int \int \exp(-\beta E(\mathbf{q}^N, \mathbf{p}^N)) d\mathbf{q}^N d\mathbf{p}^N \quad (2.25)$$

where  $N$  number of particles are represented by their position coordinates  $\mathbf{q}^N$  and momentum coordinates  $\mathbf{p}^N$ ,  $\beta$  is the Boltzmann-constant,  $h$  is the Planck's constant and  $E$  is the total energy, which can be decomposed into the sum of the potential energy ( $U$ , only depending on the position coordinates) and the kinetic energy ( $K$ , only depending on the momentum coordinates):

$$E(\mathbf{q}^N, \mathbf{p}^N) = U(\mathbf{q}^N) + K(\mathbf{p}^N) \quad (2.26)$$

The expected value of a momenta-independent measurable physical quantity of a system ( $M$ ) which is thus accessible by Monte Carlo simulation can be given as:

$$\langle M \rangle = \frac{\int M(\mathbf{q}^N) \exp(-\beta U(\mathbf{q}^N)) d\mathbf{q}^N}{\int \exp(-\beta U(\mathbf{q}^N)) d\mathbf{q}^N} \quad (2.27)$$

Based on this equation, equally probable random configurations are created and weighted with the Boltzmann factor. Due to exponential function in the weight, only the configurations corresponding to the lowest values of  $U(\mathbf{q}^N)$  contribute non-negligibly to the average. Further, the necessary number of microstates is far too high to be calculated in a simulation during a reasonable time. To remedy the problem, Metropolis *et al.* applied a solution based on the use of configurations bearing relatively low  $U$  value that are chosen with higher probability.<sup>182</sup> Hence, sample configurations are generated with a probability proportional to their Boltzmann factor followed by unweighted averaging. In the Metropolis sampling, a chain of random configurations is generated by the

displacement or rotation of a particle. After a random change is made in the configuration, the potential energy of the new configuration is calculated. If the energy decreases, the new state is always accepted, but if the energy increases, the trial move is only accepted by the following probability:

$$P_{accepted} = \min(1, \exp(-\beta\Delta U(\mathbf{q}^N))) \quad (2.28)$$

where  $\Delta U$  is the difference between the potential energy of the new and the previous configuration.

### 2.4.2 Grand-Canonical ( $\mu, V, T$ ) Monte Carlo

In the grand canonical ensemble, our system can exchange both energy and particles with the surrounding environment, hence, the chemical potential of the particles ( $\mu$ ) as well as the temperature ( $T$ ) and volume ( $V$ ) of the system are constant. This ensemble is ideal for investigating adsorption processes. The grand canonical partition function can be given as:<sup>181</sup>

$$Q(\mu, V, T) = \sum_{N=1}^{\infty} \frac{1}{N!} \frac{1}{h^{3N}} \int \int \exp(-\beta(E^N(\mathbf{q}^N, \mathbf{p}^N) - N\mu)) d\mathbf{q}^N d\mathbf{p}^N \quad (2.29)$$

A sum has to be used to account for the changing number of the particles in the system. The expected value of a measurable physical quantity of a system ( $M$ ) can be described as:

$$\langle M \rangle = \frac{\sum_{N=1}^{\infty} \frac{1}{N!} \left(\frac{V}{\Lambda^3}\right)^N \int M^N(\mathbf{s}^N) \exp(-\beta(U^N(\mathbf{s}^N) - N\mu)) d\mathbf{s}^N}{\sum_{N=1}^{\infty} \frac{1}{N!} \left(\frac{V}{\Lambda^3}\right)^N \int \exp(-\beta(U^N(\mathbf{s}^N) - N\mu)) d\mathbf{s}^N} \quad (2.30)$$

where  $V$  is the volume of the system,  $\mu$  is the chemical potential and  $\Lambda$  is the thermal de Broglie wavelength, given as:

$$\Lambda = \sqrt{\frac{h^2 \beta}{2\pi m}} \quad (2.31)$$

where  $m$  stands for the mass of a single particle. Furthermore, it is convenient to apply the scaled coordinates  $s$ :

$$s = \mathbf{q}V^{-\frac{1}{3}} \quad (2.32)$$

During the grand canonical Monte Carlo simulation, in addition to the particle displacement steps, which was used also in the canonical ensemble, the number of the particles in the system studied can be changed by deleting or creating a particle in the system. The trial moves are accepted with the following probability:<sup>138,183</sup>

$$P_{accepted} = \min \left[ 1, \exp \left( \Delta \left( N \ln \left( \frac{V}{\Lambda^3} \right) - \ln(N!) - \beta(U(\mathbf{s}^N) - N\mu) \right) \right) \right] \quad (2.33)$$

In practice, it is advisable to change the number of particles in one step by  $\Delta N = \pm 1$ . If a new particle is inserted into a randomly selected point of our system ( $\Delta N = + 1$ ), the probability of the acceptance of this step reduces to:

$$P_{accepted} = \min \left[ 1, \exp \left( \ln \left( \frac{V}{\Lambda^3} \right) - \ln(N + 1) - \beta\Delta U + \beta\mu \right) \right] \quad (2.34)$$

Similarly, in the case when a randomly selected particle is removed from the system ( $\Delta N = - 1$ ), the probability of acceptance simplifies to:

$$P_{accepted} = \min \left[ 1, \exp \left( -\ln \left( \frac{V}{\Lambda^3} \right) + \ln N - \beta\Delta U - \beta\mu \right) \right] \quad (2.35)$$

If the density of our system is high, the insertion of a new particle can be very ineffective, because the inserted particle would almost certainly overlap with the already existing particles in the system. This problem can be avoided by using cavity-biased insertions.<sup>184,185</sup> By this method, particle insertions are only attempted in such points of the simulation box that are located in a spherical cavity the radius of which is larger than a pre-defined value, and hence no particles are found within this radius.

## 2.5 Molecular Dynamics Simulation

In molecular dynamics simulations, configurations are generated by solving the Newtonian equations of motions for a system built up from  $N$  particles. The equations of motion can be described by the Lagrangian form:

$$\frac{d}{dt} \left[ \frac{\partial L(\mathbf{q}, \mathbf{p})}{\partial \mathbf{p}} - \frac{\partial L(\mathbf{q}, \mathbf{p})}{\partial \mathbf{q}} \right] = 0 \quad (2.36)$$

where  $L(\mathbf{p}, \mathbf{q})$  is the Lagrangian function, which can be defined as the sum of the kinetic and the potential energy:  $L(\mathbf{p}, \mathbf{q}) = K(\mathbf{p}) + U(\mathbf{q})$ . The force ( $F_i$ ) acting on a the  $i$ -th particle, having the mass  $m_i$ , can be expressed as:

$$F_i = m_i \ddot{\mathbf{q}}_i \quad (2.37)$$

The equations of motion in the Hamiltonian formalism can be written as:

$$\dot{\mathbf{q}}_i = \frac{\mathbf{p}_i}{m_i} \quad , \quad \dot{\mathbf{p}}_i = F_i \quad (2.38)$$

Having calculated all the forces between the particles, the Newtonian equations of motion can be integrated. Several methods have been developed to solve the above equations of motion; these algorithms alternately calculate the force and solve the equations of motion based on the accelerations obtained from the updated forces. In each step of the simulation, particle displacements are calculated by solving a set of second-order differential equations consisting of  $3N$  equation. One of the most well-known and widely used algorithms is the Verlet algorithm. In the first step of the computation, particle coordinates ( $q$ ) are written as their Taylor expansion according to time ( $t$ ):<sup>138,183,186</sup>

$$q(t + \Delta t) = q(t) + v(t)\Delta t + \frac{f(t)}{2m}\Delta t^2 + \frac{\Delta t^3}{3!}\ddot{q} + \varphi(\Delta t^4) \quad (2.39)$$

$$q(t - \Delta t) = q(t) - v(t)\Delta t + \frac{f(t)}{2m}\Delta t^2 - \frac{\Delta t^3}{3!}\ddot{q} + \varphi(\Delta t^4) \quad (2.40)$$

The sum of the two equations yields:

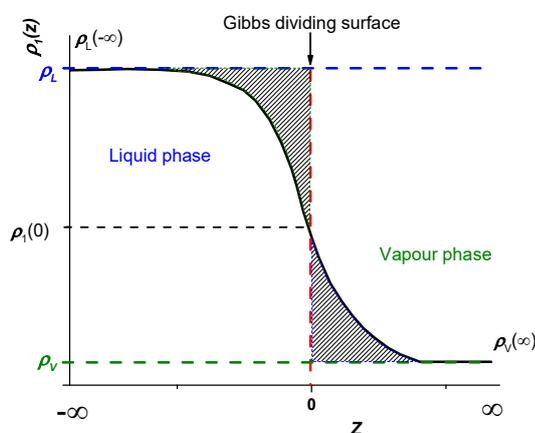
$$q(t + \Delta t) + q(t - \Delta t) = 2q(t) + \frac{f(t)}{m}\Delta t^2 + \varphi(\Delta t^4) \quad (2.41)$$

The advantage of the Verlet algorithm is that in generating the next positions of the particles it does not use their velocities. The velocities of the particles, if needed, can be calculated from their positions as:

$$v(t) = \frac{q(t + \Delta t) - q(t - \Delta t)}{2\Delta t} + \varphi(\Delta t^2) \quad (2.42)$$

## 2.6 The Identification of Truly Interfacial Molecules (ITIM)

The different properties of the interfaces between two phases have been in the center of physico-chemical investigations for decades. Various experimental methods (e.g., sum frequency generation and second harmonic generation spectroscopy, X-ray, and neutron reflectometry) have been developed to help the molecular level understanding by selectively probing the region of the interface.<sup>187</sup> Experimental works have been complemented by computer simulation studies to provide an even better, atomistic resolution model of the systems of interest. However, the exact determination of the full set of the interfacial particles in a computer simulation is a difficult task. When simulating an interface between two non-crystalline phases, one has to take into account the fact that the surface is not flat on the microscopic scale, instead, it is corrugated by capillary waves. Although the importance of this fact was already recognized and taken into account in the very first simulations of liquid-vapor interfaces,<sup>188,189</sup> in the vast majority of the early simulations of such systems the effect of the capillary waves was simply neglected, and the flat Gibbs dividing surface was used in the analyses instead of the real, wavy surface of the liquid phase. The “dividing surface” as a simple approach was introduced by Gibbs, and it assumes a molecularly sharp, planar interface.<sup>187,190</sup> Further, the bulk properties are assumed to remain constant up to the dividing surface separating the two bulk phases. In one-component systems, the Gibbs dividing surface is chosen as such that the particles have no surface excess, as illustrated in Figure 2.5. Further, in the case of mixtures it is conventionally chosen in such a way that the surface excess of the major component (i.e., solvent) is zero.



*Figure 2.5: The schematic representation of the Gibbs dividing surface at liquid–vapor interface in the case of one-component system, showing the variation in density  $\rho_1(z)$  as the function of  $z$  coordinate, liquid density  $\rho_L$  is the liquid density and  $\rho_V$  is the vapor density*

This treatment was later shown to introduce a systematic error of unknown magnitude into several calculated properties (e.g., structure, composition) of the interface,<sup>191,192</sup> and this systematic error can even propagate to the thermodynamic properties of the system.<sup>193</sup>

Capillary Wave Theory (CWT) can describe the macroscopically planar liquid-liquid and liquid-vapor interfaces on the molecular scale more accurately.<sup>194</sup> It is a widely accepted fact that the interface of liquids are perturbed by waves parallel to the surface. The spreading of a wave with  $\lambda$  wavelength can be described with the term of gravitational waves and the capillary waves as the following equation

$$vc^2 = g + \frac{\gamma}{m\rho}v^2 = g \left[ 1 + \frac{1}{2}(av)^2 \right] \quad (2.43)$$

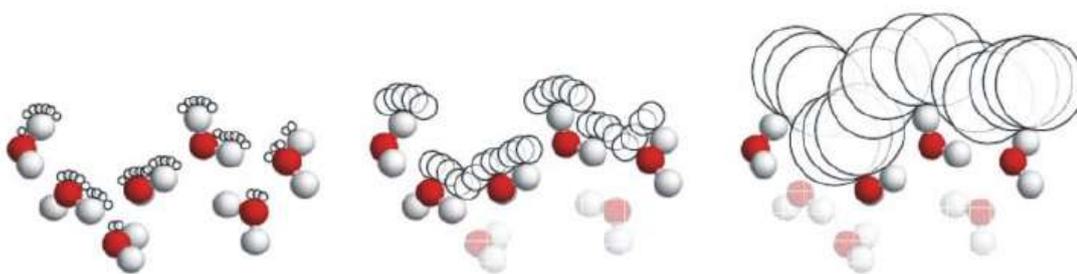
where  $v$  is the wavenumber,  $c$  is the phase velocity,  $g$  is the gravitational constant,  $\gamma$  is the surface tension and  $a$  is the capillary length. The length of the capillary is the size of the transition range between the gravitational and capillary waves. In this theory, the interface is well defined locally, but the position of the surface  $z = \xi(x, y)$  fluctuates along the  $z$  axis which is perpendicular to the surface. To overcome the long-wave fluctuations, one can define the intrinsic density profile,<sup>195</sup> which is the density profile using a local, instantaneous reference frame fixed to this real, capillary wave corrugated interface. Thus, we can write the intrinsic density profile as

$$\rho(z) = \left\langle A^{-1} \sum_{i=1}^N \delta(z - z_i + \xi(x_i, y_i)) \right\rangle \quad (2.44)$$

where  $(x_i, y_i, z_i)$  is the position of the  $i$ th particle,  $\xi(x_i, y_i)$  is the fluctuation of the surface and  $A$  is the surface area. Proper treatment of the capillary waves in a computer simulation requires the determination of the real, capillary wave corrugated intrinsic liquid surface (and define distances relative to this surface rather than to the (flat) Gibbs dividing surface) and correspondingly, the determination of the full list of the truly interfacial molecules (i.e., the ones that stay right at the boundary of the two phases). Following the pioneering work of Chacón and Tarazona,<sup>196</sup> several approaches have been proposed<sup>191,197-199</sup> in the last two decades to determine the intrinsic surface at various types of interfaces.

In 2007, Pártay<sup>191</sup> and her colleagues presented a method to identify the truly interfacial molecules (ITIM) at interfaces in computer simulations with molecular scale

resolution. This method turned out to be an excellent compromise between computational cost and accuracy.<sup>199</sup> The ITIM method is based on probe spheres with appropriately chosen radius, which are moved from the bulk of one phase to the direction of the other phase along straight test lines that are perpendicular to the interface. When the probe sphere touches a molecule of the other phase, it stops, and the molecule is marked as being interfacial. The ITIM method is illustrated in Figure 2.6. This procedure needs to be repeated along a dense enough set of test lines to identify all the interfacial molecules. The intrinsic surface can then be defined through the positions of the interfacial atoms (i.e., the ones that stopped the probe).



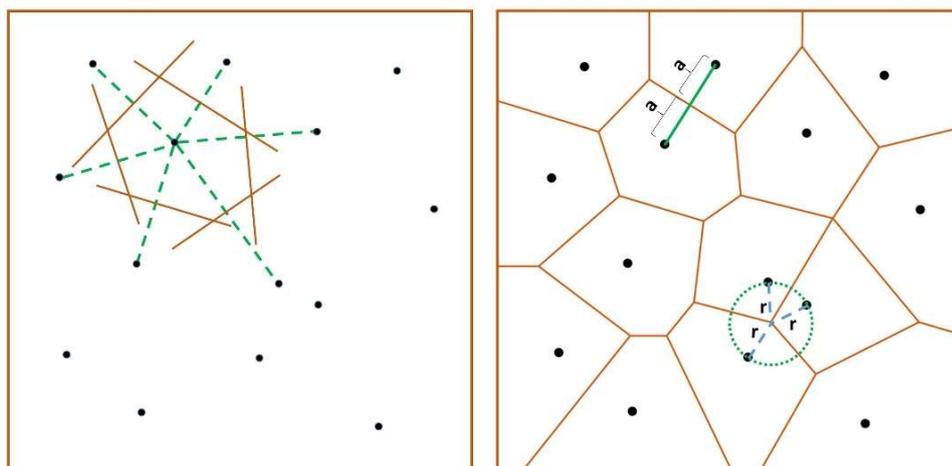
*Figure 2.6: Representation of the ITIM method with different probe sphere radius ( $\alpha$ ) in the case of water surface.  $\alpha = 0.5 \text{ \AA}$  (left),  $\alpha = 2.0 \text{ \AA}$  (middle) and  $\alpha = 10.0 \text{ \AA}$  (right)<sup>191</sup>*

To apply the ITIM method successfully, the value of several parameters has to be chosen appropriately. The most important among these parameters is the radius of the probe sphere, which is clearly a free parameter, but its choice strongly influences the final set of the molecules that are identified as being interfacial. Thus, if the probe sphere radius is chosen to be zero, a huge number of molecules from the bulk liquid phase could also be considered as being interfacial. On the other hand, when the probe sphere radius approaches infinity, the number of the interfacial molecules detected will be only one, i.e., the one located farthest from the bulk phase along the macroscopic interface normal axis, and no other molecules can be touched by the probe even if they are located at the interface. It should be emphasized, however, that the presence of such a free parameter is inherent in any method aiming at determining the intrinsic liquid surface, as the choice of this parameter defines the length scale on which the surface is looked at.<sup>199</sup> Therefore, the probe sphere radius is conveniently chosen to be of the same order of magnitude as the size of the particles building up the phase considered. The effect of the probe sphere radius is shown in Figure 2.6.

The other important parameter is the arrangement and density of the test lines that guide the path of the probe sphere. The test points in the plane of the interface where the test lines traverse are conventionally generated along a square grid. To obtain a sufficiently accurate result in the mapping of the surface, the neighboring test points need to be considerably closer to each other than the probe sphere radius. The choice of the number of test points should be based on a reasonable compromise between the precision of the calculation and the used computing capacity. Former studies indicated that, in the case of small molecules, the distance of the neighboring test points of about 0.5 Å represents a sufficient choice.<sup>199</sup> Finally, for detecting the touching position with the probe, the radius of each atom in the studied phase also has to be defined. A physically reasonable choice for the atomic radii values could be the use of the van der Waals radii or half of the Lennard-Jones distance parameter of the atoms.<sup>191,199,200</sup>

## 2.7 Voronoi Analyses

Following the identification of the surface molecules in a system, such in the case of the formamide-water mixtures, the Voronoi method has already proved to be suitable to study the structure of the surface layer through the distribution of the molecules in it.<sup>145,151,201,202</sup> In mathematics, the Voronoi diagram or Voronoi tessellation is a type of metric space division. The established regions determined by distances to a given set of objects in the space.<sup>203</sup> In a two-dimensional system of seeds, the Voronoi cell of a given seed is the locus of points, which are closer to this seed than to any other one. In practice, polygons can be obtained by drawing perpendicular bisectors between every two neighboring seeds. These polygons are the Voronoi cells and their boundaries are the line segments of the Voronoi diagram. The line segments are equidistant to the two nearest seeds. The determination of Voronoi cells, in the case of random points in the plane, is illustrated in Figure 2.7. This method can be applied for points in the three-dimensional space as well. However in that case, the bisectors will be planes, instead of lines, and the Voronoi cells will be polyhedra instead of polygons.



*Figure 2.7: Illustration of the Voronoi cells formation: the dashed green lines (left diagram) indicate the distance between the seeds, while the lines perpendicular to them indicate the bisectors between the two adjacent points. Image of Voronoi cells and relationships between the distances of neighboring seeds. (right diagram)*

This method beside a lot of other scientific application<sup>204,205</sup> also suitable for investigating the surface structure of condensed systems. The seeds can represent the molecules at the surface and a Voronoi cell stands for the surface area belonging to one molecule. Hence, the self-aggregation of the molecules can be detected in multicomponent systems by calculating the Voronoi cell area distribution for one of the components. In case of a homogeneous distribution, Voronoi cells will have similar size and shape and their area distribution will have Gaussian shape. When the distribution of a component is not homogeneous, Voronoi cells will become aspherical and will show significant differences in their areas. For example, the cells' areas at the edge of the aggregation will be considerably bigger comparing to the cells' areas at the middle. Consequently, the area distribution of the selected component will show an exponentially decreasing tail.<sup>206</sup>

### 3. Computational Details

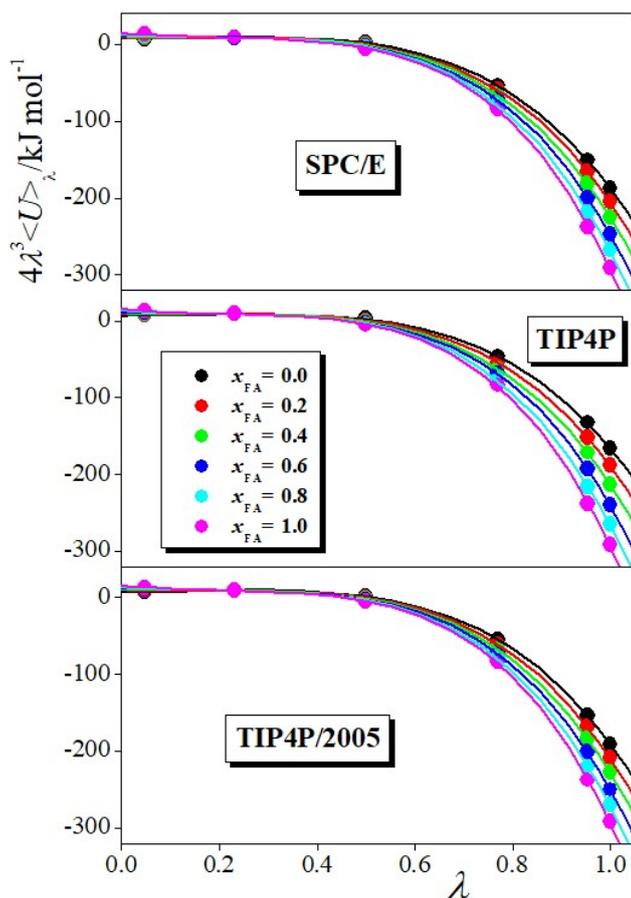
In this chapter, the details of the computer simulations performed are presented in detail. First the Monte Carlo simulation carried out to investigate the thermodynamics of mixing of water and formamide with different formamide-water model pairs are described. This is followed by presenting the technical details of the molecular dynamic simulation of the vapor-liquid interface of water-formamide mixtures of different compositions. Finally, the grand canonical Monte Carlo investigation of the adsorption of formamide molecules at the surface of both crystalline ( $I_h$ ) and amorphous (LDA) ice under atmospheric and interstellar conditions, is described. It should be noted that, in the latter two studies, we used the model combination that had turned out to be the best in describing the thermodynamics of mixing of the two molecules in the first study.

#### 3.1. The Mixing of Different Models of Formamide and Water

The values of  $A_A$ ,  $A_B$ , and  $A_{A+B}$  in eq. 2.7 have been calculated using the method of thermodynamic integration, which is described in the chapter 2 in detail. Monte Carlo simulations have been performed at six different (virtual) temperatures, corresponding to the  $\lambda$  values of 0.046911, 0.230765, 0.5, 0.769235, 0.953089, and 1. The first five  $\lambda$  values correspond to the 5-point Gaussian quadrature,<sup>207</sup> while simulations at the non-fictitious state  $\lambda = 1$  have been performed in order to also evaluate  $U_A$ ,  $U_B$ , and  $U_{A+B}$ . To perform the integral of eq. 2.13, a fourth order polynomial has been fitted to the  $4\lambda^3 \langle U_Y \rangle_\lambda$  vs.  $\lambda$  data points, and the fitted polynomial has been integrated analytically. These points, along with the fitted polynomial, are shown in Figure 3.1, as obtained in selected systems. The integrand of eq. 2.13 has been evaluated by performing a Monte Carlo simulation at each  $\lambda$  value considered for each system. The real temperature of the systems, corresponding to  $\lambda = 1$  has been 298 K. In every simulation, 512 molecules have been placed in a cubic basic simulation box, the edge length of which,  $L$ , has been set in order to reproduce the experimental mass density<sup>208</sup> of the system,  $\rho$ . Besides the two neat systems, mixtures of nine different compositions, corresponding to the formamide mole

fraction ( $x_{FA}$ ) values of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9 have been simulated. The characteristics of the different systems simulated are collected in Table 3.1.

All simulations have been performed with the program MMC.<sup>209</sup> In a Monte Carlo step, a randomly chosen molecule has been randomly displaced to a maximum distance of 0.25 Å, and randomly rotated by no more than 10°. The geometry of all molecules has been kept unchanged (rigid model) in the simulations. At least 25% of the trial moves have been successful in every case. All interactions have been truncated to zero beyond the center-center cut-off distance of 12.4 Å, the O and N atoms being regarded as the centers of the water and formamide molecules, respectively. The long range part of the electrostatic interaction has been accounted for using the reaction field correction method<sup>138,169,172</sup> under conducting boundary conditions. The systems have been equilibrated through  $5 \times 10^7$  Monte Carlo steps; the value of  $\langle U_\gamma \rangle$  has then been evaluated in a consecutive,  $10^8$  Monte Carlo steps.



**Figure 3.1:** *Integrand of the thermodynamic integration (eq. 2.13), obtained at six  $\lambda$  points for the combination of the CHARMM model of formamide with the SPC/E (top panel), TIP4P (middle panel), and TIP4P/2005 (bottom panel) models of water (full circles), together with the fourth order polynomials fitted to these data (solid curves) at the formamide mole fraction values of 0 (black), 0.2 (red), 0.4 (green), 0.6 (dark blue), 0.8 (light blue) and 1.0 (magenta).*

The mixing of five different models of formamide with three widely used water models had been considered. Since simulations have been performed at six (virtual) temperatures and nine compositions for each model combination and also for the neat systems, the total number of simulations performed is 858. The formamide potentials considered include the united atom OPLS model,<sup>106</sup> the all-atom OPLS/AA model<sup>210</sup> modified by Pérez de la Luz et al.<sup>116</sup> (referred to here as OPLS/AA\_mod), the model belonging to the CHARMM force field,<sup>211</sup> and those proposed by Cordeiro<sup>98</sup> and by Macchiagodena et al.,<sup>105</sup> referred to here as MMPB. For water, we consider the three-site SPC/E<sup>212</sup> and four-site TIP4P<sup>213</sup> and TIP4P/2005<sup>214</sup> potentials. It should be noted that, without additional testing, the CHARMM force field is supposed to be used in combination with the SPC/E, while OPLS with the TIP4P water model. However, the present work represents a far more stringent test of the model combinations considered than the ones on which such “thumb rules” are based, hence all possible combinations of the above models are included in the following analyses.

*Table 3.1: Properties of the Systems Simulated.  $x_{FA}$  is the Mole Fraction of the Formamide Molecules in the Simulation Box with an edge length  $L$ , While  $N_{FA}$  and  $N_{wat}$  Stand for the Number of Formamide and Water Molecules, Respectively.*

| $x_{FA}$ | $N_{FA}$ | $N_{wat}$ | $\rho / \text{g cm}^{-3}$ | $L / \text{\AA}$ |
|----------|----------|-----------|---------------------------|------------------|
| 0        | 0        | 512       | 0.997                     | 24.85            |
| 0.1      | 51       | 461       | 1.028                     | 25.81            |
| 0.2      | 102      | 410       | 1.051                     | 26.69            |
| 0.3      | 154      | 358       | 1.068                     | 27.52            |
| 0.4      | 205      | 307       | 1.082                     | 28.32            |
| 0.5      | 256      | 256       | 1.093                     | 29.08            |
| 0.6      | 307      | 205       | 1.103                     | 29.80            |
| 0.7      | 358      | 154       | 1.111                     | 30.49            |
| 0.8      | 410      | 102       | 1.118                     | 31.15            |
| 0.9      | 461      | 51        | 1.124                     | 31.78            |
| 1        | 512      | 0         | 1.130                     | 32.39            |

In the water models considered, the O atom is the only center of a Lennard-Jones interaction, and the H atoms bear a fractional charge of  $q_H$ . The negative fractional charge compensating them is located on the O atom in the SPC/E model, but it is displaced from

the O atom by  $r_{OM}$  to a non-atomic interaction site, M, along the bisector of the H-O-H bond angle,  $\alpha_{HOH}$ , in the TIP4P and TIP4P/2005 models. The parameters of the water models considered are collected in Table 3.2. The OPLS model of formamide consists of only five interaction sites; the CH group is treated in this model as a united atom. The OPLS/AA\_mod, CHARMM, and Cordeiro models have interaction sites on all six atoms, while in the MMPB model the fractional charge is displaced from the O atom to two non-atomic interaction sites, denoted as L, located at the distance of 0.31 Å from the O atom along its two lone pair directions. According to the molecular geometry of formamide, all interaction sites are arranged in a planar way in all of these models. The interaction parameters of the formamide models considered are collected in Table 3.3.

**Table 3.2: Geometry and Interaction Parameters of the Water Models Used.**

| model                     | $r_{OH}/\text{Å}$ | $r_{OM}/\text{Å}$ | $\alpha_{HOH}/\text{deg}$ | $\sigma/\text{Å}$ | $\epsilon/\text{kJ mol}^{-1}$ | $q_H/e$ |
|---------------------------|-------------------|-------------------|---------------------------|-------------------|-------------------------------|---------|
| SPC/E <sup>212</sup>      | 1.000             | 0                 | 109.47                    | 3.1660            | 0.650                         | 0.4238  |
| TIP4P <sup>213</sup>      | 0.9572            | 0.1500            | 104.52                    | 3.1540            | 0.649                         | 0.5200  |
| TIP4P/2005 <sup>214</sup> | 0.9572            | 0.1546            | 104.52                    | 3.1589            | 0.775                         | 0.5564  |

All models considered are pairwise additive; the interaction energy of two molecules,  $i$  and  $j$ , being within the center-center cut-off distance, is calculated as

$$u_{ij} = \sum_{\alpha=1}^{N_i} \sum_{\beta=1}^{N_j} \frac{1}{4\pi\epsilon_0} \frac{q_\alpha q_\beta}{r_{i\alpha,j\beta}} + 4\epsilon_{\alpha\beta} \left[ \left( \frac{\sigma_{\alpha\beta}}{r_{i\alpha,j\beta}} \right)^{12} - \left( \frac{\sigma_{\alpha\beta}}{r_{i\alpha,j\beta}} \right)^6 \right] \quad (3.1)$$

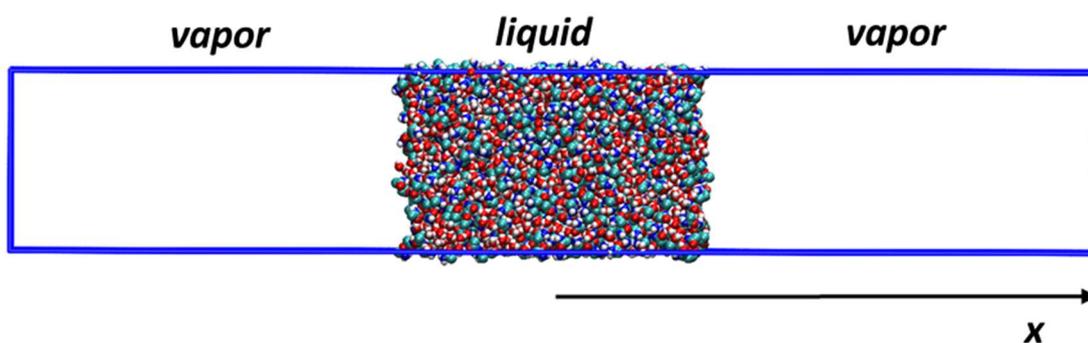
Here indices  $\alpha$  and  $\beta$  run over the  $N_i$  and  $N_j$  interaction sites of molecules  $i$  and  $j$ , respectively,  $\epsilon_0$  is the vacuum permittivity,  $r_{i\alpha,j\beta}$  is the distance of site  $\alpha$  on molecule  $i$  from site  $\beta$  on molecule  $j$ ,  $q_\alpha$  and  $q_\beta$  are the fractional charges corresponding to the respective interaction sites, while  $\epsilon_{\alpha\beta}$  and  $\sigma_{\alpha\beta}$  are the energy and distance parameters, respectively, of the Lennard-Jones interaction of sites  $\alpha$  and  $\beta$ , related to the corresponding parameters of the individual sites through the Lorentz-Berthelot rule,<sup>138</sup> i.e.,  $\epsilon_{\alpha\beta} = (\epsilon_\alpha \epsilon_\beta)^{1/2}$  and  $\sigma_{\alpha\beta} = (\sigma_\alpha + \sigma_\beta)/2$ .

*Table 3.3: Interaction Parameters of the Formamide Models Used. <sup>a</sup>Non-atomic interaction site; <sup>b</sup>H atom in cis relative position with the carboxylic oxygen; <sup>c</sup>H atom in trans relative position with the carboxylic oxygen*

| model                      | site           | $\sigma/\text{\AA}$ | $\epsilon/\text{kJ mol}^{-1}$ | $q/e$  |
|----------------------------|----------------|---------------------|-------------------------------|--|
| OPLS <sup>106</sup>        | CH             | 3.80                | 0.4815                        | 0.500  |
|                            | O              | 2.96                | 0.8792                        | -0.500   |
|                            | N              | 3.25                | 0.7118                        | -0.850   |
|                            | H              | -                   | -                             | 0.425  |
| OPLS/AA_mod <sup>116</sup> | C              | 3.80625             | 0.307524                      | 0.1398   |
|                            | O              | 3.00440             | 0.615048                      | -0.5283  |
|                            | H(C)           | 2.45630             | 0.043932                      | 0.1753   |
|                            | N              | 3.29875             | 0.497896                      | -0.4163  |
|                            | H              | -                   | -                             | 0.3121 <sup>b</sup><br>0.3174 <sup>c</sup>     |
|                            |                |                     |                               |  |
| CHARMM <sup>211</sup>      | C              | 3.564               | 0.293                         | 0.42   |
|                            | O              | 3.029               | 0.502                         | -0.51  |
|                            | H(C)           | 2.352               | 0.090                         | 0.08   |
|                            | N              | 3.296               | 0.837                         | -0.69  |
|                            | H              | 0.400               | 0.192                         | 0.35   |
|                            |                |                     |                               |  |
| Cordeiro <sup>98</sup>     | C              | 3.75                | 0.440                         | 0.340  |
|                            | O              | 2.96                | 0.880                         | -0.460   |
|                            | H(C)           | 2.75                | 0.159                         | 0.120  |
|                            | N              | 3.25                | 0.712                         | -0.830   |
|                            | H              | -                   | -                             | 0.415  |
| MMPB <sup>105</sup>        | C              | 3.75                | 0.43932                       | 0.154544                                       |
|                            | O              | 2.96                | 0.87864                       | 0  |
|                            | L <sup>a</sup> | -                   | -                             | -0.201887                                      |
|                            | H(C)           | 2.42                | 0.06276                       | 0.131902                                       |
|                            | N              | 3.25                | 0.71128                       | -0.566816                                      |
|                            | H              | -                   | -                             | 0.339036 <sup>b</sup><br>0.345109 <sup>c</sup> |

### 3.2 The Liquid–Vapor Interface of Formamide–Water Mixtures

Molecular dynamics simulations of the liquid-vapor interface of water-formamide mixtures of different compositions, including the two neat systems, have been performed on the canonical ( $N, V, T$ ) ensemble at the temperature of 300 K. The basic simulation box has contained 4000 molecules, among which 0, 400, 800, 1200, 1600, 2000, 2400, 2800, 3200, 3600, and 4000 have been formamide in the different systems. These systems are referred to here as the 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100% FA systems, respectively. The  $X$ ,  $Y$  and  $Z$  edges of the basic box have been 300 Å, 50 Å, and 50 Å long, respectively,  $X$  being the axis perpendicular to the macroscopic liquid surface (see Figure 3.2).



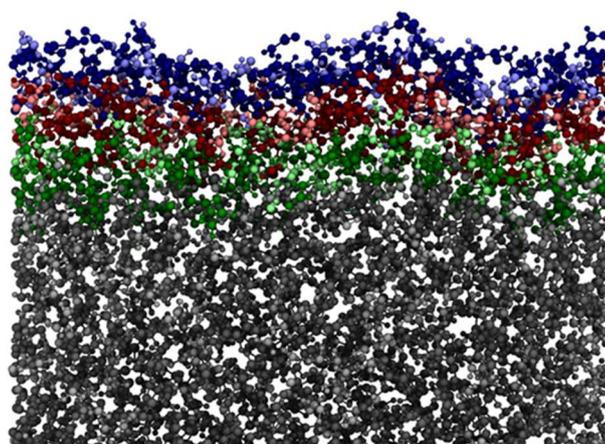
*Figure 3.2: Equilibrium snapshot of the 50% FA system. The simulation box contains the liquid phase between the two vapor phases. The  $X$  axis is shown at the bottom of the figure.*

Formamide molecules have been described by the CHARMM27 force field,<sup>211</sup> while water was modelled by the rigid SPC/E potential.<sup>212</sup> The choice of these potential models has been dictated by the fact that this model combination turned out to be the best in describing the mixing properties of water and formamide from the calculations described in the previous chapter and which results discussed in details in the chapter 4.1. All molecules have been treated as rigid bodies in the simulations; their geometries have been kept unchanged using the LINCS algorithm.<sup>215</sup> All interactions have been truncated to zero beyond the center-center cut-off distance of 15 Å; the long range part of the electrostatic interaction has been taken into account by means of the smooth variant of the Particle Mesh Ewald method<sup>168</sup> using a mesh spacing of 1.2 Å.

The simulations have been performed with the GROMACS 5.1.4 program package.<sup>216</sup> The temperature of the systems have been controlled with the Nosé-Hoover thermostat<sup>217,218</sup> with a time constant of 0.2 ps. The equations of motion have been integrated in time steps of 1 fs. At the beginning of each simulation, the required number

of water and formamide molecules have been placed in a basic simulation box the edges  $Y$  and  $Z$  of which have been  $50 \text{ \AA}$ , while the length of the  $X$  edge has roughly corresponded to the density of the liquid phase. After proper energy minimization, these bulk liquid systems have been equilibrated for 2 ns. Then the  $X$  edge of the basic box has been increased to  $300 \text{ \AA}$ , resulting in a system containing the liquid-vapor interface. The interfacial systems have been further equilibrated for 4 ns. Finally, during the 4 ns long production phase of the simulations, 4000 equilibrium sample configurations, separated by 1 ps long trajectories each, have been dumped for the subsequent analyses. All results have been averaged over the two liquid-vapor interfaces present in the basic box.

The intrinsic surface of the liquid phase has been determined on the sample configurations by the ITIM method.<sup>191</sup> In the ITIM analysis, a probe sphere of the radius of  $1.25 \text{ \AA}$  has been moved along test lines parallel with the  $X$  edge of the basic box, arranged in a  $125 \times 125$  grid. The entire analysis has been repeated three times per system; hence, the molecules constituting the first three molecular layers at the liquid surface have been separately identified. An equilibrium snapshot of the surface portion of the 50% FA system is shown in Figure 3.3 as taken out from the simulation, in which the molecules belonging to the first three layers at the liquid surface are separately indicated.



*Figure 3.3: Equilibrium snapshot of the 50% FA system, as taken out from the simulation. Molecules belonging to the first, second, and third molecular layer beneath the liquid surface are marked by blue, red, and green colors, respectively; molecules located beneath the third subsurface molecular layer are marked by grey color. Lighter and darker shades of the colors correspond to the water and formamide molecules, respectively. The snapshot only shows the  $X > 0 \text{ \AA}$  half of the basic box.*

Further, the surface tensions of the systems has been calculated through the difference of the lateral and normal pressure components.<sup>229</sup> The results for the simulated systems are collected in Table 3.4. As is seen, both models underestimate the surface tensions of the corresponding neat liquids (i.e.,  $72.8 \text{ mN/m}$  for water and  $58.2 \text{ mN/m}$  for

FA) by about 15%, which is also reflected in the somewhat too low surface tension values of the liquid. However, apart from this slight downshift, the composition dependence of the surface tension follows the experimental trend.

*Table 3.4: The calculated surface tension of the simulated systems.*

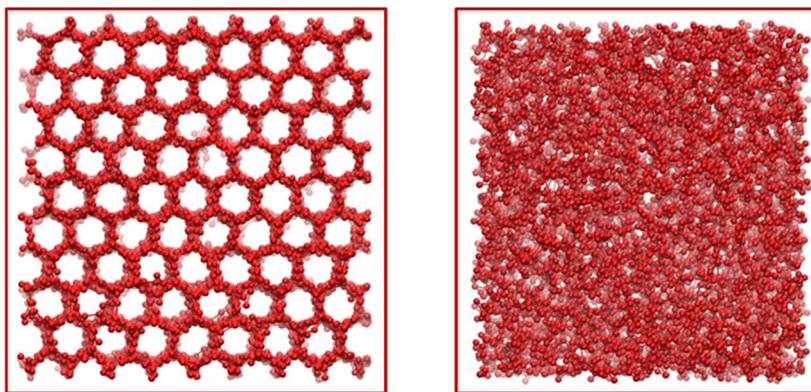
| System (FA)                 | 0%   | 10%  | 20%  | 30%  | 40%  | 50%  | 60%  | 70%  | 80%  | 90%  | 100% |
|-----------------------------|------|------|------|------|------|------|------|------|------|------|------|
| $\gamma / \text{mN m}^{-1}$ | 60.2 | 55.2 | 53.4 | 53.1 | 53.0 | 52.7 | 52.5 | 51.5 | 50.8 | 51.4 | 49.0 |

### 3.3 Adsorption of Formamide at the Surface of Amorphous and Crystalline Ices

The adsorption of formamide at the surface of both  $I_h$  and LDA ice has been simulated on the grand canonical  $(\mu, V, T)$  ensemble by the GCMC<sup>138,146</sup> method. Simulations involving amorphous ice have been performed at the temperatures of 50, 100, and 200 K, while adsorption on  $I_h$  ice has only been simulated at 200 K. The number of the water molecules in the basic box has been fixed to 2880 in every simulation based on earlier simulations take place in similar environment.<sup>136,175,219–224</sup> It should be emphasized that due to the periodic boundary conditions, the basic box has contained two interfaces. To obtain the adsorption isotherms, with each adsorbent and every temperature a set of GCMC simulations have been performed in which the chemical potential,  $\mu$ , of formamide has been systematically varied from values corresponding to practically no formamide molecule to those at which formamide fills the available space in the basic box. The number of the adsorbed formamide molecules has then been determined as a function of the chemical potential. The formamide chemical potential values considered in the different systems are collected in Tables 3.5-3.8. The X, Y, and Z edges of the rectangular basic simulation box have been 100 Å, 35.926 Å, and 38.891 Å, respectively, axis X being perpendicular to the macroscopic plane of the ice surface.

Water and formamide molecules have been described by the three-site SPC/E potential<sup>212</sup> and by the CHARMM27 force field,<sup>211</sup> respectively, since this model combination was previously found to best reproduce the mixing properties of water and formamide based on the simulation was described in chapter 3.1. All molecules have been treated as rigid bodies, and all interactions have been truncated to zero beyond the center–center cutoff distance of 12.5 Å.<sup>211</sup>

The simulations have been performed using the program MMC.<sup>209</sup> In every Monte Carlo step either a randomly chosen molecule has been randomly translated by no more than 0.25 Å and randomly rotated around a randomly chosen space-fixed axis by no more than 15°, or a formamide molecule has been tried to be either inserted to or deleted from the system. Particle displacement and insertion/deletion attempts have been performed with equal probabilities. Insertion and deletion attempts have been done according to the cavity biased scheme of Mezei;<sup>184,185</sup> thus, insertions have only been attempted into pre-existing empty cavities of the radius of at least 2.5 Å. Such cavities have been searched for in the basic box along a 100 × 100 × 100 grid, regenerated after every 10<sup>6</sup> Monte Carlo steps. Insertion and deletion attempts have been accepted or rejected according to the cavity biased scheme.<sup>184,185</sup> The probability of finding a suitable cavity,  $P_{\text{cav}}$ , occurring in these acceptance criteria has simply been estimated as the ratio of the number of suitable cavities found and grid points tested in configurations consisting of exactly N formamide molecules.<sup>184,185</sup> Acceptance of the particle displacement trials has been decided according to the standard Metropolis criterion.<sup>138,182</sup> The fraction of the successful particle displacement and insertion/deletion trials has turned out to be about 5% and 0.0001%, respectively, at the lowest temperatures. In preparation of the  $I_h$  phase of ice, the water molecules have been arranged in 18 layers, consisting of 160 molecules each, along the interface normal axis, X, in such a way that the two surfaces present in the basic box correspond to the (0001) surface of  $I_h$  ice. The LDA phase has been created by thermalizing the  $I_h$  phase at 300 K in a 10<sup>8</sup> Monte Carlo steps long run, followed by a 2 × 10<sup>8</sup> Monte Carlo steps long run at 350 K and another 10<sup>8</sup> Monte Carlo steps long run again at 300 K. Finally, the system has been quenched to 200 K and equilibrated for 10<sup>8</sup> Monte Carlo steps at this temperature. The top view of the prepared ice surface is shown in Figure 3.4.



*Figure 3.4: The structure of the crystalline ( $I_h$ , left) and low density amorphous (LDA, right) ice*

At the beginning of each simulation, two formamide molecules have been placed in the basic box far from the ice phase. The systems have been equilibrated for  $1.5 \times 10^9$  to  $2 \times 10^{10}$  Monte Carlo steps until both the number of the formamide molecules present in the basic box and the total energy of the system stopped showing even traces of a tendentious change in a trajectory of at least  $10^9$  Monte Carlo steps. The need for such an unusually long equilibration period came from the unusually small acceptance rate of the trial moves, in particular, that of the insertion/deletion attempts, at low temperatures. The number of the formamide molecules present in basic box has then been averaged over a subsequent  $10^9$  Monte Carlo steps long equilibrium run. Finally, at selected chemical potential values, indicated also in Tables 3.5-3.8, 5000 equilibrium sample configurations per system, separated from each other by  $5 \times 10^5$  Monte Carlo steps, have been dumped for detailed analyses from  $2.5 \times 10^9$  Monte Carlo steps long production runs. Equilibrium snapshots of the 50 K system, simulated at these four selected chemical potential values, covering the adsorption range from only a few adsorbed formamide molecules to multilayer adsorption, are shown in Figure 3.5.

*Table 3.5. Data of the Adsorption Isotherm of Formamide on LDA Ice at 50 K, as Obtained from the Simulations. The  $\mu$  Value Corresponding to the Point of Condensation Is Also Indicated (Bold Characters).*

| $\mu/\text{kJ mol}^{-1}$ | $\langle N \rangle$   | $p/p_0$                | $\Gamma/\mu\text{mol m}^{-2}$ |
|--------------------------|-----------------------|------------------------|-------------------------------|
| -63.54 <sup>a</sup>      | 3.020                 | $3.66 \times 10^{-55}$ | 0.178                         |
| -62.29                   | 9.017                 | $7.35 \times 10^{-54}$ | 0.535                         |
| -61.46                   | 11.03                 | $5.43 \times 10^{-53}$ | 0.654                         |
| -60.21                   | 13.20                 | $1.09 \times 10^{-51}$ | 0.773                         |
| -58.13                   | 32.52                 | $1.62 \times 10^{-49}$ | 1.90                          |
| -56.06                   | 37.34                 | $2.40 \times 10^{-47}$ | 2.20                          |
| -53.98 <sup>b</sup>      | 60.28                 | $3.57 \times 10^{-45}$ | 3.58                          |
| -51.90                   | 68.07                 | $5.29 \times 10^{-43}$ | 4.04                          |
| -49.82                   | 81.12                 | $7.85 \times 10^{-41}$ | 4.81                          |
| -48.99 <sup>c</sup>      | 115.4                 | $5.80 \times 10^{-40}$ | 6.83                          |
| -48.57                   | 515.2                 | $1.58 \times 10^{-39}$ | 30.6                          |
| -48.16                   | 896.6                 | $4.29 \times 10^{-39}$ | 53.3                          |
| -47.74 <sup>d</sup>      | 907.8                 | $1.17 \times 10^{-38}$ | 54.0                          |
| -46.91                   | 910.2                 | $8.61 \times 10^{-38}$ | 54.1                          |
| -46.08                   | 912.1                 | $6.36 \times 10^{-37}$ | 54.2                          |
| -45.25                   | 906.0                 | $4.70 \times 10^{-36}$ | 53.8                          |
| <b>-11.43</b>            |                       | <b>1.00</b>            |                               |
| <sup>a</sup> state I     | <sup>b</sup> state II | <sup>c</sup> state III | <sup>d</sup> state IV         |

*Table 3.6. Data of the Adsorption Isotherm of Formamide on LDA Ice at 100 K, as Obtained from the Simulations. The  $\mu$  Value Corresponding to the Point of Condensation Is Also Indicated (Bold Characters).*

| $\mu/\text{kJ mol}^{-1}$ | $\langle N \rangle$   | $p/p_0$                | $\Gamma/\mu\text{mol m}^{-2}$ |
|--------------------------|-----------------------|------------------------|-------------------------------|
| -69.74 <sup>a</sup>      | 7.011                 | $3.60 \times 10^{-25}$ | 0.416                         |
| -67.25                   | 7.548                 | $7.22 \times 10^{-24}$ | 0.449                         |
| -65.59                   | 9.103                 | $5.34 \times 10^{-23}$ | 0.535                         |
| -63.92                   | 10.40                 | $3.94 \times 10^{-22}$ | 0.618                         |
| -63.09                   | 17.01                 | $1.07 \times 10^{-22}$ | 1.01                          |
| -61.43                   | 25.02                 | $7.92 \times 10^{-21}$ | 1.49                          |
| -60.18                   | 28.01                 | $3.55 \times 10^{-20}$ | 1.66                          |
| -58.93                   | 37.95                 | $1.59 \times 10^{-19}$ | 2.26                          |
| -57.27 <sup>b</sup>      | 57.86                 | $1.18 \times 10^{-18}$ | 3.44                          |
| -56.02                   | 75.13                 | $5.27 \times 10^{-18}$ | 4.46                          |
| -54.78                   | 86.83                 | $2.36 \times 10^{-17}$ | 5.16                          |
| -53.95 <sup>c</sup>      | 104.1                 | $6.42 \times 10^{-17}$ | 6.18                          |
| -53.11                   | 221.9                 | $1.74 \times 10^{-16}$ | 13.2                          |
| -51.45 <sup>d</sup>      | 894.0                 | $1.29 \times 10^{-15}$ | 53.1                          |
| -49.79                   | 895.1                 | $9.52 \times 10^{-15}$ | 53.2                          |
| -48.54                   | 896.5                 | $4.27 \times 10^{-14}$ | 53.3                          |
| -47.29                   | 898.2                 | $1.91 \times 10^{-13}$ | 53.4                          |
| <b>-22.95</b>            |                       | <b>1.00</b>            |                               |
| <sup>a</sup> state I     | <sup>b</sup> state II | <sup>c</sup> state III | <sup>d</sup> state IV         |

*Table 3.7. Data of the Adsorption Isotherm of Formamide on LDA Ice at 200 K, as Obtained from the Simulations. The  $\mu$  Value Corresponding to the Point of Condensation Is Also Indicated (Bold Characters).*

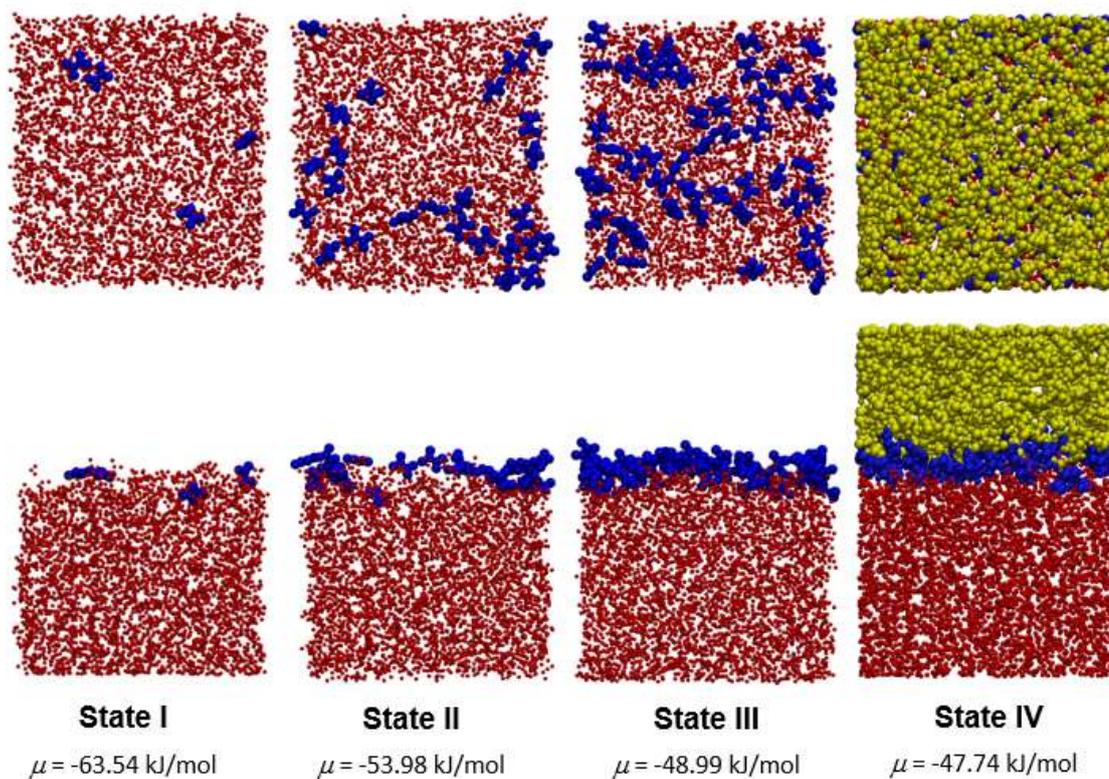
| $\mu/\text{kJ mol}^{-1}$ | $\langle N \rangle$   | $p/p_0$                | $\Gamma/\mu\text{mol m}^{-2}$ |
|--------------------------|-----------------------|------------------------|-------------------------------|
| -76.36                   | 0.0021                | $7.93 \times 10^{-10}$ | $1.25 \times 10^{-4}$         |
| -73.04                   | 0.0443                | $5.86 \times 10^{-9}$  | $2.63 \times 10^{-3}$         |
| -70.21                   | 1.028                 | $3.21 \times 10^{-8}$  | $6.11 \times 10^{-2}$         |
| -68.05                   | 2.086                 | $1.18 \times 10^{-7}$  | 0.124                         |
| -64.73 <sup>a</sup>      | 9.434                 | $8.70 \times 10^{-7}$  | 0.561                         |
| -62.73                   | 24.10                 | $2.89 \times 10^{-6}$  | 1.43                          |
| -61.90                   | 41.68                 | $4.76 \times 10^{-6}$  | 2.48                          |
| -61.40 <sup>b</sup>      | 50.22                 | $6.43 \times 10^{-6}$  | 2.98                          |
| -60.57                   | 57.58                 | $1.06 \times 10^{-5}$  | 3.42                          |
| -59.74                   | 93.25                 | $1.75 \times 10^{-5}$  | 5.54                          |
| -58.91 <sup>c</sup>      | 118.29                | $2.88 \times 10^{-5}$  | 7.03                          |
| -58.07                   | 338.19                | $4.75 \times 10^{-5}$  | 20.1                          |
| -57.24                   | 872.0                 | $7.83 \times 10^{-5}$  | 51.8                          |
| -56.41 <sup>d</sup>      | 878.0                 | $1.29 \times 10^{-4}$  | 52.2                          |
| -54.75                   | 896.1                 | $3.51 \times 10^{-4}$  | 53.2                          |
| -53.09                   | 901.6                 | $9.54 \times 10^{-4}$  | 53.6                          |
| <b>-41.52</b>            |                       | <b>1.000</b>           |                               |
| <sup>a</sup> state I     | <sup>b</sup> state II | <sup>c</sup> state III | <sup>d</sup> state IV         |

**Table 3.8. Data of the Adsorption Isotherm of Formamide on  $I_h$  Ice at 200 K, as Obtained from the Simulations. The  $\mu$  Value Corresponding to the Point of Condensation Is Also Indicated (Bold Characters).**

| $\mu/\text{kJ mol}^{-1}$ | $\langle N \rangle$   | $p/p_0$                | $\Gamma/\mu\text{mol m}^{-2}$ |
|--------------------------|-----------------------|------------------------|-------------------------------|
| -78.03                   | 0.0002                | $2.92 \times 10^{-10}$ | $1.19 \times 10^{-5}$         |
| -76.36                   | 0.0008                | $7.93 \times 10^{-10}$ | $4.75 \times 10^{-5}$         |
| -73.04                   | 0.0075                | $5.86 \times 10^{-9}$  | $4.46 \times 10^{-4}$         |
| -69.71                   | 0.0679                | $4.33 \times 10^{-8}$  | $4.03 \times 10^{-3}$         |
| -67.22                   | 1.322                 | $1.94 \times 10^{-7}$  | $7.86 \times 10^{-2}$         |
| -66.39                   | 2.307                 | $3.20 \times 10^{-7}$  | 0.137                         |
| -64.73 <sup>a</sup>      | 5.635                 | $8.70 \times 10^{-7}$  | 0.335                         |
| -63.06                   | 11.09                 | $2.36 \times 10^{-6}$  | 0.659                         |
| -61.40 <sup>b</sup>      | 42.57                 | $6.43 \times 10^{-6}$  | 2.53                          |
| -60.57                   | 54.67                 | $1.06 \times 10^{-5}$  | 3.25                          |
| -59.74                   | 112.3                 | $1.75 \times 10^{-5}$  | 6.67                          |
| -58.91 <sup>c</sup>      | 119.8                 | $2.88 \times 10^{-5}$  | 7.12                          |
| -58.07                   | 127.6                 | $4.75 \times 10^{-5}$  | 7.58                          |
| -57.41                   | 136.6                 | $7.08 \times 10^{-5}$  | 8.12                          |
| -56.74                   | 771.0                 | $1.06 \times 10^{-4}$  | 45.8                          |
| -56.41 <sup>d</sup>      | 748.1                 | $1.29 \times 10^{-4}$  | 44.5                          |
| -55.91                   | 761.4                 | $1.74 \times 10^{-4}$  | 45.2                          |
| -55.25                   | 768.0                 | $2.60 \times 10^{-4}$  | 45.6                          |
| -54.75                   | 770.9                 | $3.51 \times 10^{-4}$  | 45.8                          |
| <b>-41.52</b>            |                       | <b>1.000</b>           |                               |
| <sup>a</sup> state I     | <sup>b</sup> state II | <sup>c</sup> state III | <sup>d</sup> state IV         |

Since the surface of the amorphous ice phase, unlike that of crystalline ice, is corrugated, on the molecular length scale, by capillary waves, the adsorbed formamide molecules that belong to the first molecular layer (i.e., that are in direct contact with the amorphous ice phase) cannot be identified as simply as in the case of crystalline ice. Thus, while in the latter case the first adsorbed molecular layer is conventionally defined through the first minimum of the adsorbate density profile,<sup>175,221–223,225</sup> at the surface of amorphous ice these molecules have to be identified by means of an appropriate intrinsic surface analyzing method. Therefore, here we have identified the first layer formamide molecules adsorbed at the surface of amorphous ice using the ITIM method. In accordance with earlier findings concerning the optimal performance of the ITIM method for systems of small molecules,<sup>191,199</sup> here we used a probe sphere of the radius of 1.5 Å and a grid spacing of 0.4 Å along both the Y and Z axes. In deciding whether a formamide molecule is touched by the probe, its atoms have been represented by spheres with the diameter equal to the corresponding Lennard-Jones distance parameter,  $\sigma$ . Since practically no formamide molecules have been found, in any case, either to be dissolved

in the bulk ice phase, or to be in the vapor phase isolated from the adsorption layer, no modification of the original ITIM algorithm had to be made. The first layer formamide molecules identified this way are also indicated in Figure 3.5.



*Figure 3.5: Equilibrium snapshots of the system at 50K, both from top (top line) and side (bottom line) views, as obtained at four different chemical potential values from our simulations. Water molecules are shown by red, first layer formamide molecules (as determined by the ITIM method) are shown by blue, while outer formamide molecules are shown by yellow color.*

## 4. Results and Discussion

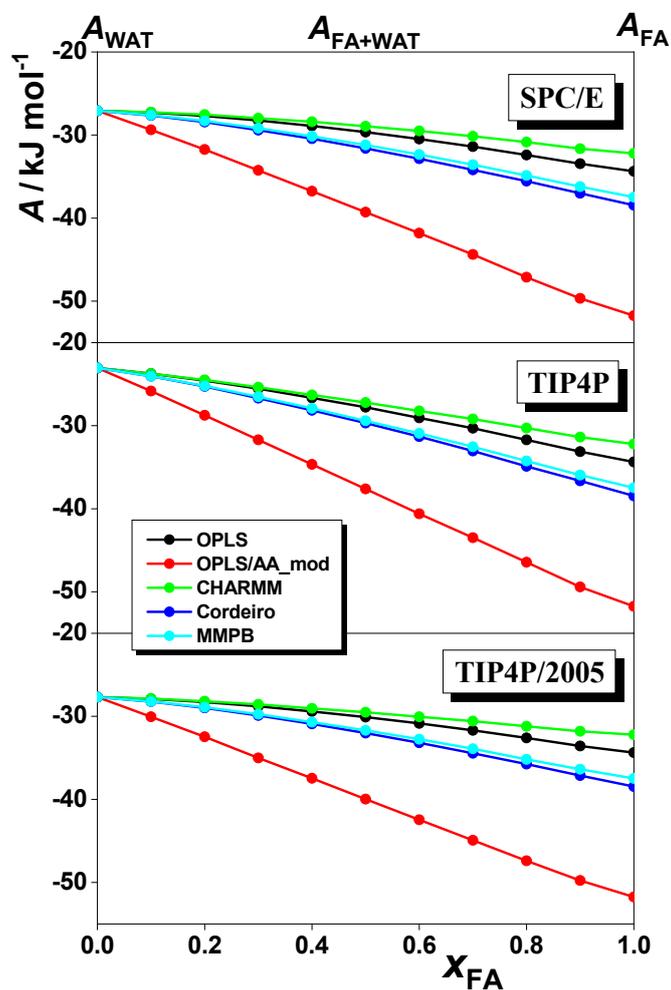
### 4.1. Miscibility and Thermodynamics of Mixing of Different Models of Formamide and Water

To obtain an appropriate model pair, what is able to reproduce the experimental properties of the formamide-water binary system, the chosen force fields which describe formamide (OPLS, OPLS/AA\_mod, CHARMM, Cordeiro and MMPB) and water (SPC/E, TIP4P and TIP4P/2005) systems have to be compatible with each other. To test these model pairs, their thermodynamic properties were calculated, along with the miscibility. The details of the Monte Carlo simulation used in thermodynamic integration were described in chapter 3.1. Here, the analysis of the results and the choice of the most reliable model pair is presented, which was subsequently used to investigate the phenomenon of the formamide enrichment at the vapor-liquid interface and the adsorption of formamide molecules on various ice surfaces.

#### 4.1.1 Thermodynamic properties of mixing

The calculated Helmholtz free energy ( $A$ ), energy ( $U$ ), and entropy ( $S$ ) in formamide-water systems, including the two neat systems, with respect to the ideal gas state are shown in Figures 4.1, 4.2 and 4.3, respectively, as obtained from the thermodynamic integration for the various model combinations considered. These quantities show a nearly linear composition dependence in most cases, suggesting that the mixing of the two components is probably close to the ideal mixing, where  $A_{A+B} = x_A A_A + x_B A_B$ . Markedly different behavior was previously seen for mixtures of acetone both with water<sup>121,122</sup> and methanol.<sup>226</sup> The free energy of the liquid decreases with increasing formamide mole fraction (see Figure 4.1). This free energy decrease is clearly of energetic origin, as both the energy and the entropy of the systems decrease as the formamide concentration increases. It is also seen that these quantities can be rather different for different model combinations. This difference mainly originates from the different thermodynamic properties of the formamide models considered: the difference

of the Helmholtz free energy, energy, and entropy of neat formamide can be as large as 50%, 35%, and 20%, respectively, when different potentials are considered.



*Figure 4.1: Helmholtz free energy of the SPC/E (top panel), TIP4P (middle panel), and TIP4P/2005 (bottom panel) water, and OPLS (black), OPLS/AA\_mod (red), CHARMM (green), Cordeiro (dark blue), and MMPB (light blue) formamide models of various compositions with respect to the ideal gas state. The data obtained by thermodynamic integration are shown by full circles; the lines connecting the symbols only serve as guides to the eye. Error bars are typically below  $\pm 0.2$  kJ/mol.*

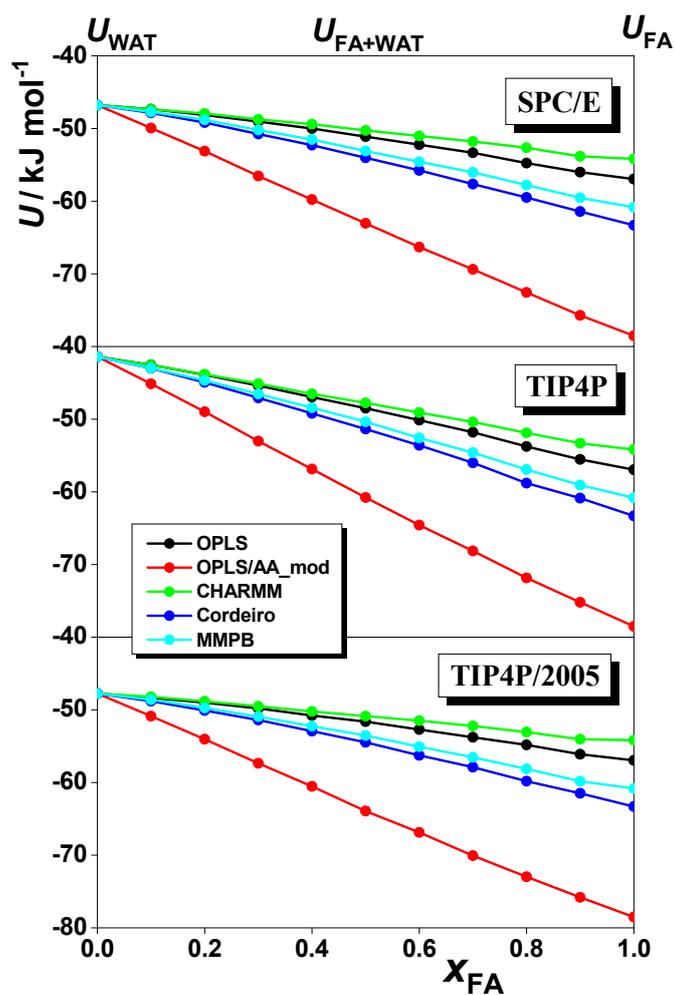


Figure 4.2: Energy of the SPC/E (top panel), TIP4P (middle panel), and TIP4P/2005 (bottom panel) water, and OPLS (black), OPLS/AA\_mod (red), CHARMM (green), Cordeiro (dark blue), and MMPB (light blue) formamide models of various compositions with respect to the ideal gas state. The data obtained by thermodynamic integration are shown by full circles; the lines connecting the symbols only serve as guides to the eye. Error bars are typically below  $\pm 0.2$  kJ/mol.

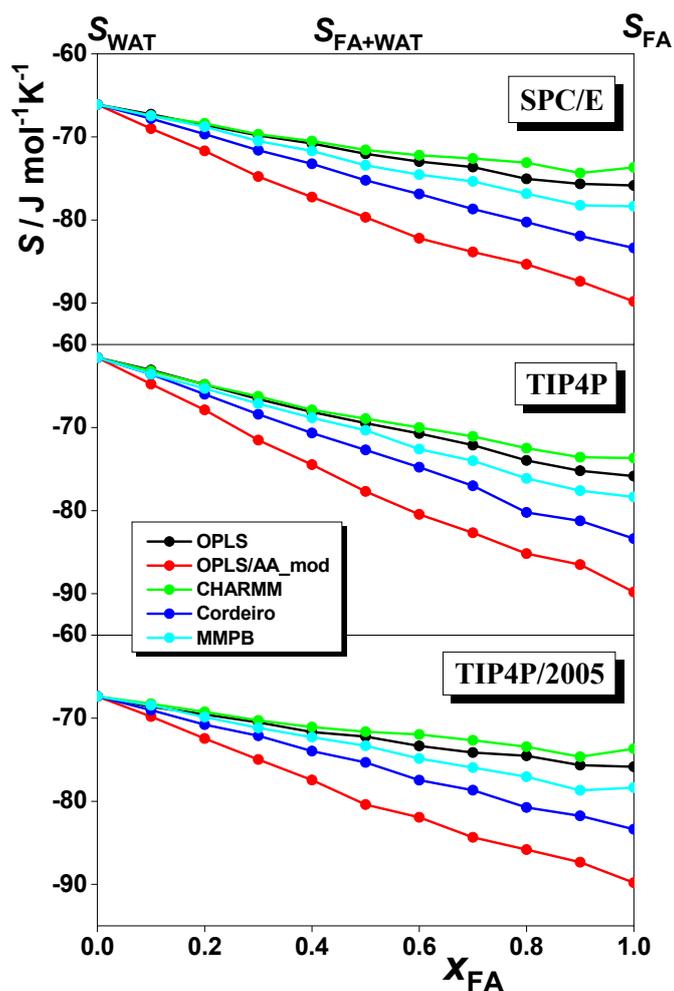


Figure 4.3: Entropy of mixtures of the SPC/E (top panel), TIP4P (middle panel), and TIP4P/2005 (bottom panel) water, and OPLS (black), OPLS/AA\_mod (red), CHARMM (green), Cordeiro (dark blue), and MMPB (light blue) formamide models of various compositions with respect to the ideal gas state. The data obtained by thermodynamic integration are shown by full circles; the lines connecting the symbols only serve as guides to the eye. Error bars are typically below  $\pm 1 \text{ J/mol K}$ .

The Helmholtz free energy, energy, and entropy change accompanying the mixing of neat liquid water and formamide are plotted in Figures 4.4, 4.5 and 4.6 respectively, as obtained with the different model combinations. As is clear from Figure 4.4, the free energy of mixing is negative in every case, indicating that the mixing is always thermodynamically preferred phenomena. The magnitude of  $A^{\text{mix}}$  is rather small, being always below 2 kJ/mol. For comparison, the value of  $RT$  is about 2.5 kJ/mol at 298 K. Similarly, small values are obtained for  $U^{\text{mix}}$  and  $S^{\text{mix}}$  in every case, their values being always within  $RT/2$  and  $R$ , respectively. (The value of  $RT/2$ , i.e., the average kinetic energy of the molecules along one single degree of freedom, is shown for comparison in

Figures 4.4, and 4.5, while that of  $R/2$  in Figure 4.6.) As a consequence, the obtained  $A^{\text{mix}}$ ,  $U^{\text{mix}}$ , and  $S^{\text{mix}}$  values are also rather close to each other as obtained with different model combinations, their difference never exceeds 1.3 kJ/mol, 1.8 kJ/mol, and 3.5 J/mol K, respectively. The free energy of mixing is the smallest in magnitude as obtained with the OPLS, Cordeiro, and MMPB models of formamide (the results obtained with these three models are always within error bars to each other), and deepest as obtained with the OPLS/AA\_mod model. No marked difference is seen between the water models in this respect. The OPLS/AA\_mod model is rather close to the model of ideal solution (see dotted line in Figure 4.4) followed by the CHARMM model.

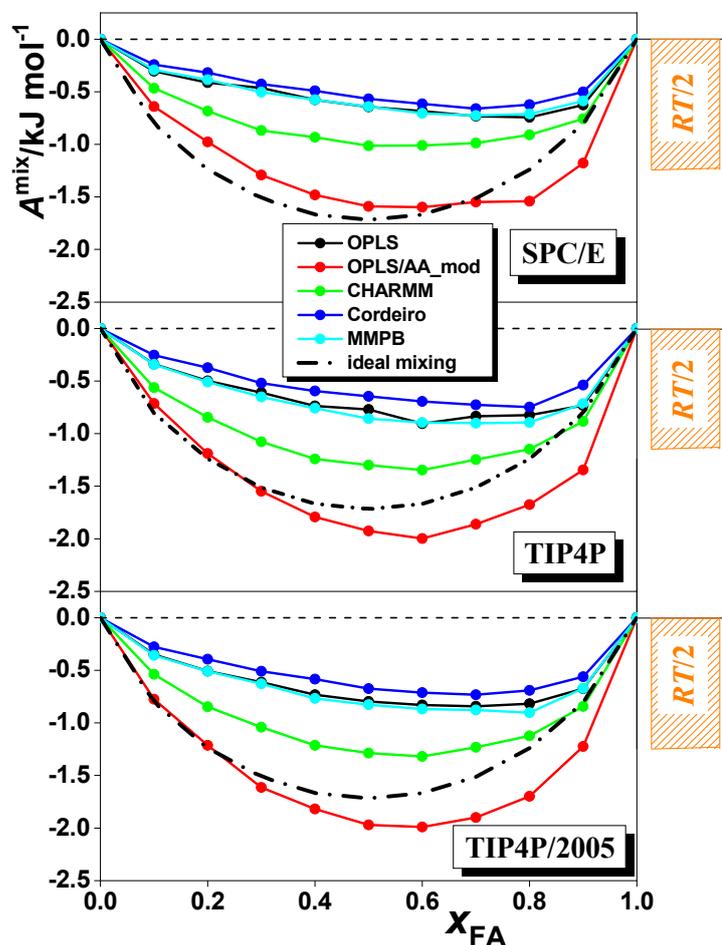


Figure 4.4: Helmholtz free energy of mixing, of the SPC/E (top panel), TIP4P (middle panel), and TIP4P/2005 (bottom panel) models of water with the OPLS (black), OPLS/AA\_mod (red), CHARMM (green), Cordeiro (dark blue), and MMPB (light blue) models of formamide. Error bars are typically below  $\pm 0.2$  kJ/mol. The calculated data are shown by full circles; the lines connecting the symbols only serve as guides to the eye. Zero value is indicated by dashed line the Helmholtz free energy of ideal mixing is shown by dotted line in panel. The average kinetic energy of the molecules along one degree of freedom of  $RT/2$  is also shown for reference by orange bars.

The energy of mixing shows a qualitatively different behavior when the OPLS/AA\_mod model of formamide is considered than what is seen with the other formamide models. Namely, it is always below zero in mixtures of the OPLS/AA\_mod model, but it is positive in all other cases. Although the deviation from the experimental values<sup>113,114</sup> is roughly of the same magnitude with all formamide models considered, the experimental data supports the qualitative behavior of the OPLS, CHARMM, Cordeiro, and MMPB models rather than that of OPLS/AA\_mod in this respect. Thus, according to the experimental data,<sup>113,114</sup> the observed negative free energy of mixing of water and formamide is of entropic origin, as the energy of mixing is very small and positive. Consistently, the entropy of mixing is found to be positive for all model combinations.

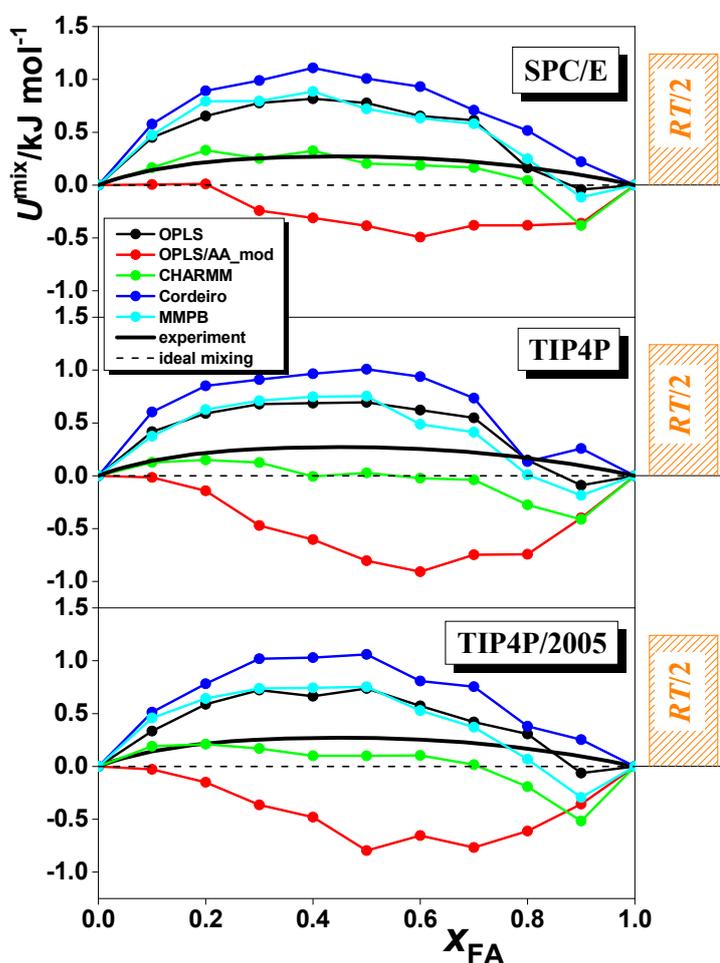


Figure 4.5: Energy of mixing of the SPC/E (top panel), TIP4P (middle panel), and TIP4P/2005 (bottom panel) models of water with the OPLS (black), OPLS/AA\_mod (red), CHARMM (green), Cordeiro (dark blue), and MMPB (light blue) models of formamide. Error bars are typically below  $\pm 0.2$  kJ/mol and  $\pm 1$  J/mol K. The calculated data are shown by full circles; the lines connecting the symbols only serve as guides to the eye. Zero value is indicated by dashed line (ideal mixing), experimental data of the energy of mixing<sup>113,114</sup> is shown by a thick line in panel. The average kinetic energy of the molecules along one degree of freedom of  $RT/2$  is also shown for reference by orange bars.

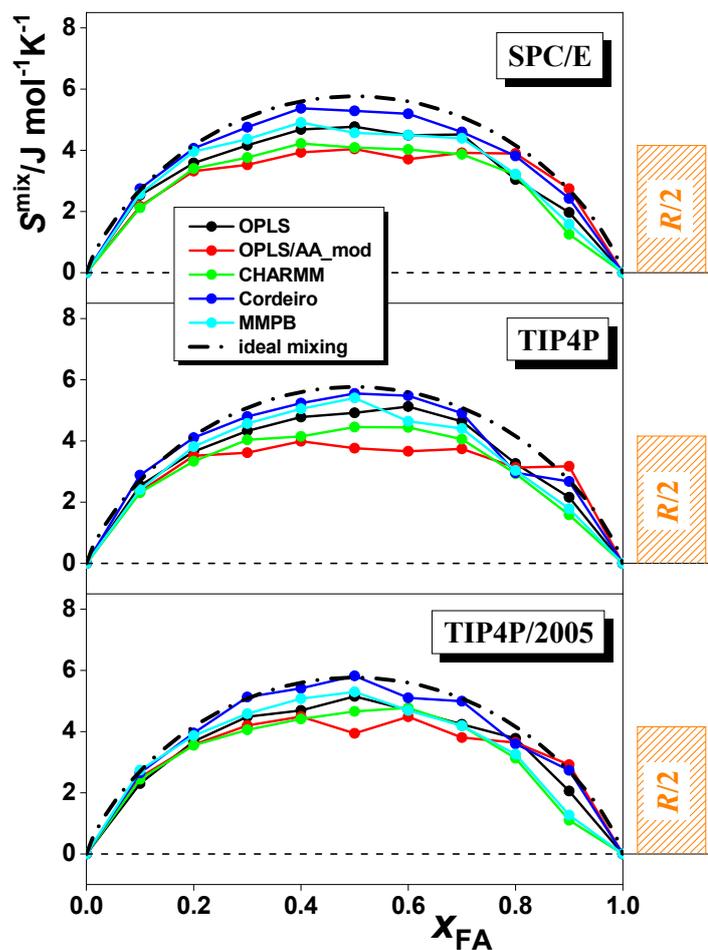


Figure 4.6: Entropy of mixing of the SPC/E (top panel), TIP4P (middle panel), and TIP4P/2005 (bottom panel) models of water with the OPLS (black), OPLS/AA\_mod (red), CHARMM (green), Cordeiro (dark blue), and MMPB (light blue) models of formamide. Error bars are typically below  $\pm 0.2$  kJ/mol and  $\pm 1$  J/mol K. The calculated data are shown by full circles; the lines connecting the symbols only serve as guides to the eye. Zero value is indicated by dashed line, the entropy of ideal mixing is shown by dotted line in panel. The average kinetic energy of the molecules along one degree of freedom of  $R/2$  is also shown for reference by orange bars.

Figure 4.6 also includes the entropy change corresponding to ideal mixing, i.e.,  $-R[x_{\text{FA}} \ln x_{\text{FA}} + (1-x_{\text{FA}}) \ln(1-x_{\text{FA}})]$ . As is clearly seen from this figure, the obtained  $S^{\text{mix}}$  values are rather close to this ideal term (shown as dotted line); their difference never exceeds 2.5 J/molK, and, considering the OPLS, Cordeiro, and MMPB models, this difference always remains below 1 J/molK, being within the error bars of the calculation. The largest entropy of mixing values in the cases of the three water models are obtained with the Cordeiro model, while the smallest ones with the OPLS/AA\_mod model of formamide, the  $S^{\text{mix}}(x_{\text{FA}})$  curves obtained with the former model being closer to the values corresponding to the entropy of ideal mixing. The obtained negative free energy of mixing

in the case of OPLS, CHARMM, Cordeiro and MMPB formamide models with water is simply the consequence of the ideal term in the entropy of mixing. It should again be stressed that both the energy and entropy of mixing values turned out to be very close to the term corresponding to ideal mixing (i.e., 0 for  $U^{\text{mix}}$  and  $-R[x_{\text{FA}} \ln x_{\text{FA}} + (1-x_{\text{FA}}) \ln(1-x_{\text{FA}})]$  for  $S^{\text{mix}}$ ) for every model combination and at every composition, confirming our previous suggestion, based on the nearly linear dependence the Helmholtz free energy, energy, and entropy of the liquid phase on  $x_{\text{FA}}$  that the mixing of these two compounds is nearly ideal.

As it has been emphasized previously, a negative value of  $A^{\text{mix}}$  is a necessary but not sufficient condition of miscibility; the two components are fully miscible with each other if and only if the value of the  $D$  parameter, defined in eq. 2.5, is positive in the entire composition range. To determine the  $D(x_{\text{FA}})$  data for the different model combinations, instead of numerical second derivative of the  $A^{\text{ex}}/RT$ , we fitted the Margules equation:<sup>227</sup>

$$A^{\text{ex}} / RT = x_{\text{FA}}(1-x_{\text{FA}}) \left[ A_{12}x_{\text{FA}} + A_{21}(1-x_{\text{FA}}) - (C_{12}x_{\text{FA}} + C_{21}(1-x_{\text{FA}}))x_{\text{FA}}(1-x_{\text{FA}}) + Bx_{\text{FA}}^2(1-x_{\text{FA}})^2 \right] \quad (4.1)$$

where  $A_{12}$ ,  $A_{21}$ ,  $C_{12}$ ,  $C_{21}$ , and  $B$  are the Margules parameters, to the calculated  $A^{\text{ex}}(x_{\text{FA}})$  data, and calculated the analytical second derivative of this fitted function. The quality of the Margules fit is illustrated in the inset of Figure 4.7 for the mixtures of the OPLS/AA\_mod model of formamide with the SPC/E and TIP4P/2005 water models.

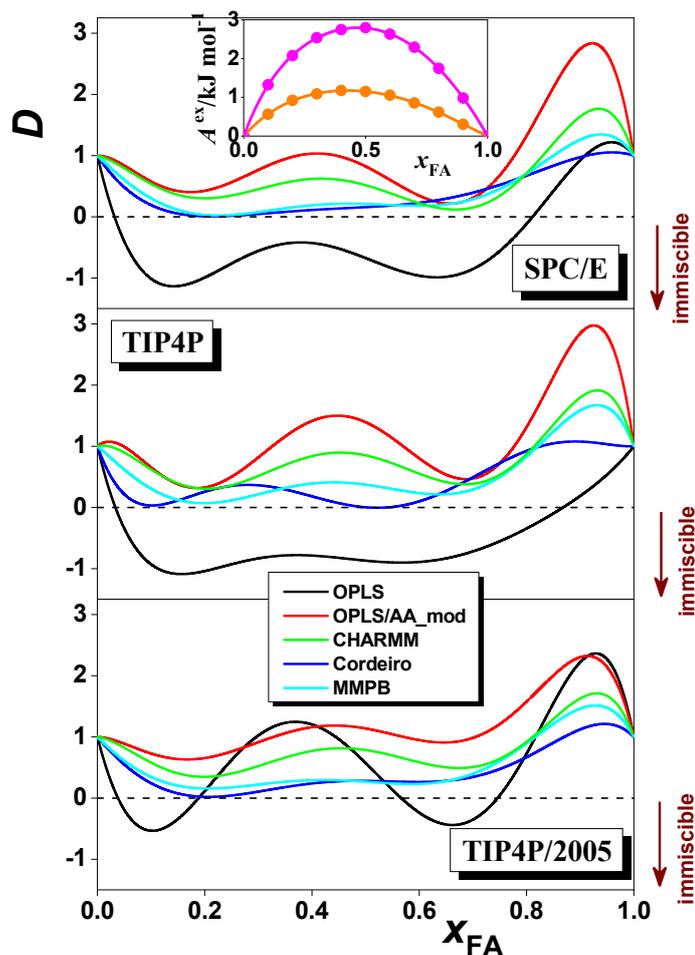


Figure 4.7: Composition dependence of the parameter  $D$ , defined by eq. 2.5, in mixtures of the SPC/E (top panel), TIP4P (middle panel), and TIP4P/2005 (bottom panel) water, and OPLS (black), OPLS/AA\_mod (red), CHARMM (green), Cordeiro (dark blue), and MMPB (light blue) formamide models. Estimated uncertainty of the data is below  $\pm 0.1$ . The inset shows the calculated  $A^{ex}$  data (full circles), together with their Margules fit (eq. 4.1, solid curves) for mixtures of the OPLS/AA\_mod model of formamide with the SPC/E (orange) and TIP4P/2005 (magenta) water models.

#### 4.1.2 The Miscibility of Different Formamide–Water Model Pairs

The  $D(x_{FA})$  curves obtained with the different model combinations considered are shown in Figure 4.7. As is seen from the figure, the curves obtained with the combinations of the OPLS model of formamide with the SPC/E and TIP4P water models drop below zero in a wide composition range, extending from the  $x_{FA}$  value of 0.02-0.03 to about 0.8-0.85, indicating that these model combinations are largely immiscible with each other. The OPLS and TIP4P/2005 models show two narrow immiscibility gaps in the  $x_{FA}$  ranges of about 0.04-0.2 and 0.55-0.75. Further, the  $D(x_{FA})$  curves obtained using the Cordeiro and MMPB formamide models in combination with any of the three water

models considered are very close to zero, the former dropping always below  $D = 0.05$ , while the latter below  $D = 0.15$  at certain compositions. Although the  $D(x_{\text{FA}})$  curve is positive in the entire composition range in the case of these six model combinations, its minimum is close to, or it is even below the estimated uncertainty of the  $D$  data of about 0.1, thus, immiscibility of some of these model combinations might occur in a narrow composition range. Further, full miscibility of these model combinations might also depend on the details of the simulation (e.g., long range correction, the use of rigid model, etc.).

On the other hand, the  $D$  values obtained using the OPLS/AA\_mod formamide model with any of the three water models considered remains safely above zero in the entire composition range, indicating that the OPLS/AA\_mod model is indeed fully miscible with all these three water models. On the other hand, as it has been pointed out previously, these model combinations are qualitatively incompatible with existing experimental data<sup>113,114</sup> in respect to the energy of mixing, as they correspond to negative rather than positive energy of mixing values in the entire composition range (Figure 4.5).

The  $D(x_{\text{FA}})$  curves obtained with the CHARMM model of formamide exhibit a clear minimum at  $x_{\text{FA}} = 0.067$  in the case of SPC/E and at  $x_{\text{FA}} = 0.2$  in the cases of TIP4P and TIP4P/2005 water models. The value of this minimum is around 0.12 when the SPC/E, 0.31 when the TIP4P and 0.35 when the TIP4P/2005 water model is used. Hence, the CHARMM model of formamide is fully miscible with all three tested water models. Further, as seen from Figure 4.5, these model combinations are also qualitatively compatible with the experimental data<sup>113,114</sup> concerning the sign of the energy of mixing. Considering also the fact that the experimental energy of mixing data is somewhat better reproduced when the SPC/E rather than the TIP4P or TIP4P/2005 models of water are used, the CHARMM model of formamide had been selected in combination with the SPC/E water model for the subsequent investigations.

Finally, it should be noted that whenever the value of  $D$  is found to drop below zero, it remains relatively small in magnitude, never going below about -1. This means that the thermodynamic driving force of the immiscibility of the corresponding model combinations is rather small. Therefore, in simulations of the bulk liquid phase, when demixing is also suppressed by the use of periodic boundary conditions, the occurrence of a visible phase separation might require much longer simulations than what is usually

done.<sup>228</sup> However, if the liquid mixture is in contact with an apolar object, such as in the presence of the liquid-vapor interface, phase separation can occur rather quickly.

## 4.2 Liquid–Vapor Interface of Formamide–Water Mixtures

In the previous chapter we have selected the model pair that best describes the mixing properties of formamide and water. In this chapter, we discuss in detail the results of molecular dynamics simulations of the liquid-vapor interface of water-formamide mixtures of different compositions, performed using this model combination.

### 4.2.1 Density Profiles

The (non-intrinsic) number density profiles of the water and formamide molecules as well as the mass density profiles of the entire systems and their surface layer along the macroscopic surface normal axis,  $X$ , are shown in Figure 4.8, as obtained in selected systems. As is seen, not only the mass density profiles change smoothly between the liquid and vapor phase, but the number density profiles of both molecules also show rather little structure at the interface: apart from a tiny peak occurring at the liquid side of the interface on the formamide density profiles in the systems of low formamide mole fraction, these profiles also change smoothly between the two phases. This result is in contrast with previous studies concerning the liquid surface of aqueous solutions of methanol,<sup>201</sup> acetonitrile,<sup>151</sup> dimethyl sulfoxide,<sup>202</sup> and HCN,<sup>145</sup> and suggests that, in contrast with the above solutes, formamide does not show considerable adsorption at the surface of its aqueous solutions.

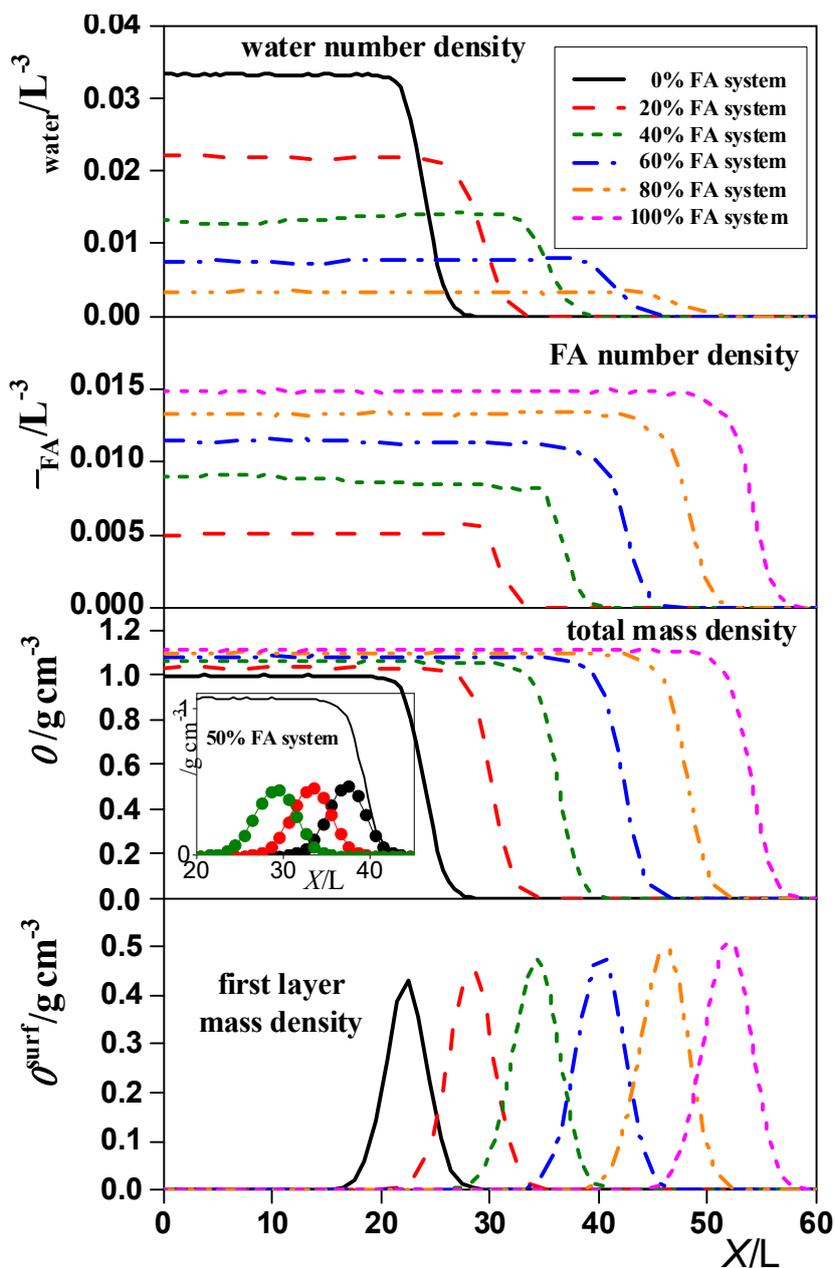
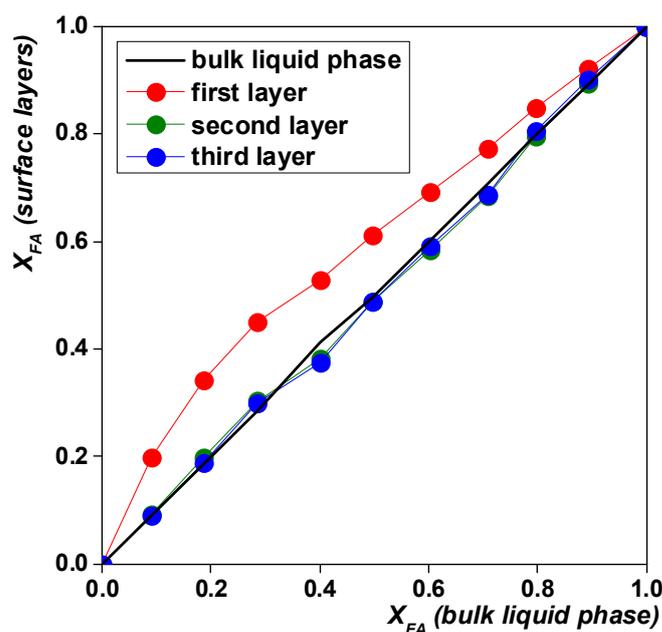


Figure 4.8: Number density profile of the water (top panel) and formamide (second panel) molecules as well as the mass density profile of the entire system (third panel) and of the surface layer of the liquid phase (bottom panel) along the interface normal axis,  $X$ , as obtained from the simulations of different systems. Black solid lines: 0% FA system, red dashed lines: 20% FA system, green dotted lines: 40% FA system, blue dash-dotted lines: 60% FA system, orange dash-dot-dotted lines: 80% FA system, magenta short dashed lines: 100% FA system. The inset shows the mass density profile of the first three layers (with black, red, and green colors, respectively) in the equimolar system. All profiles shown are averaged over the two interfaces present in the basic simulation box.

To further investigate this point, we show the composition of the first three molecular layers at the liquid surface (in terms of the formamide (FA) mole fraction) as a function of the composition of the bulk liquid phase in Figure 4.9. The FA mole fraction in the bulk liquid phase has simply been estimated from the average molecular number densities of water and formamide in the  $10 \text{ \AA}$  wide liquid slab in the middle of the basic

simulation box (i.e., at  $|X| < 5 \text{ \AA}$ ). As is seen from Figure 4.9, FA is only very weakly adsorbed at the liquid surface. Thus, in systems of low FA concentration, the mole fraction of FA in the surface layer is about twice as much as in the bulk liquid phase, while in systems of higher FA concentration the difference between the surface and bulk mole fraction values remains below 50%. This behavior is in a clear contrast with what was seen at the surface of the aqueous solution of, e.g., HCN, where an order of magnitude difference was observed between the surface and bulk mole fractions of the solute.<sup>145</sup> Furthermore, again in a clear contrast with the behavior of the aqueous solutions of HCN,<sup>145</sup> acetonitrile,<sup>151</sup> and acetone,<sup>152</sup> this adsorption is strictly restricted to the first molecular layer at the liquid surface; the composition of the second and third layers always agrees with that of the bulk liquid phase. The comparison of the composition of the first three subsurface molecular layers with those obtained earlier in other aqueous solutions<sup>145,151,152</sup> is shown in (Figure 4.10).



*Figure 4.9: Composition (in terms of FA mole fraction) of the first (red), second (green), and third (blue) molecular layer as the function of the bulk liquid phase composition. The lines connecting the symbols are just guides to the eye. For reference, the straight line corresponding to the bulk phase composition itself is also shown (black solid line).*

When comparing the mass density profiles of the entire systems with that of their surface layer (Figure 4.8), it is seen that the surface layer extends well into the  $X$  range of constant mass density in every case. Conversely, the density profile of the second and even the third layer (not shown) strongly overlap with the  $X$  range of intermediate densities between the two phases in every system. This finding stresses the importance of employing an intrinsic method in analyzing the surface of these systems, clearly demonstrating that the identification of the interfacial region through the non-intrinsic density profile results in the misidentification of a large number of molecules both as interfacial and non-interfacial ones.

As it is expected,<sup>230</sup> the density profiles of the individual molecular layers beneath the liquid surface can be very well fitted by Gaussian functions. The width of these Gaussians,  $\sigma$ , can serve as an estimate of the width of the corresponding layer. The  $\sigma$  values obtained in the first, second, and third molecular layer beneath the surface in the various systems (denoted by  $\sigma_{L1}$ ,  $\sigma_{L2}$ , and  $\sigma_{L3}$ , respectively) are summarized in Table 4.1. It is seen that upon approaching the interface the  $\sigma$  values become smaller, indicating that the corresponding molecular layers get progressively more compact. Thus, the second and third layers are 3-5% and 5-10% thicker than the first layer in every case. This finding is again in contrast with what was seen at the surface of other aqueous liquid mixtures,<sup>145,152,231</sup> and suggests the presence of a strong hydrogen-bonding network in the surface layer, which can partly compensate the energy loss of the surface molecules due to the vicinity of the vapor phase, and which makes the surface layer unusually compact. This point is further addressed in a subsequent section.

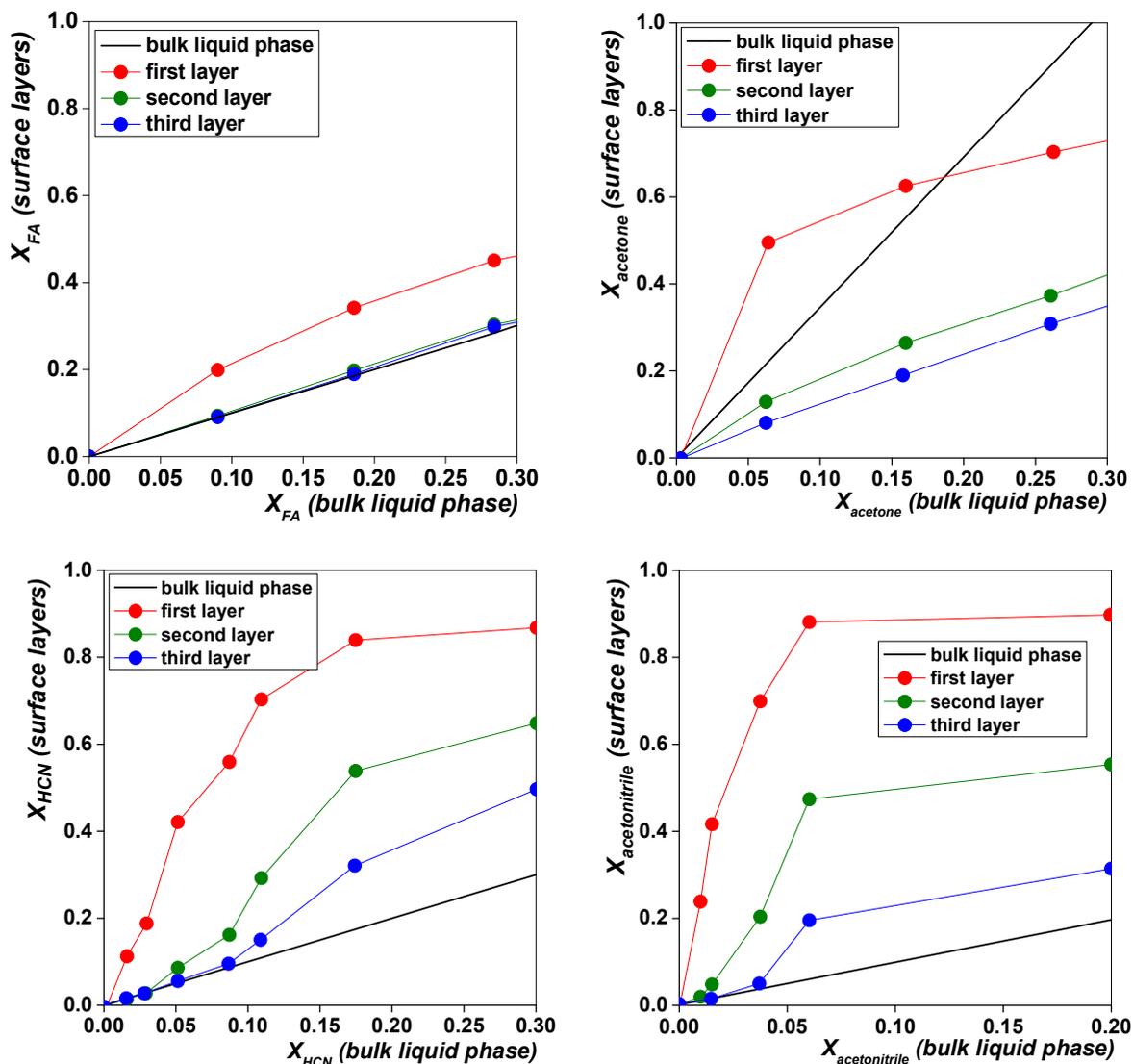


Figure 4.10: Comparison of composition (in terms of mole fraction) of the first (red), second (green), and third (blue) molecular layer as the function of the bulk liquid phase composition in the case of FA (top left), acetone (top right)<sup>152</sup> HCN (bottom left),<sup>145</sup> and acetonitrile (bottom right).<sup>151</sup> The lines connecting the symbols are just guides to the eye. For reference, the straight line corresponding to the bulk phase composition itself is also shown (black solid line).

*Table 4.1: Various Properties of the Systems Simulated<sup>a</sup>*

| system  | $\sigma_{L1}/\text{\AA}$ | $\sigma_{L2}/\text{\AA}$ | $\sigma_{L3}/\text{\AA}$ | $\tau_{\text{surf}}/\text{ps}$ |      | $\tau_{\text{D}}/\text{ps}$ |      | $D_{\parallel}/\text{\AA}^2\text{ps}^{-1}$ |       |
|---------|--------------------------|--------------------------|--------------------------|--------------------------------|------|-----------------------------|------|--|-------|
|         |                          |                          |                          | water                          | FA   | water                       | FA   | water                                      | FA    |
| 0% FA   | 3.79                     |                          |                          | 12.5                           | -    | 8.0                         | -    | 0.547                                      | -     |
| 10% FA  | 3.97                     | 4.08                     | 4.22                     | 11.3                           | 67.8 | 11.7                        | 13.3 | 0.448                                      | 0.393 |
| 20% FA  | 4.08                     | 4.27                     | 4.46                     | 11.1                           | 59.8 | 13.6                        | 15.8 | 0.422                                      | 0.366 |
| 30% FA  | 4.15                     | 4.37                     | 4.55                     | 11.1                           | 55.0 | 14.9                        | 17.9 | 0.412                                      | 0.345 |
| 40% FA  | 4.20                     | 4.42                     | 4.59                     | 11.3                           | 53.1 | 17.8                        | 19.1 | 0.360                                      | 0.336 |
| 50% FA  | 4.27                     | 4.50                     | 4.63                     | 11.3                           | 51.0 | 17.8                        | 20.3 | 0.376                                      | 0.331 |
| 60% FA  | 4.28                     | 4.50                     | 4.61                     | 11.0                           | 49.9 | 19.8                        | 21.6 | 0.353                                      | 0.323 |
| 70% FA  | 4.31                     | 4.52                     | 4.61                     | 11.3                           | 49.2 | 20.3                        | 22.3 | 0.356                                      | 0.324 |
| 80% FA  | 4.32                     | 4.54                     | 4.60                     | 11.6                           | 48.9 | 19.7                        | 24.0 | 0.380                                      | 0.312 |
| 90% FA  | 4.42                     | 4.61                     | 4.66                     | 11.3                           | 48.9 | 18.8                        | 24.4 | 0.408                                      | 0.316 |
| 100% FA | 4.34                     |                          |                          | -                              | 48.9 | -                           | 26.0 | -  | 0.307 |

<sup>a</sup>Error bars affect the last shown decimal digit.

#### 4.2.2 Lateral Distribution of the Surface Molecules

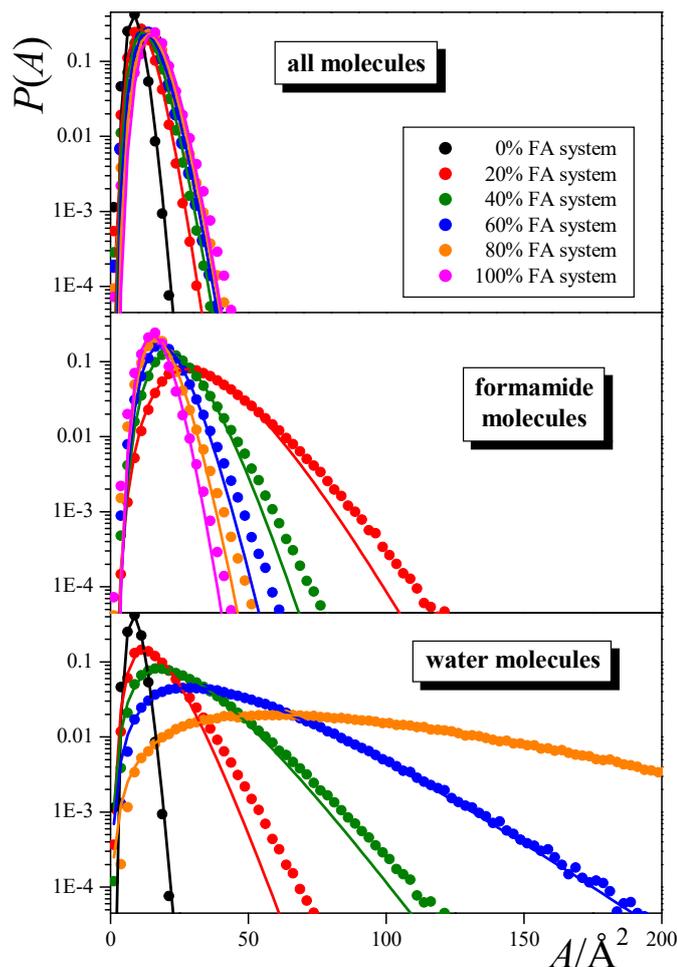
As it has been discussed above, the water and formamide molecules do not show considerable microscopic separation along the macroscopic surface normal axis,  $X$ , as the compositions of the subsequent molecular layers are rather similar to each other and to that of the bulk liquid phase. To further analyze the arrangement of the two kinds of molecules around each other, here we investigate the possibility of their lateral microscopic self-association within the surface layer. For this purpose, we have projected the centers-of-mass of the surface layer molecules to the macroscopic plane of the interface,  $YZ$ , and calculated the area distribution of the Voronoi cells of these projections. In a two-dimensional set of seeds (i.e., the projections of the molecular centers), the Voronoi cell of a given seed is the locus of the points in the plane that are closer to this seed than to any other one.<sup>203–205</sup> If the seeds are evenly distributed in the plane, the area of their Voronoi cells,  $A$ , follows a gamma distribution,<sup>232–234</sup> i.e.,

$$P(A) = a A^{\nu-1} \exp(-\nu\rho A). \quad (4.2)$$

In this equation,  $\nu$  and  $\rho$  are free parameters, while the factor  $a$  normalizes this probability distribution to unity. On the other hand, if the seeds show considerable self-association, leaving large empty areas between them, the Voronoi cells of the seeds located at the boundary of these empty domains are considerably larger than those being inside the self-associates. As a consequence, the corresponding  $P(A)$  distribution deviates considerably from eq. 4.2, exhibiting a long, exponentially decaying tail at large area values.<sup>206</sup>

To check the possible self-association of the like molecules in binary systems, the  $P(A)$  distribution should be calculated considering only one type of the molecules, and disregarding the other one. Thus, self-associates of the disregarded component are transformed to empty domains, and hence the  $P(A)$  distribution of the component considered exhibits the exponentially decaying tail. On the other hand, in the lack of such self-associates of the disregarded component, the  $P(A)$  distribution of the component considered still follows eq. 4.2.<sup>235</sup>

Figure 4.11 shows the  $P(A)$  distribution of the projections of the surface molecules to the  $YZ$  plane, considering both kinds of molecules as well as only formamides and only waters, as obtained in selected systems. As is seen, when one component is disregarded, the resulting distribution is broader than that obtained by considering both components, and gets progressively broader with increasing mole fraction of the disregarded component. More importantly, the  $P(A)$  distributions obtained by considering only one of the components can always be reasonably well fitted by eq. 4.2. To magnify the small deviations from the fit at large area values, the obtained data are shown on a logarithmic scale in Figure 4.11. Yet, slight deviations of the simulated data from eq. 4.2 is only seen at large  $A$  values in the systems of low FA concentration, indicating that no considerable self-association of the like molecules occurs even within the surface layer. In other words, water and FA molecules turn out to be miscible with each other on the molecular scale even at the surface of their mixtures. This behavior of FA is again in a clear contrast with that of HCN,<sup>145</sup> acetonitrile,<sup>151</sup> and acetone,<sup>152</sup> and is in accordance with the claim that the mixing of FA with water is nearly ideal.<sup>110,111,236</sup>



*Figure 4.11: Area distribution of the Voronoi polygons corresponding to the projections of the surface molecules to the macroscopic plane of the surface, YZ, when considering all surface molecules (top panel), only the surface formamide molecules (middle panel), and only the surface water molecules (bottom panel) in the different systems simulated (symbols). The color coding of the various systems is the same as in Fig.4.8. Solid lines show the best fit to the simulation data according to eq. 4.2.*

### 4.2.3 Hydrogen Bonding at the Liquid Surface

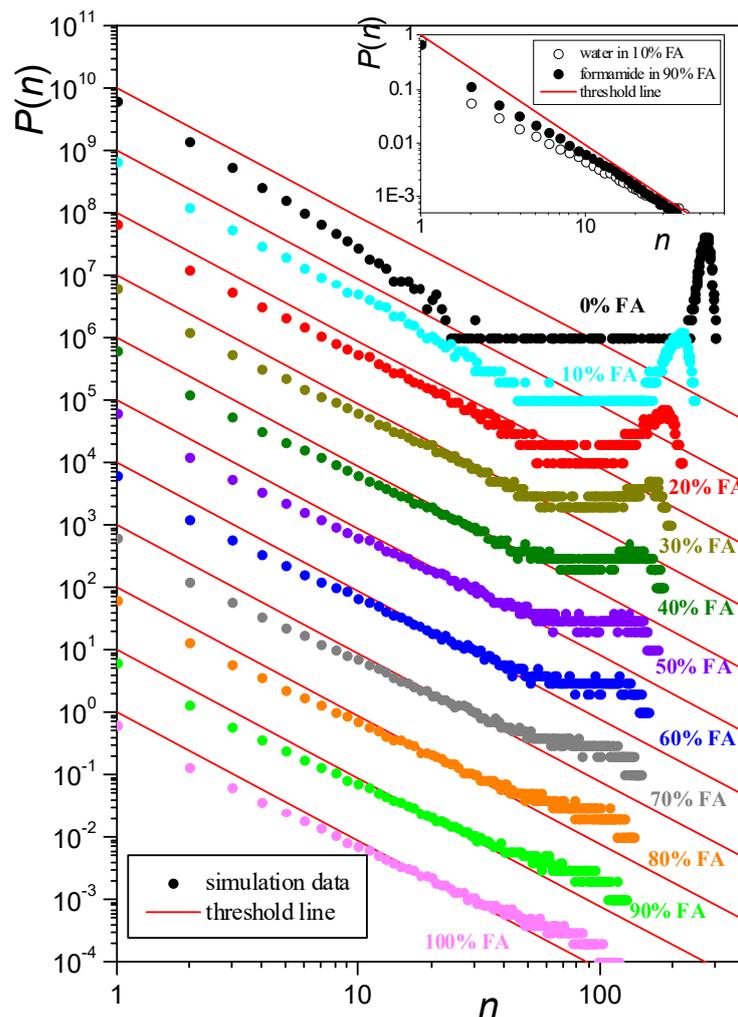
The observed molecular scale miscibility of the water and formamide molecules in the surface layer suggests that, similarly to the bulk liquid phase,<sup>100</sup> they can substitute each other in the hydrogen bonding network, and together form an infinite, percolating two-dimensional network at the liquid surface. The formation of such a percolating mixed network of water and formamide molecules has already been observed in three dimensions in the bulk liquid phase.<sup>100</sup> Further, a two-dimensional percolating hydrogen-bonding network of the molecules is also known to exist at the surface of water at room temperature.<sup>237,238</sup>

To address this point, we have calculated the size distribution of the hydrogen-bonding clusters in the surface layer of the systems simulated. The size of a cluster,  $n$ , is simply the number of molecules forming it (irrespective of whether they are water or FA). In order to consider only the two-dimensional, surface clusters, all these molecules have to be part of the surface layer. Two molecules belong to the same cluster if they are connected by a chain of hydrogen-bonded molecule pairs. Two molecules are considered here as being hydrogen-bonded to each other if (i) the distance of the hydrogen-donor (i.e., water O or formamide N) and hydrogen-acceptor (i.e., water O or formamide O) atoms is less than 3.5 Å, and (ii) the distance of the donated H atom and its acceptor is less than 2.4 Å. These cut-off distances correspond to the first minimum position of the corresponding partial pair correlation functions. (It should be noted that although formamide is able to form also weak, C-H $\cdots$ O type hydrogen-bonds, here we use the term ‘hydrogen-bond’ referring only to strong, O-H or N-H donated hydrogen-bonds.)

The size distribution of the surface clusters,  $P(n)$ , is shown in Figure 4.12 as obtained in selected systems. At the percolation threshold, the  $P(n)$  distribution of an infinitely large system follows a power law, i.e.,

$$P(n) \sim n^{-\alpha}. \quad (4.3)$$

with the universal exponent  $a = 2.05$  in two-dimensional systems.<sup>239</sup> The critical line corresponding to the  $P(n)$  distribution at the percolation threshold is also indicated in Figure 4.12. As is seen, the calculated  $P(n)$  distribution exceeds the critical line at large cluster sizes in every case, indicating that the mixed hydrogen-bonding clusters of the water and FA molecules within the surface layer indeed span the entire liquid surface. By contrast, the  $P(n)$  distribution of both the clusters built up solely by water molecules at the surface of the 10% FA system, and those built up solely by formamide molecules at the surface of the 90% FA system drop below the critical line at small  $n$  values, and remain below it in the entire  $n$  range (see the inset of Figure 4.12). This finding shows that, contrary to the mixed clusters, the surface clusters of only like molecules do not span the liquid surface even if this component is in a large excess in the surface layer.



*Figure 4.12: Size distribution of the hydrogen bonded clusters, formed by both molecules, within the surface layer of the systems simulated (full black circles), together with the critical line of eq. 4.2 corresponding to the percolation threshold (solid red lines), shown on a double logarithmic scale. The data corresponding to the 90% FA, 80% FA, 70% FA, 60% FA, 50% FA, 40% FA, 30% FA, 20% FA, 10% FA, and 0% FA systems are shifted upwards by 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 orders of magnitude, respectively, for better visibility. The inset compares the size distribution of the H-bonded clusters of solely the water molecules in the surface layer of the 10% FA system (open circles), and that of solely the FA molecules in the surface layer of the 90% FA system (filled circles) with the critical line of eq. 4.2 corresponding to the percolation threshold (solid red line).*

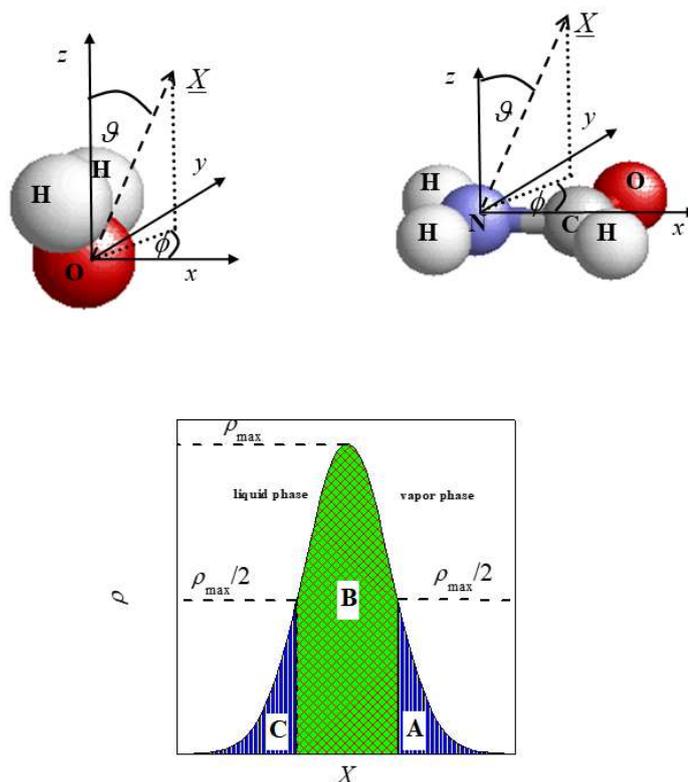
#### 4.2.4 Surface Orientation

The full description of a rigid molecule relative to an external direction (and, thus, also to an external plane) requires two independent orientational variables. Therefore, the full description of the orientational statistics of such molecules relative to an external direction or plane can only be done through the bivariate joint distribution of these two variables.<sup>176,240</sup> Clearly, the two angular polar coordinates of the external vector in a local

Cartesian frame fixed to the individual molecules represent a sufficient choice of such an independent orientational variable pair. However, considering also the fact that the polar angle  $\mathcal{G}$  is the angle between two general spatial vectors, but  $\phi$  is formed by two vectors that are restricted to lay in a given plane (i.e., the  $xy$  plane of the local frame) by definition, uncorrelated orientation of the molecules with the external plane (liquid surface) corresponds to a uniform bivariate distribution only if  $\cos\mathcal{G}$  and  $\phi$  are chosen to be the orientational variables.<sup>176,240</sup>

In order to analyze the orientational distribution relative to the macroscopic plane of the liquid surface,  $YZ$ , here we define the above local Cartesian frames in the following way. For water, its origin is located at the O atom, axis  $z$  points along the main symmetry axis of the molecule, oriented in such a way that the  $z$  coordinates of the H atoms are positive, axis  $y$  is parallel with the line connecting the two H atoms, while axis  $x$  is the molecular normal. For FA, the origin is at the N atom, axis  $z$  is the molecular normal, axis  $x$  points along the N-C bond from the N to the C atom, while axis  $y$  is perpendicular to both axes  $x$  and  $z$ , and is oriented in such a way that the  $y$  coordinate of the O atom is positive. The definition of these local frames as well as of the polar angles  $\mathcal{G}$  and  $\phi$ , describing the alignment of the macroscopic surface normal vector,  $\underline{X}$ , in these frames are illustrated in Figure 4.13. By our convention, the surface normal vector,  $\underline{X}$ , is oriented in such a way that it points from the liquid to the vapor phase. It should also be noted that the symmetry of the molecules restricts the range of  $\cos\mathcal{G}$  and  $\phi$  that corresponds to different orientations. Hence, the above frames have always been chosen in such a way that the relations  $\phi \leq 90^\circ$  (for water) and  $0 \leq \cos\mathcal{G}$  (for FA) hold.

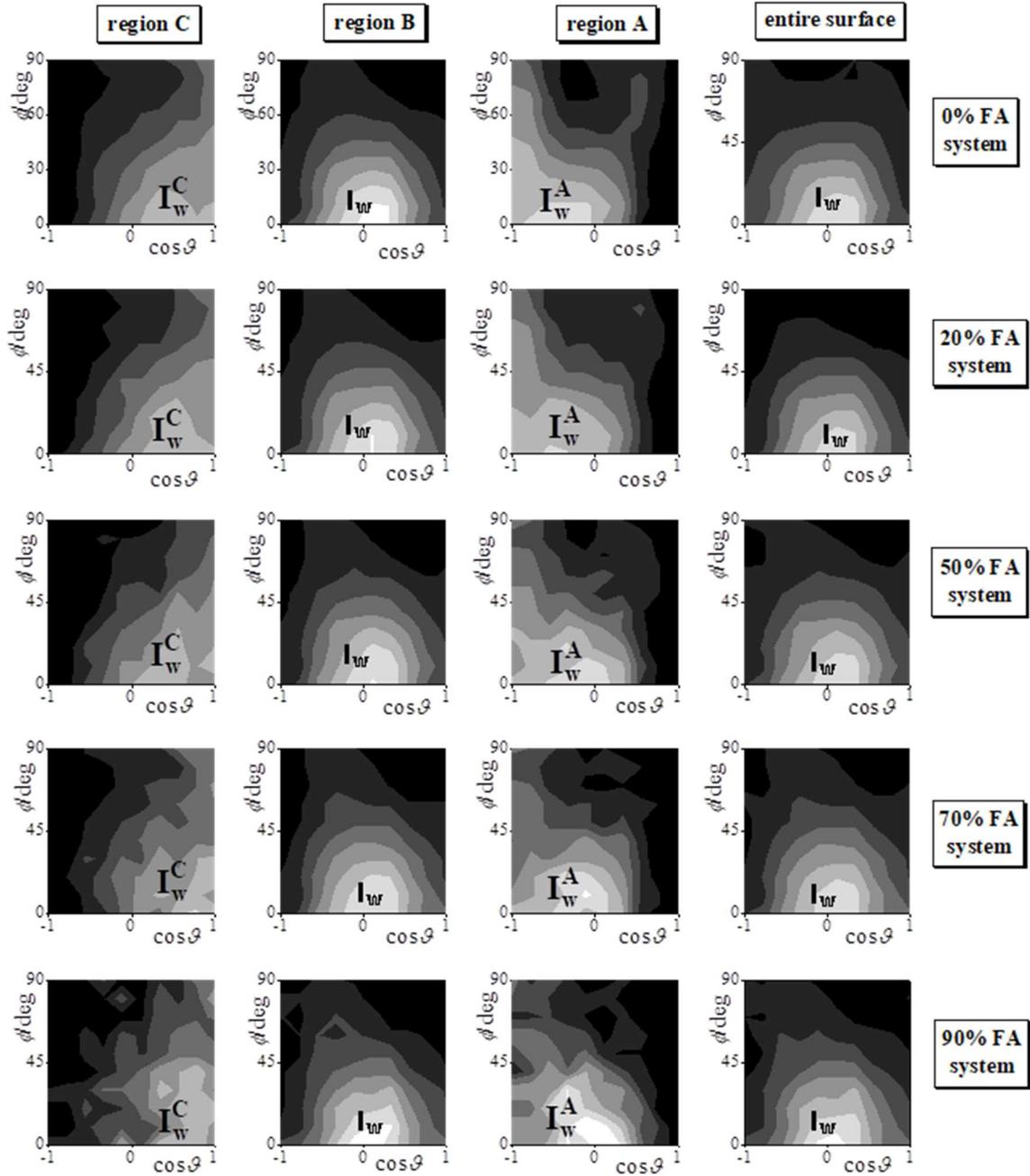
In order to investigate also the effect of the local curvature of the liquid surface on the preferred orientations, we have also calculated the  $P(\cos\mathcal{G},\phi)$  orientational distributions, besides the entire surface layer, also in its separate regions A, B, and C. These regions are defined through the mass density profile of the surface layer as follows. Region B covers the  $X$  range in which the density of the surface layer exceeds half of its maximum value, while regions C and A cover the  $X$  ranges located closer to and farther from the bulk liquid phase, respectively, than that of region B. Thus, region C covers typically the negatively curved troughs, while region A the positively curved crests of the molecularly rough liquid surface. The definition of zones A, B, and C of the surface layer is illustrated in Figure 4.13.



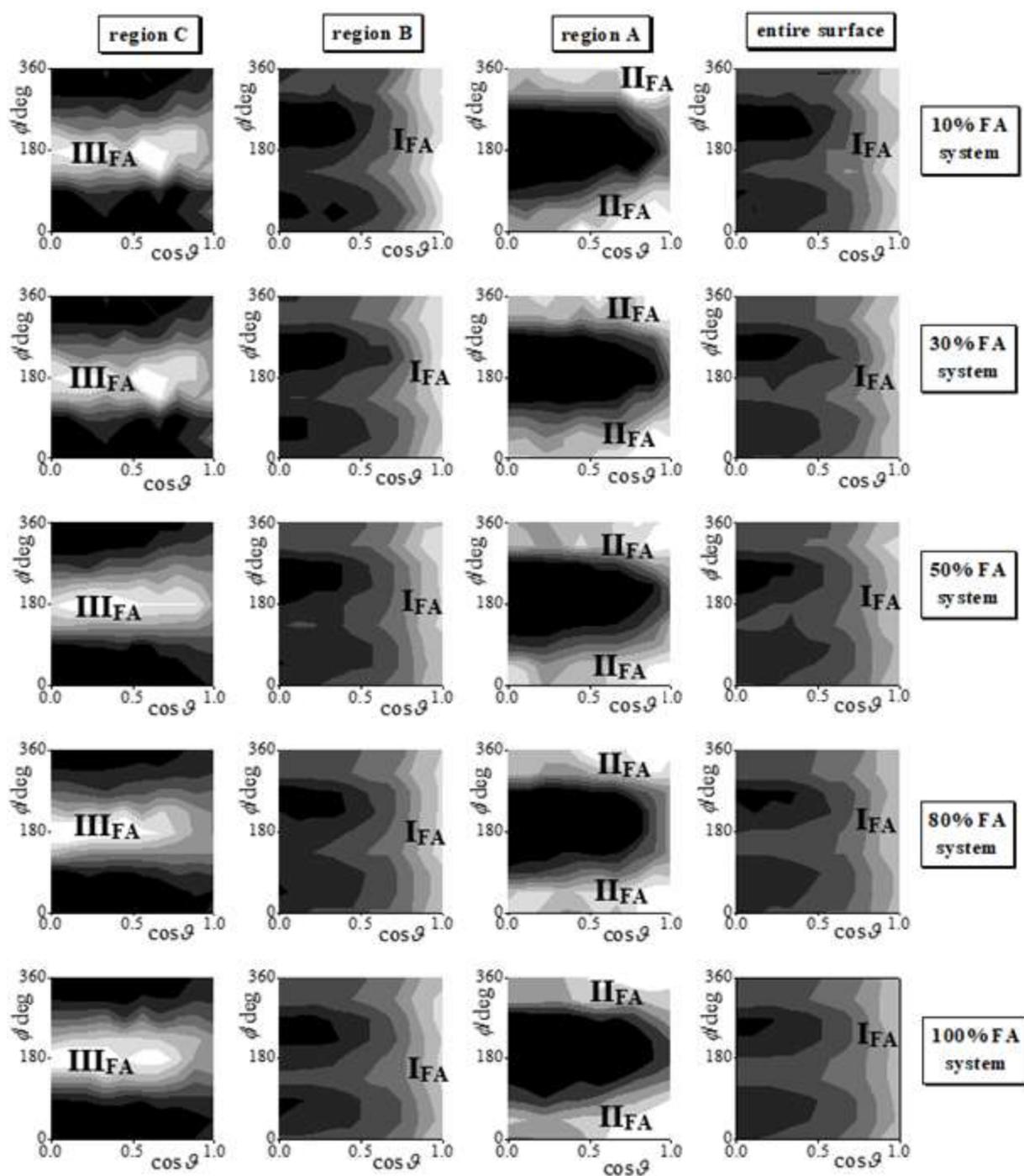
*Figure 4.13: Top: definition of the local Cartesian frames fixed to the individual water (left) and formamide (right) molecules, and of the polar angles  $\vartheta$  and  $\phi$  in these frames. Bottom: definition of the regions A, B, and C of the surface layer.*

The  $P(\cos \vartheta, \phi)$  orientational maps of the water and FA molecules are shown in Figures 4.14 and 4.15, respectively, as obtained in the entire surface layer as well as in its separate regions A, B, and C in selected systems. The preferred orientation of the water molecules in the entire surface layer as well as in its highest populated region, B, corresponds to  $\cos \vartheta = 0$  and  $\phi = 0^\circ$ , while that of the formamide molecules to  $\cos \vartheta = 1$ . (It should be noted that in the case of  $\cos \vartheta = 1$  the vector  $\underline{X}$  is projected to a single point in the  $xy$  plane of the local frame, see Figure 4.13, and hence all points of the  $P(\cos \vartheta, \phi)$  orientational map laying along the  $\cos \vartheta = 1$  line correspond to the same orientation.) These preferred orientations, denoted here as  $I_w$  (for water) and  $I_{FA}$  (for formamide) and illustrated in Figure 4.16.a, correspond to the parallel alignment of the respective molecules with the macroscopic plane of the surface,  $YZ$ . The planar FA molecules can donate and accept hydrogen bonds within their molecular plane. Further, in cases of water-donated H atoms, the donor water molecule can still lay in this plane, while in cases of water-accepted hydrogen-bonds the water molecule only has to slightly twist out from this plane, as reflected in the small shift of the  $I_w$  peak of the water orientation map to

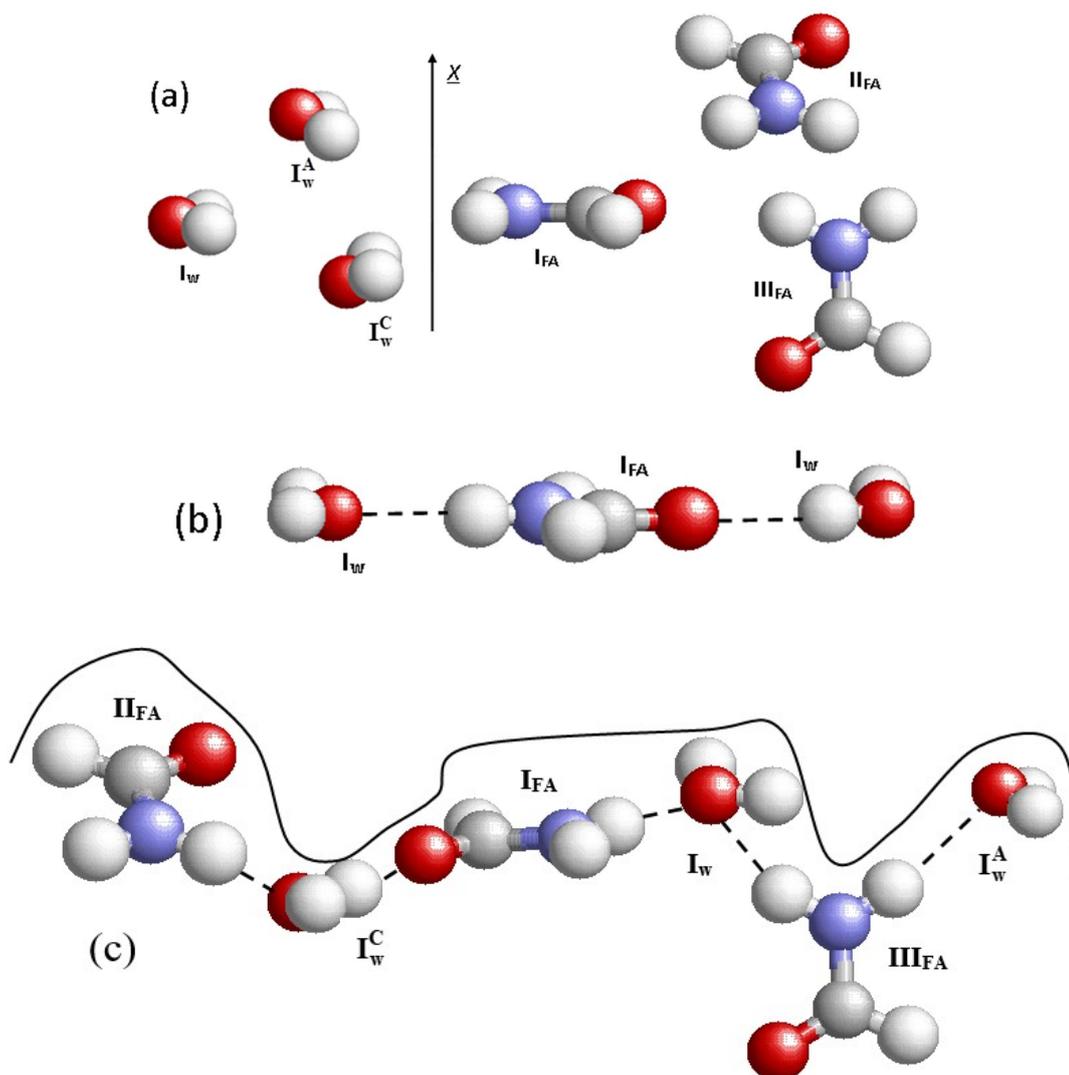
somewhat larger  $\cos\vartheta$  values with increasing FA concentration. Thus, the observed preferential alignment of both molecules with the macroscopic plane of the surface enables them to form a strong hydrogen-bonding network at the liquid surface. The possible hydrogen bonding alignments of a FA and a water molecule, both of which are aligned in their preferred surface orientation, are illustrated in Figure 4.16.b.



**Figure 4.14:** *Oriental maps of the water molecules in the entire surface layer (last column) as well as in separate regions C (first column), B (second column), and A (right column). For the definition of these regions, see the text and Fig. 4.13. Lighter shades of grey correspond to higher probabilities. The notation of the orientations corresponding to the different peaks is also indicated.*



*Figure 4.15: Orientational maps of the formamide molecules in the entire surface layer (last column) as well as in its separate regions C (first column), B (second column), and A (right column). For the definition of these regions, see the text and Fig. 4.13. Lighter shades of grey correspond to higher probabilities. The notation of the orientations corresponding to the different peaks is also indicated.*



*Figure 4.16: (a) Illustration of the orientations preferred by the water and formamide molecules in the entire surface layer as well as in any of its separate regions A, B, and C, relative to the surface normal vector (pointing, by our convention, towards the vapor phase),  $X$ . (b) Illustration of the possible hydrogen bonds formed by a water and a formamide molecule, both aligned according to their main orientational preference, in the surface layer. (c) Illustration of possible hydrogen bonding alignments of water and formamide molecules, both oriented according to their preferences, in portions of different local curvatures of the surface layer. The solid curved is a schematic representation of the liquid surface.*

Both molecules prefer somewhat different orientations at the tips of the crests (in region A) and in the bottom of the troughs (in region C), i.e., at the strongly positively and negatively curved portions, respectively, of the molecularly rough liquid surface. Thus, the peak of the water orientation map is shifted to negative  $\cos \vartheta$  values in region A, and to positive  $\cos \vartheta$  values in region C. In the corresponding orientations, marked as  $I_w^A$  and  $I_w^C$ , respectively, and shown also in Figure 4.16.a, the water molecule is tilted from the parallel alignment with the  $YZ$  plane, and points towards the bulk liquid phase

(in orientation  $I_{\text{w}}^{\text{A}}$ ) and towards the vapor phase (in orientation  $I_{\text{w}}^{\text{C}}$ ) with its H atoms. Further, in region A the peak of the preferred orientation of the FA molecules is shifted to smaller  $\cos \vartheta$  values along the  $\phi = 0^\circ$  (or  $\phi = 360^\circ$ , being equivalent with the former, see Figure 4.13) line. In this orientation, marked here as  $\text{II}_{\text{FA}}$ , the plane of the formamide molecule is tilted somewhat from the  $YZ$  plane, pointing with the CHO group to the vapor, and with the  $\text{NH}_2$  group to the liquid phase. In region C, the preferred orientation of the FA molecules corresponds to  $\cos \vartheta = 0$  and  $\phi = 180^\circ$ , and highly preferred alignments also extend to larger  $\cos \vartheta$  values. The corresponding orientation, marked here as  $\text{III}_{\text{FA}}$ , is perpendicular to the macroscopic plane of the surface, and the  $\text{NH}_2$  group points straight to the vapor phase. The extension of this peak to larger  $\cos \vartheta$  values corresponds to the tilt of the molecule from the perpendicular alignment in such a way that, within the constraint set by the alignment of the molecular plane, the  $\text{NH}_2$  group still points as straight to the vapor phase as possible. Orientations  $\text{II}_{\text{FA}}$  and  $\text{III}_{\text{FA}}$  are also illustrated in Figure 4.16.a.

The orientational preferences observed in regions A and C are all governed by the requirement of maximizing the number of possible hydrogen bonds a surface molecule can form. Thus, in orientation  $I_{\text{w}}^{\text{A}}$  the water molecule sacrifices one of its hydrogen-bonds, while three of its possible hydrogen bonding directions (i.e., two O-H bonds and a lone pair) point flatly toward the bulk liquid phase, straddling along the locally positively curved surface. Conversely, in orientation  $I_{\text{w}}^{\text{C}}$  one hydrogen bonding direction (i.e., a lone pair) points straight to the bulk liquid phase, while the other three ones straddle along the locally negatively curved surface, pointing flatly to the vapor phase. In this way, water molecules oriented in alignments  $I_{\text{w}}^{\text{A}}$  and  $I_{\text{w}}^{\text{C}}$  in regions A and C, respectively, can safely maintain three and four hydrogen-bonds, respectively, with their neighbors.

Similarly, in alignment  $\text{II}_{\text{FA}}$  the formamide molecule sacrifices one of the two possible hydrogen-accepting directions of its O atom, while the other one as well as the two N-H bonds point flatly toward the bulk liquid phase. In alignment  $\text{III}_{\text{FA}}$ , preferred at locally negatively curved portions of the surface, the two N-H bonds point flatly toward the vapor phase, straddling along the curved surface. Thus, FA molecules in alignment  $\text{II}_{\text{FA}}$  in region A can maintain three, while those of alignment  $\text{III}_{\text{FA}}$  in region C four

hydrogen bonds with their neighbors. It is also clear that a FA molecule of alignment  $\text{II}_{\text{FA}}$  can easily form a hydrogen-bond with a water molecule of  $\text{I}_{\text{W}}^{\text{C}}$ , while a formamide of orientation  $\text{III}_{\text{FA}}$  can hydrogen bond to a water of  $\text{I}_{\text{W}}^{\text{A}}$ . Possible hydrogen bonding patterns of neighboring surface FA and water molecules, both aligned in one of their preferred orientations, and their relation with the local curvature of the molecularly rough surface are illustrated in Figure 4.16.c.

#### 4.2.5 Dynamics of the Surface Molecules

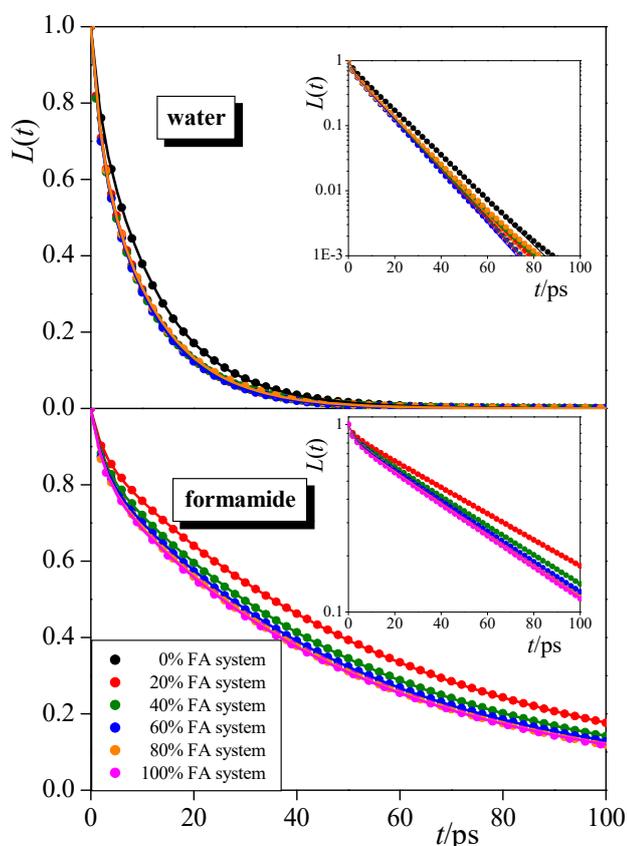
To investigate the dynamics of the molecules within the surface layer, we have calculated their survival probability and mean residence time as well as their lateral diffusion coefficient,  $D_{\parallel}$ , within the surface layer. The survival probability,  $L(t)$ , is defined as the probability that a molecule that belongs to the surface layer at  $t_0$  will stay in this layer up to  $t_0+t$ . Since the departure of the molecules from the liquid surface is governed by first order kinetics, the  $L(t)$  survival probability is expected to follow exponential decay, and hence can be fitted by the function  $\exp(-t/\tau_{\text{surf}})$ , where  $\tau_{\text{surf}}$  is the mean residence time of the molecules in the surface layer. However, besides leaving the liquid surface permanently, the molecules may departure from the surface layer also temporarily, due to a fast oscillatory move. Therefore, the  $L(t)$  data can be well fitted by the sum of two decaying exponentials, i.e., by the function

$$L(t) = A_1 \exp(-t/\tau_{\text{osc}}) + A_2 \exp(-t/\tau_{\text{surf}}) . \quad (4.4)$$

The first term of these two exponentials corresponds to the fast oscillatory move,  $\tau_{\text{osc}}$  being its characteristic time, while the second one corresponds to the permanent departure of the molecules from the surface.

The  $L(t)$  data obtained both for the water and FA molecules in selected systems are shown in Figure 4.17. To emphasize the exponential decay of the curves, the insets show the same data on a logarithmic scale. The sum of the two exponential functions fitted to these data is also indicated. As is seen, the obtained fit is almost perfect in every case. The characteristic time of the fast, oscillatory move,  $\tau_{\text{osc}}$ , is always found to be  $2.0 \pm 0.5$  ps, being smaller for the water than the FA molecules. The mean surface residence time values,  $\tau_{\text{surf}}$ , are collected in Table 4.1 for both molecules. As is seen,

formamide molecules stay considerably, i.e., 4-6 times longer at the liquid surface than waters. Further, while the mean surface residence time of the water molecules does not show any apparent composition dependence (although it is slightly larger in neat water than in the mixed systems), the  $\tau_{\text{surf}}$  value of FA clearly decreases with increasing formamide concentration, being about 30% smaller at the surface of neat formamide than in the 10% FA system. This finding indicates that the presence of water stabilizes the stay of the formamide molecules at the surface of the liquid phase of their mixtures.

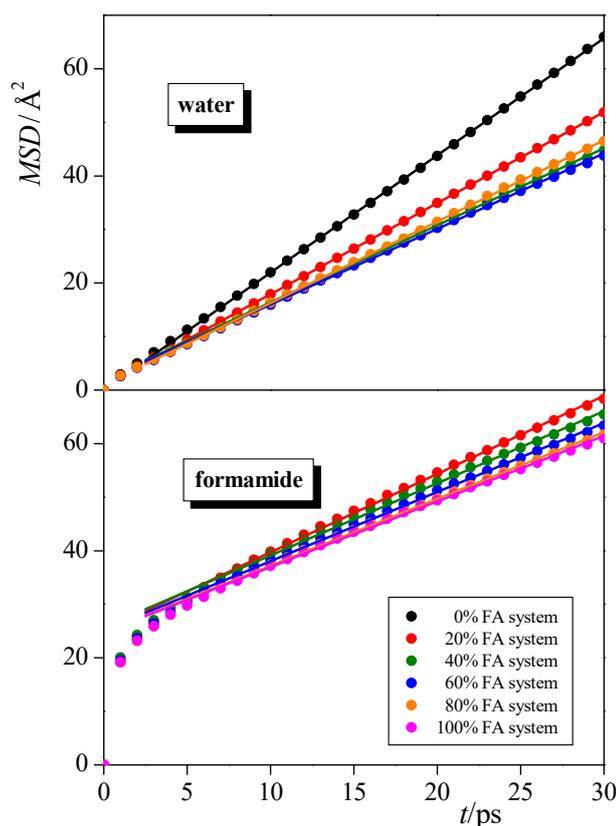


*Figure 4.17: Survival probability of the water (top panel) and formamide (bottom panel) molecules within the surface layer in the various systems simulated (symbols). The color coding of the various systems is the same as in Fig.4.8. Solid lines show the best fit to the simulation data by the sum of two exponentially decaying functions. The insets show the same data on a logarithmic scale, to emphasize the exponential decay of the survival probability.*

The lateral diffusion coefficient of the molecules,  $D_{\parallel}$ , can be obtained through the Einstein relation,<sup>138</sup> i.e.,

$$D_{\parallel} = \frac{MSD}{4t}, \quad (4.5)$$

where  $MSD$  is the mean square displacement of the molecules within the macroscopic plane of the liquid surface,  $YZ$ , by fitting a straight line to the calculated  $MSD$  vs.  $t$  data. The time evolution of the  $MSD$  of the individual molecules has only been calculated during which they have remained in the surface layer. The linear fitting of the  $MSD(t)$  data has been done, however, only above 5 ps, in order to ensure that the molecules are no longer in the ballistic regime, and their mobility is already governed by diffusive motion. The  $MSD(t)$  data corresponding both to the water and the FA molecules, together with the straight lines fitted to them, are shown in Figure 4.18 as obtained in selected systems. Further, the  $D_{\parallel}$  values obtained for both types of molecules from eq. 4.5 are summarized in Table 4.1.



**Figure 4.18:** Mean square displacement of the water (top panel) and formamide (bottom panel) molecules within the macroscopic plane of the surface,  $YZ$ , as a function of time (symbols). The color coding of the various systems is the same as in Fig.4.8. Solid lines show the straight lines fitted to the simulation data above 5 ps.

In analyzing the lateral diffusion of the molecules within the surface layer, first the relation of their mean surface residence time,  $\tau_{\text{surf}}$ , and the characteristic time of their diffusion,  $\tau_{\text{D}}$ , has to be checked. Thus, in cases when  $\tau_{\text{surf}}$  is larger than  $\tau_{\text{D}}$ , the molecules exhibit considerable diffusion within the surface layer during their stay there. On the other hand, if the relation  $\tau_{\text{D}} > \tau_{\text{surf}}$  holds, surface diffusion cannot be discussed meaningfully, as the surface lifetime of the molecules is simply too short for considerable diffusion occurring within this time. In this case the mobility of the molecules is inherently three-dimensional, not being limited to the liquid surface even for a short time. The characteristic time of the two dimensional lateral diffusion can be calculated as<sup>232,241</sup>

$$\tau_{\text{D}} = \frac{A_{\text{m}}}{4D_{\parallel}}, \quad (4.6)$$

where

$$A_{\text{m}} = \frac{2L_{\text{Y}}L_{\text{Z}}}{\langle N_{\text{surf}} \rangle}, \quad (4.7)$$

is the average area per molecule in the surface layer,  $L_{\text{Y}}$ ,  $L_{\text{Z}}$ , and  $\langle N_{\text{surf}} \rangle$  being the lengths of the basic box in the  $Y$  and  $Z$  directions and the mean number of the surface molecules, respectively, and the factor of 2 in eq. 4.7 accounts for the two interfaces present in the basic box.

The  $\tau_{\text{D}}$  values obtained for the water and FA molecules in the different systems simulated are also included in Table 4.1. As is seen, for water,  $\tau_{\text{surf}}$  is larger than  $\tau_{\text{D}}$  only at the surface of the neat system; these two time scales are roughly equal in the 10% FA system, while in the systems of higher formamide concentration the lateral diffusion of the water molecules clearly occur on a larger time scale than their mean surface residence time, indicating that water molecules typically leave the surface of these systems before they could make considerable diffusion there. Since the mean surface residence time of the water molecules is insensitive to the composition of the system, the reason for this behavior is the increase of  $\tau_{\text{D}}$  with increasing formamide concentration. In other words, our finding indicates that formamide molecules stabilize the water molecules at their position in the surface during the entire time scale of their stay at the surface of such mixtures. It is also seen that, for formamide,  $\tau_{\text{surf}}$  is always noticeably, i.e., by a factor of 2-5 larger than  $\tau_{\text{D}}$ , indicating that formamide molecules do diffuse at the liquid surface of

their aqueous mixtures. The lateral surface diffusion coefficient of formamide is rather insensitive to the composition, as it only exhibits a slight (i.e., about 30%) decrease in the entire composition range with increasing FA mole fraction. It is also seen that the surface diffusion of the FA molecules is considerably slower in their aqueous mixtures than that of the water molecules at the surface of neat water, i.e., where they do exhibit a noticeable diffusion at the liquid surface.

### **4.3 Adsorption of Formamide at the Surface of Amorphous and Crystalline Ices under Interstellar and Tropospheric Conditions.**

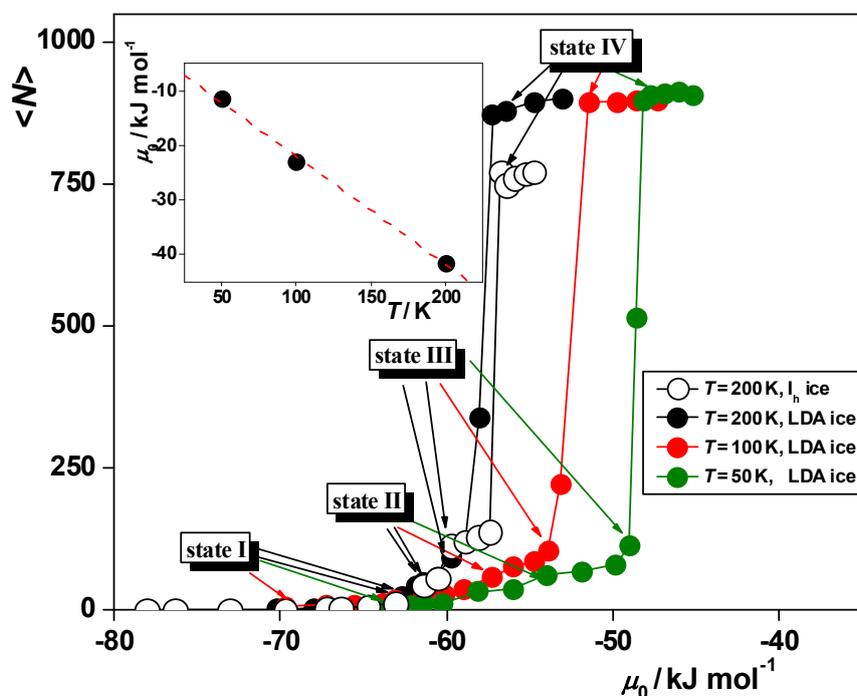
Having determined the best available force field combination for the investigation of formamide-water systems, we have studied in detail the vapor-liquid interface of such mixtures, as discussed in the previous chapter. Another important area concerning the investigation of this binary system is the adsorption of formamide at various ice surfaces. Details of the Monte Carlo simulation performed in this investigation have been described in chapter 3.3. Here we discuss in detail the results obtained from this investigation.

#### ***4.3.1. Adsorption Isotherms***

The adsorption isotherms obtained from the simulations are shown in Figure 4.19 in the form of number of adsorbed formamide molecules as a function of their chemical potential. Since negligible fraction of the formamide molecules are either dissolved in the bulk ice phase or stay isolated from the adsorption layer in the vapor phase in every case, we simply took the total number of formamide molecules in the basic box as the number of the adsorbed molecules. As is seen, the isotherm obtained on  $I_h$  ice rises continuously up to the chemical potential value of about  $-60$  kJ/mol. This rising part is followed by a plateau in the  $\mu$  range of about  $-60$  to  $-57$  kJ/mol, corresponding to an adsorption monolayer. This monolayer is rather stable, as evidenced by the fact that it exists in an about 3 kJ/mol wide range of chemical potentials. The plateau is followed by a very steeply rising part of the isotherm, corresponding to multilayer adsorption. This steeply rising part of the isotherm ends at the chemical potential value of about  $-56$  kJ/mol, where the basic box becomes filled with formamide molecules.

The isotherm obtained at 200 K on LDA ice is very similar to that corresponding to crystalline ice, with two notable differences. First, the plateau observed between  $-60$

and  $-57$  kJ/mol on  $I_h$  ice is missing on LDA ice; instead, the isotherm rises continuously even in this  $\mu$  range, exceeding the isotherm corresponding to  $I_h$  ice. This finding suggests that the corrugated geometry of the LDA surface, resulting also in corrugations of the first adsorption layer, promotes multilayer adsorption. As a consequence, multilayer adsorption on LDA starts at  $\mu$  values at which the adsorbed monolayer is still stable on  $I_h$  ice. Second, the basic box can host considerably more formamide molecules in the presence of LDA than  $I_h$  ice. This difference simply reflects that the density of LDA ice is higher than that of  $I_h$  ice. The shape of the  $\langle N \rangle (\mu)$  isotherm does not change noticeably with the temperature; it is simply shifted to higher chemical potential values as the temperature decreases.



*Figure 4.19: Adsorption isotherms of formamide, in the form of number of adsorbed molecules vs chemical potential, at the surface of  $I_h$  ice at 200 K (black open symbols) as well as at the surface of LDA ice (full symbols) at 200 K (black), 100 K (red), and 50 K (green), as obtained from our simulations. The lines connecting the points are just guides to the eye. The state points at which sample configurations have been dumped for detailed analyses are marked by the arrows (see the text). The inset shows the temperature dependence of the chemical potential corresponding to the point of condensation. Black circles indicate simulation results, and red line is a straight line fitted to these data.*

An important feature of the obtained isotherms is that they also have a point on their very sharply rising part, preceding immediately the filling of the basic box, in most of the cases. Usually this very sharp rise of the isotherm is interpreted as the condensation of the adsorbate. Our present finding indicates, however, that condensation is preceded by the occurrence of a thick, multimolecular adsorption layer. Therefore, the possibility that the thickness of the adsorption layer of formamide can be comparable with the width of the slab between the two ice surfaces in the basic box (being about 30 Å), or might even well exceed it, cannot be excluded. If this is indeed the case, the filling of the basic box simply indicates the point at which the adsorption layer gets as thick as this slab rather than the point of condensation of formamide. To check this possibility and determine the real point of condensation without the above finite size effect error, we have performed a set of GCMC simulations of neat formamide (i.e., in the absence of the ice phase) in the same basic box at each of the three temperatures considered. Since the point of condensation is a value characteristic to the adsorbate itself, being independent from the type and even from the presence or absence of any adsorbent (at least as long as its surface is macroscopically flat), its value can unambiguously be determined in the lack of the ice phase. Indeed, we observed a very sharp transition from a practically empty to a filled basic box at a given chemical potential value at every temperature considered. Thus, the chemical potential value corresponding to the point of condensation,  $\mu_0$ , has turned out to be -41.521 kJ/mol at 200 K, -22.947 kJ/mol at 100 K, and -11.432 kJ/mol at 50 K, indeed well above the  $\mu$  values corresponding to the sharply rising part of the isotherms. The temperature dependence of the obtained  $\mu_0$  values is shown in the inset of Figure 4.19. As is seen, the  $\mu_0(T)$  points are laying along a straight line. Considering the fact that, for a one component system, the chemical potential is equivalent with the molar free energy and hence, apart from a minus sign, its temperature derivative with the molar entropy,  $S^m$  (i.e.,  $S^m = -(\partial\mu/\partial T)$ ), our result suggests that the molar entropy of formamide is constant (i.e., temperature independent) at the point of condensation, being  $198.5 \pm 11.0$  J/mol K.

In order to further analyze the adsorption isotherms, we have converted them to the more conventional  $\Gamma$  vs.  $p_{\text{rel}}$  form, where  $\Gamma$  is the surface density of formamide, calculated as

$$\Gamma = \frac{\langle N \rangle}{2YZ}, \quad (4.8)$$

where, the factor 2 in the denominator reflects the presence of two ice surfaces in the basic box. Further,  $p_{\text{rel}}$  is the relative pressure of the vapor phase (i.e., its pressure,  $p$ , normalized by that of the saturated vapor,  $p_0$ ). The value of  $p_{\text{rel}}$  can be calculated based on an ideal gas equation from that of  $\mu$ , using also the chemical potential value corresponding to the point of condensation,  $\mu_0$ , as<sup>242</sup>

$$p_{\text{rel}} = \frac{p}{p_0} = \frac{\exp(\mu / k_{\text{B}}T)}{\exp(\mu_0 / k_{\text{B}}T)}, \quad (4.9)$$

where  $k_{\text{B}}$  stands for the Boltzmann factor. It should be noted that in converting the  $\langle N \rangle(\mu)$  isotherms to the  $\Gamma(p_{\text{rel}})$  form, we have omitted the points corresponding to filled basic box (as they are affected by finite size effect error).

The obtained  $\Gamma(p_{\text{rel}})$  isotherms are shown in Figure 4.20. As is seen, at 200 K, the isotherms start rising almost linearly; this rise gradually decreases turning into a plateau, and the plateau part is followed by a second, steeply rising part of the isotherms. The first rising part corresponds to the gradual building up of the first molecular layer; the plateau reflects the presence of a more or less saturated monolayer, while the second, steeply rising part corresponds to multilayer adsorption. At lower temperatures, the plateau of the isotherm gets progressively shorter and occurs at lower surface densities, suggesting that multilayer adsorption starts before the complete saturation of the first molecular layer. Thus, the obtained isotherms represent a transition between type II and type III isotherms (both describing multilayer adsorption) according to the IUPAC convention. Considering that type II isotherms correspond to systems where the adsorbate–adsorbent interaction is considerably stronger, while type III isotherms correspond to systems where it is considerably weaker than the lateral interaction between the adsorbed molecules, this result is in a clear accordance with earlier claims<sup>110,111,236</sup> and with our previous findings that the mixing of water and formamide is nearly ideal, and hence there is no marked difference between the interactions of the like and unlike molecules. Furthermore, it is

also seen that the steepness of the first, rising part of the isotherm, corresponding to the building up of the first molecular layer, increases very strongly with decreasing temperature, which results in a very sharp rise of the low temperature isotherms even at extremely low  $p_{\text{rel}}$  values. We tried to fit the  $\Gamma(p_{\text{rel}})$  data up to their plateau region by the Langmuir isotherm,<sup>243,244</sup> i.e.,

$$\Gamma = \Gamma_{\text{max}} \frac{p_{\text{rel}}K}{1 + p_{\text{rel}}K}, \quad (4.10)$$

where the parameters  $\Gamma_{\text{max}}$  and  $K$  are the surface density of the saturated monolayer and the Langmuir partition coefficient, measuring the partitioning of the adsorbate molecules between the adsorption layer and the vapor phase, respectively. However, we could not get a reasonable fit in any case (see Figure 4.20), suggesting that there might well be strong lateral interaction between the adsorbed formamide molecules. This point will be investigated in detail later in this chapter. Nevertheless, the Langmuir function can still be fitted to the nearly linearly rising part of the isotherms (i.e., to the points of very low  $p_{\text{rel}}$ , corresponding to considerably smaller  $\Gamma$  values than that of the plateau region) in the simulated data sets. Although the  $\Gamma_{\text{max}}$  values resulting from these fits are meaningless, as they are affected by far too large numerical error, the  $K$  values can still serve at least as rough estimates of the formamide partitioning between the adsorption layer and the vapor phase. This way, the value of  $K$  has turned out to be about  $5 \times 10^4$  at 200 K,  $6 \times 10^{18}$  at 100 K, and  $10^{43}$  at 50 K. Although these values are still affected by large error bars, their very high orders of magnitude indicate strongly increasing affinity of formamide to the LDA surface with decreasing temperatures.

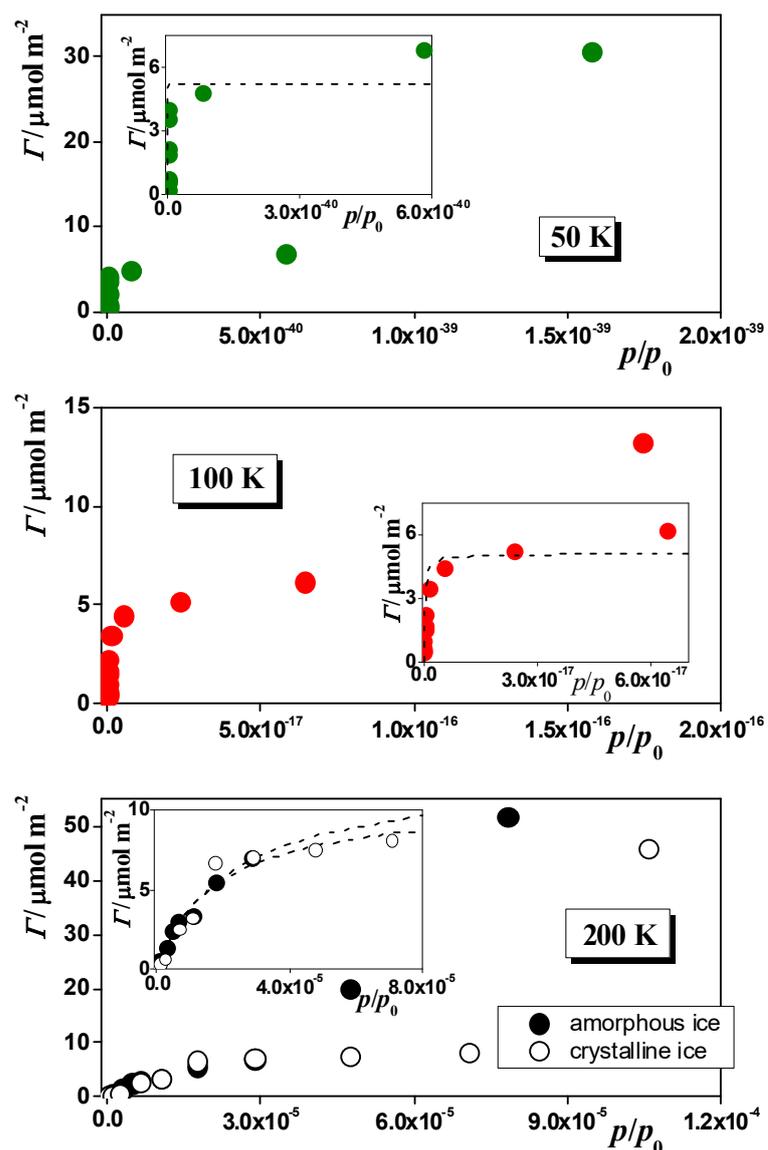


Figure 4.20: Adsorption isotherms of formamide, in the form of surface density vs relative pressure, at the surface of  $I_h$  ice at 200 K (bottom panel, open symbols) as well as at the surface of LDA ice (full symbols) at 200 K (bottom panel, black), 100 K (middle panel, red), and 50 K (top panel, green), as obtained from our simulations. The insets show these data up to the point where the adsorption layer is still monomolecular (symbols), together with their best Langmuir fits (dashed curves).

On the basis of the behavior of the obtained isotherms, we have selected four chemical potential values in each system, at which sample configurations have been collected for detailed analyses. The lowest of these chemical potential (referred to here as state I) always corresponds to only a few adsorbed molecules per surface. At the second  $\mu$  value chosen (state II), formamide molecules form an unsaturated monolayer but cannot be isolated from each other, while at the third  $\mu$  value (state III) they form a more or less saturated monolayer. Finally, the largest of the chosen chemical potential values (state IV) corresponds to multilayer adsorption (i.e., when the basic box is already filled with

formamide molecules) in every case. The chemical potential values corresponding to these states are indicated in Tables 3.5-3.8 as well as in Figure 4.19.

### ***4.3.2 Density Profiles***

The mass density profiles of all formamide molecules as well as of only those forming the first molecular layer at the ice surface along the surface normal axis,  $X$ , are shown in Figure 4.21, as calculated in states I–IV in the various systems considered. The obtained profiles clearly confirm our earlier claim that no substantial amount of formamide is located in the vapor and in the bulk ice phase. For reference, the mass density profile corresponding to the outmost portion of the ice phase is also indicated, while the inset of the figure shows the comparison of the density profiles of the  $I_h$  and LDA ice phases at 200 K. As is seen, while the density profile of  $I_h$  ice shows the clear separation of the subsequent molecular layers in the ice crystal, that of LDA is practically constant inside the bulk phase. It is also seen that the ordered structure of  $I_h$  ice implies also a clear layering of the adsorbed formamide molecules in the case of multilayer adsorption, while this layering is effectively screened by the corrugated surface of LDA ice.

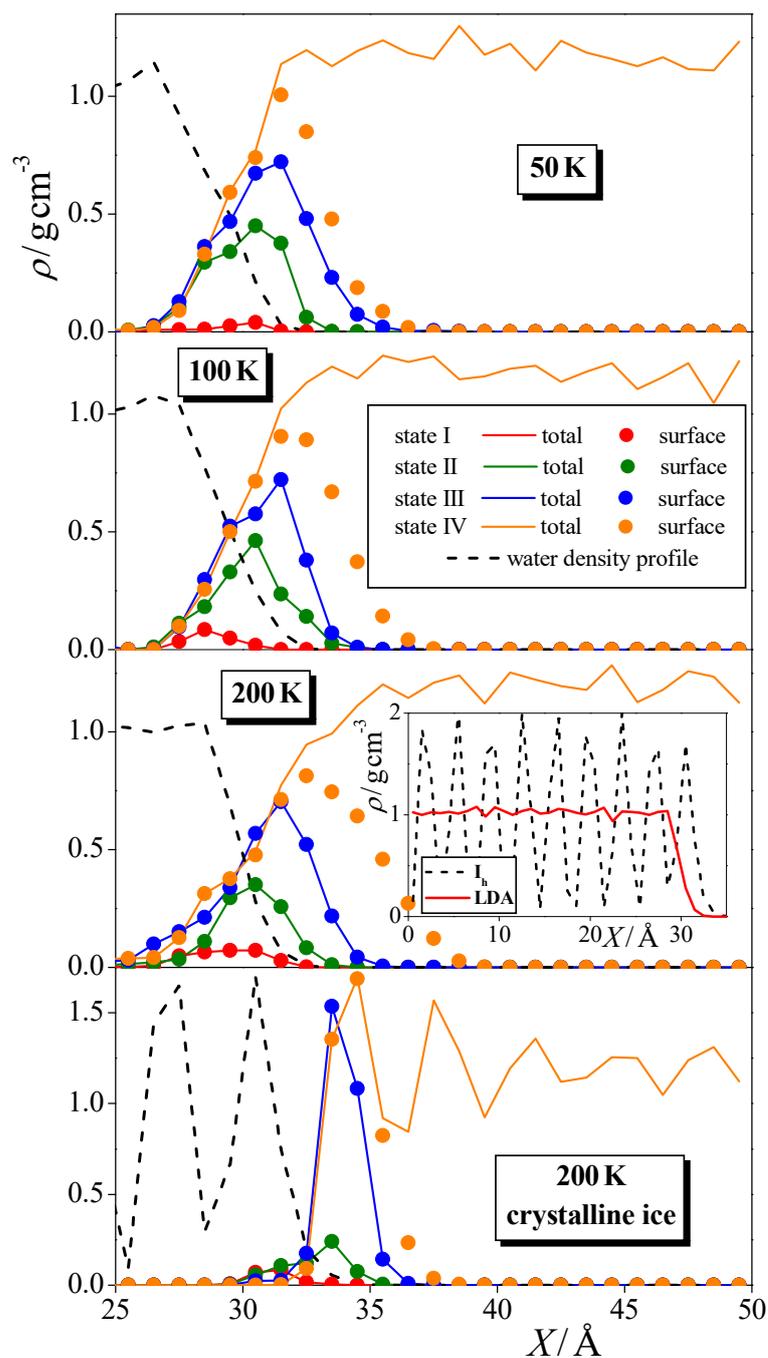


Figure 4.21: Mass density profile of the adsorbed formamide molecules (solid lines) as well as solely of their first molecular layer (circles) in states I (red), II (green), III (blue), and IV (orange), as obtained at the surface of LDA ice at 50 K (top panel), 100 K (second panel), and 200 K (third panel) as well as at the surface of  $I_h$  ice at 200 K (bottom panel). The mass density profile of the surface portion of the ice phase is also indicated for reference (black dashed lines). The inset shows the comparison of the mass density profiles of the  $I_h$  (black dashed line) and LDA (red solid line) ice phases

The comparison of the profiles corresponding to all formamide molecules and to only those belonging to the first molecular layer reveals that not even traces of multilayer

adsorption occur in states I–III. Further, since the saturated adsorption monolayer can be very well estimated by the first molecular layer in state IV, it is also evident that the adsorption monolayer in state III is clearly not yet saturated in any case (see also Figure 3.5). Since state III has always been chosen as the point immediately preceding the sharp rise of the isotherm, this finding is in a clear accordance with our earlier claim that multilayer adsorption starts before the first monolayer gets saturated.

Integration of the density profiles of the first layer formamide molecules in state IV provides the average number of first layer adsorbed molecules, and hence also the surface density of the saturated monolayer,  $\Gamma_{\max}$ . Assuming that the surface area of the LDA phase is independent from the temperature, the value of  $\Gamma_{\max}$  turns out to be  $10.5 \pm 0.7 \mu\text{mol}/\text{m}^2$  for LDA and  $9.4 \mu\text{mol}/\text{m}^2$  for  $I_h$  ice. The fact that the value of  $\Gamma_{\max}$  is significantly larger for LDA than for  $I_h$  ice reflects that considerably more adsorbed molecules can be accommodated at the corrugated surface of amorphous ice than at the flat surface of crystalline ice. Similar results were obtained earlier concerning the adsorption of methylamine at LDA and  $I_h$  ice surfaces.<sup>225,245</sup> Further, the fact that the obtained  $\Gamma_{\max}$  values are considerably larger than the surface densities corresponding to the plateau region of the  $\Gamma(p_{\text{rel}})$  isotherms, scattering between about 6 and  $8 \mu\text{mol}/\text{m}^2$  (see Figure 4.20), stresses again that multilayer adsorption precedes the saturation of the first molecular layer of formamide at the surface of both amorphous and crystalline ice.

Finally, having the  $\Gamma_{\max}$  values estimated, we can attempt to establish a relation between the calculated isotherms and the possible adsorption of formamide on LDA ice under interstellar, and on  $I_h$  ice under tropospheric conditions. For this, the  $p_0$  values corresponding to the temperatures considered are needed to be, at least, estimated. Such estimation can be given by the Antoine equation:<sup>246</sup>

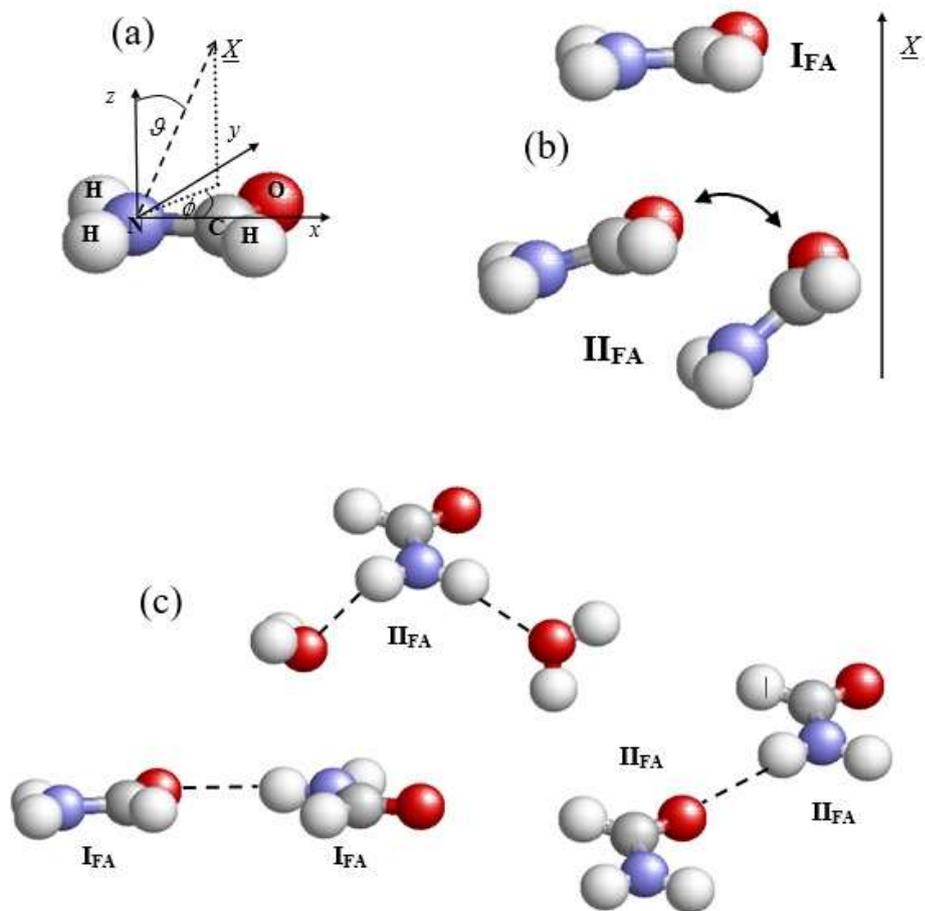
$$\ln p_0 = A - \frac{B}{T + C}, \quad (4.11)$$

where  $A$ ,  $B$ , and  $C$  are the Antoine parameters. Unfortunately, the only Antoine parameter set for formamide we are aware of, i.e.,  $A = 5.9526$ ,  $B = 8937.035 \text{ K}$ , and  $C = 27.655 \text{ K}$ , corresponding to the  $p_0$  value in Pa units, is supposed to provide reliable results only in the temperature range between 343.7 and 483.7 K,<sup>247,248</sup> well above the temperature range of our interest. Nevertheless, using the  $p_0$  value corresponding to the lower boundary of this range of 343.7 K as a very crude but certainly upper estimate of the  $p_0$  value at the

temperatures considered in this study, and converting the pressure values to bulk vapor phase concentration using the ideal gas law, the vapor phase concentration of formamide being in equilibrium with the saturated adsorption monolayer (i.e.,  $\Gamma$  being  $10.5 \mu\text{mol}/\text{m}^2$  for LDA and  $9.4 \mu\text{mol}/\text{m}^2$  for  $\text{I}_h$  ice) can be estimated as  $10^{-17}$  molecules/ $\text{dm}^3$  at 50 K,  $10^5$  molecules/ $\text{dm}^3$  at 100 K, and  $10^{15}$  molecules/ $\text{dm}^3$  at 200 K. Considering that the concentration of formamide is thought to be about  $10^{-1}$ – $10^{-2}$  molecules/ $\text{dm}^3$  at least in some parts of the ISM,<sup>55</sup> we clearly conclude that its adsorption at the LDA surface is strongly preferred at 50 K. Further, given the crudeness of the estimation of  $p_0$  we could make, being several orders of magnitude too high at the temperatures considered here, it is also quite possible that this preference of formamide for adsorption still prevails up to 100 K under interstellar conditions. Further, since the concentration of formamide in the atmosphere is thought to be about  $2.5 \times 10^{13}$  molecules/ $\text{dm}^3$ <sup>249</sup> and taking into account again that our above estimate of  $10^{15}$  molecules/ $\text{dm}^3$  is very probably deviating several orders of magnitude up from the real value, our results strongly suggest that the adsorption of formamide on  $\text{I}_h$  ice under tropospheric conditions is also a strongly preferred process.

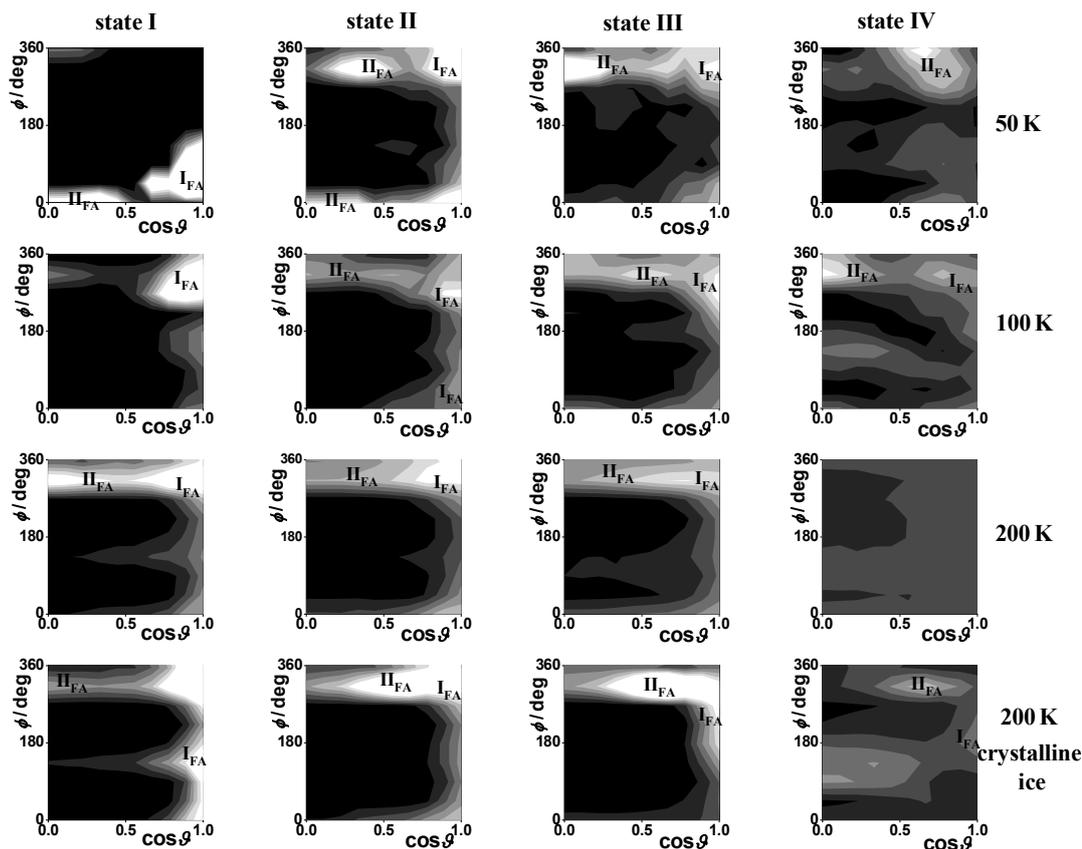
### 4.3.3 Orientation of the First Layer Formamide Molecules

In order to describe the orientation of the first layer formamide molecules relative to the macroscopic plane of the interface, YZ, here we define the local Cartesian frame fixed to the individual formamide molecules in the following way. Its origin coincides with the N atom of the formamide molecule, axis x points along the N–C bond from the N to the C atom, axis z is perpendicular to the plane of the molecule, while axis y is perpendicular to both axes x and z, and it is oriented in such a way that the y coordinate of the O atom is positive. Thus, the polar angle  $\vartheta$  is the angle between the macroscopic surface normal axis X (i.e., edge X of the basic box) and the molecule normal, while angle  $\phi$  is the angle between the projection of X to the molecular plane and the vector pointing from the N to the C atom of the formamide molecule. The definition of this local Cartesian frame as well as that of the polar angles  $\vartheta$  and  $\phi$  is illustrated in Figure 4.22.a. It should be noted that due to the planar symmetry of the formamide molecule, this local frame can always be chosen in such a way that the polar angle  $\vartheta$  does not exceed  $90^\circ$ , and thus  $0 \leq \cos \vartheta \leq 1$ .



*Figure 4.22: (a) Definition of the local Cartesian frame fixed to the individual formamide molecules and of the polar angles  $\vartheta$  and  $\phi$  describing the orientation of the molecule relative to the ice surface (see the text). (b) Preferred orientations of the formamide molecules at the surface of  $I_h$  and LDA ices. (c) Possible H-bonding arrangements of the surface formamide molecules with their formamide neighbors as well as with surface water molecules, when all molecules are aligned in one of their preferred surface orientations. Red, blue, gray, and white balls represent O, N, C, and H atoms, respectively. The dashed lines denote hydrogen bonding, and  $X$  is the macroscopic surface normal*

The  $P(\cos \vartheta, \phi)$  orientational maps of the first layer formamide molecules are shown in Figure 4.23 as obtained from our simulations. As is seen, these maps exhibit two marked peaks, corresponding to the preferred orientations  $I_{FA}$  and  $II_{FA}$  of the formamide molecules (see chapter 4.2.4). The preferred alignments  $I_{FA}$  and  $II_{FA}$  of the formamide molecules are illustrated again in Figure 4.22.b.



*Figure 4.23: Orientational maps of the first layer formamide molecules, as obtained on LDA ice at 50 K (top row), 100 K (second row), and 200 K (third row) as well as on  $I_h$  ice at 200 K (bottom row) in states I (first column), II (second column), III (third column), and IV (fourth column). Lighter shades of gray correspond to higher probabilities. The peaks corresponding to the preferred orientations of the surface formamide molecules ( $I_{FA}$  and  $II_{FA}$ ; see the text) are also marked.*

The preference of the formamide molecules for these orientations can be understood by considering the possible hydrogen bonds they can form with each other as well as with the surface water molecules in these alignments. Since all the hydrogen-bonding directions of the formamide molecule lay also within the molecular plane, formamide molecules of alignment  $I_{FA}$  can easily hydrogen-bond to each other; a formamide molecule of orientation  $I_{FA}$  can form up to four such hydrogen-bonds with its formamide neighbors. Further, since the surface of LDA ice is corrugated on the molecular length scale, two neighboring surface formamide molecules of alignment  $II_{FA}$ , located at different depths along the macroscopic surface normal axis,  $X$ , can also form a hydrogen-bond with each other in such a way that the molecule being farther from the bulk ice phase along this external axis is the hydrogen-donor, while the one being closer to the bulk ice phase is the hydrogen-acceptor partner. These possible hydrogen bonds between two surface formamide molecules, both aligned in one of their preferred orientations, are illustrated in Figure 4.22.c.

Water molecules prefer four distinct orientations at the surface of  $I_h$  ice, in each of which one hydrogen-bonding direction (i.e., an O–H bond or a lone pair) is staying perpendicular to the surface plane.<sup>245</sup> At the surface of LDA ice, these preferences are much less sharp; the corresponding peaks of the orientational map are strongly merged, indicating that surface waters can rather largely deviate from these orientations. As a consequence, the coplanar water alignment with the macroscopic plane of the surface is also preferred here.<sup>245</sup> A formamide molecule of orientation  $I_{FA}$ , laying also parallel with this plane, can thus easily form up to four hydrogen-bonds also with surface waters of this alignment. Moreover, even at the surface of  $I_h$  ice, a slight deviation from the alignment  $I_{FA}$  as well as from a preferred water orientation and/or a slight distortion of the linearity of the hydrogen-bond can also lead to the formation of several hydrogen-bonds between surface waters and  $I_{FA}$  aligned formamides. Furthermore, a formamide molecule of alignment  $II_{FA}$ , being tilted from the surface plane, can easily donate both of its  $NH_2$  hydrogen atoms to surface water molecules aligned in two of their preferred orientations. These possible hydrogen-bonding patterns between surface water and first layer formamide molecules are also illustrated in Figure 4.22.c. The possible hydrogen bonding of the first layer formamide molecules is further investigated in the subsequent subsection.

#### 4.3.4 Energetics of the Adsorption

To characterize the energetic background of the adsorption, we have calculated the  $U_b$  binding energy (i.e., the total interaction energy with the rest of the system) of the first layer formamide molecules. Further, besides  $U_b$ , its contributions coming from the interaction with the ice phase and with the other formamide molecules, denoted here as  $U_b^{ice}$  and  $U_b^{lat}$ , respectively, have also been calculated. The binding energy distributions obtained at 200, 100, and 50 K are shown in the Figure 4.24, 4.25 and 4.26 respectively

The distributions obtained on  $I_h$  and LDA ices (Figure 4.24) are rather similar to each other, with the only notable difference that they are somewhat broader in the case of amorphous than crystalline ice, and this broadening is such that the  $P(U_b^{ice})$  distributions extend to lower, while the  $P(U_b^{lat})$  ones to higher energies in the case of LDA ice. This increased importance of the formamide-water interaction at the surface of amorphous ice

can be well explained by the higher orientational flexibility of the water molecules as well as by the larger amount of accessible hydrogen-bond donor and acceptor water partners at this surface than at that of  $I_h$  ice.<sup>245</sup> The opposite effects of the ice phase on the  $P(U_b^{\text{ice}})$  and  $P(U_b^{\text{lat}})$  distributions largely cancel each other, resulting in almost identical  $P(U_b)$  distributions at the surface of the two ice phases, indicating again that the exchange of the hydrogen-bonded water and formamide neighbors of a formamide molecule has a negligible energetic consequence. In state I, the  $P(U_b^{\text{ice}})$  distribution of the 200 K systems exhibits a single peak around  $-100$  kJ/mol. With increasing chemical potential (i.e., increasing surface coverage), this peak shifts to higher energies, often being split into two distinct peaks or, at least, exhibiting a shoulder. These peaks and shoulders typically occur around the energy values of about  $-75$  kJ/mol,  $-50$  kJ/mol, and  $-25$  kJ/mol; in the case of state IV and  $I_h$  ice even a small peak of  $P(U_b^{\text{ice}})$  around zero energy is seen. Considering that the energy of a hydrogen bond is roughly  $-25$  kJ/mol, this finding indicates that in state I (i.e., at the lowest surface density considered), first layer formamide molecules typically form four hydrogen-bonds with the ice phase, and the number of their hydrogen-bonded water neighbors gradually decreases with increasing surface density.

On the other hand, the  $P(U_b^{\text{lat}})$  distribution exhibits a very high and narrow peak at zero energy in state I; this peak is still clearly seen in state II, and traces of it exist even in state III. This peak corresponds to the adsorbed formamide molecules that are isolated from each other. Interestingly, this distribution also shows a peak at around  $-25$  kJ/mol even in state I on LDA ice, indicating that adsorbed molecules are not completely isolated from each other even in this state of very low surface density (i.e., when less than five formamides are attached, on average, to each of the two ice surfaces in the basic box; see Table 3.7). This behavior is in a clear contrast with our previous observation concerning the adsorption of methylamine on LDA ice<sup>245</sup> and stresses the strong lateral interaction between the adsorbed formamide molecules. It should be noted that, by contrast with the LDA ice, no such peak is seen on  $P(U_b^{\text{lat}})$  in state I of  $I_h$  ice, indicating that in this case the adsorbed molecules are indeed isolated from each other. The increase of the chemical potential and hence that of the surface density shifts the  $P(U_b^{\text{lat}})$  distribution to lower energies, reflecting the fact that adsorbed molecules are increasingly surrounded by each

other. Thus, in LDA ice, the peak of the distribution occurs around  $-25$  kJ/mol,  $-55$  kJ/mol, and  $-80$  kJ/mol in states II, III, and IV, respectively, indicating the ability of the adsorbed formamide molecules to form up to three hydrogen-bonds with each other.

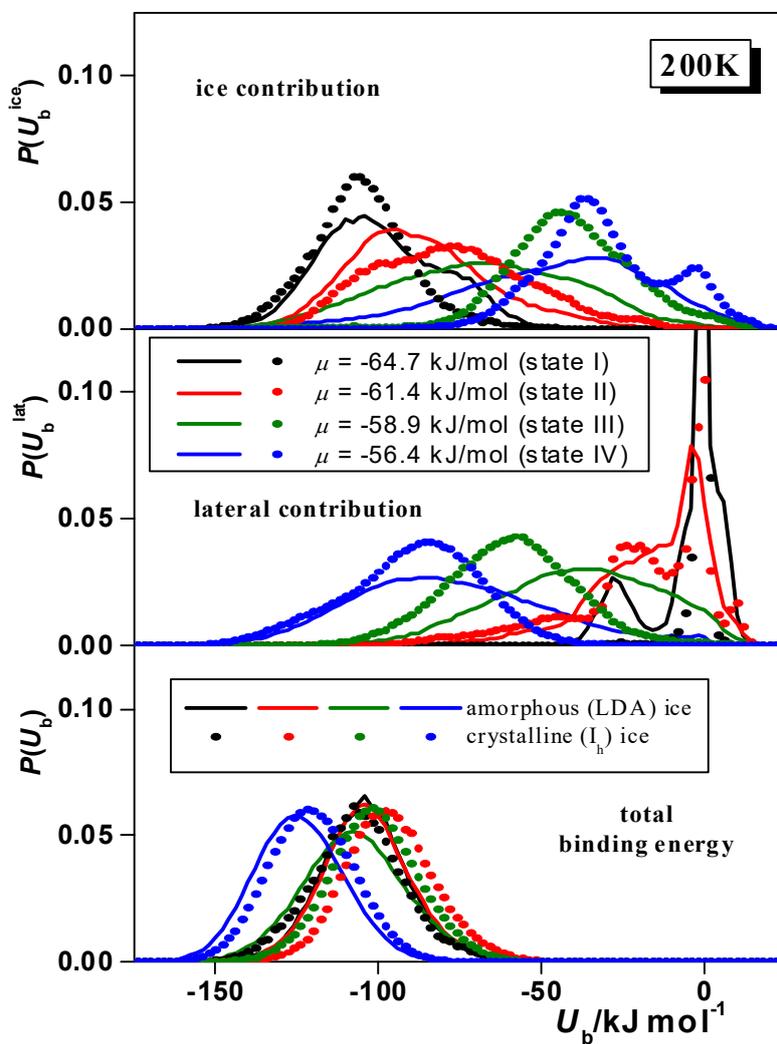


Figure 4.24: Distribution of the total binding energy ( $U_b$ ) of the first layer formamide molecules (bottom panels) as well as its contribution coming from the other formamide molecules in the system (middle panels) and from the ice phase (top panels), as obtained at 200 K, at the surface of  $I_h$  (full circles) and LDA (solid lines) ices in states I (black), II (red), III (green), and IV (blue).

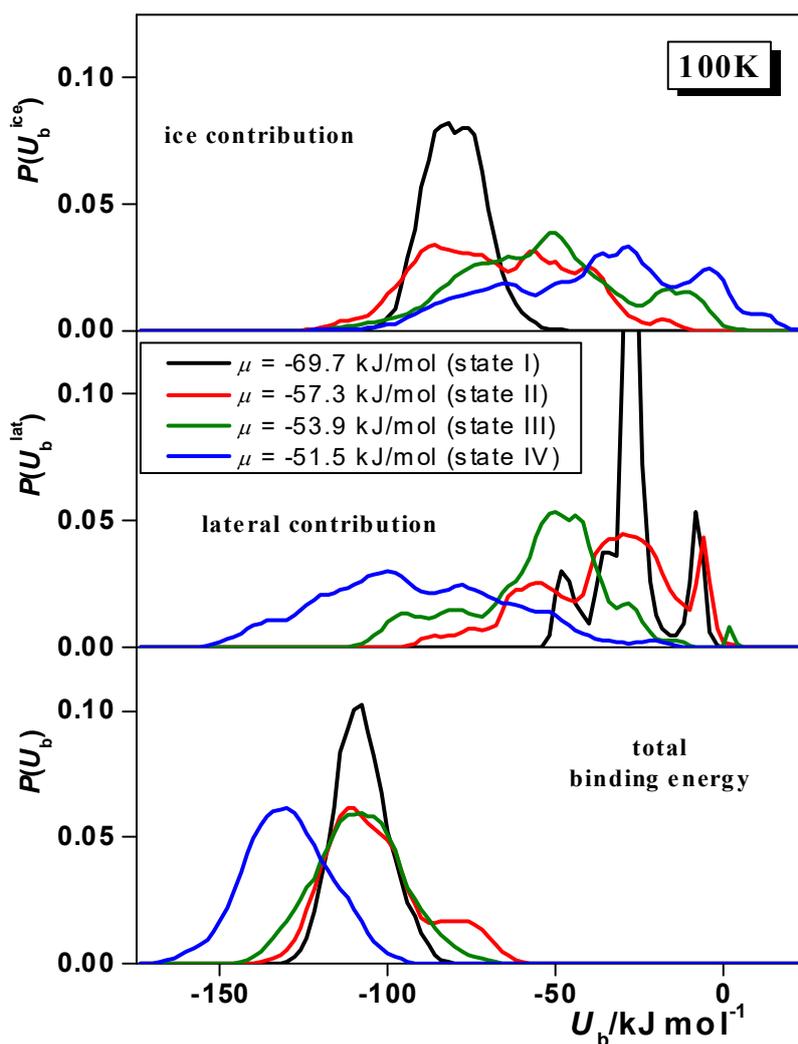


Figure 4.25: Distribution of the total binding energy ( $U_b$ ) of the first layer formamide molecules (bottom panels) as well as its contribution coming from the other formamide molecules in the system (middle panels) and from the ice phase (top panels), as obtained at 100 K, at the surface of LDA ices in states I (black), II (red), III (green), and IV (blue).

The distribution of the total binding energy,  $P(U_b)$ , exhibits a single peak around  $-100$  kJ/mol in states I–III, reflecting the compensation of the decrease of its ice, and increase of its lateral component (in magnitude) with increasing chemical potential, and indicating that adsorbed formamide molecules always prefer to form four hydrogen bonds with their neighbors, irrespective of whether they are waters or formamides. Interestingly, in state IV the  $P(U_b)$  peak occurs around  $-120$  kJ/mol. Since a formamide molecule cannot form more than four hydrogen bonds, this deep binding energy reflects also the strong but non-hydrogen bonding interaction of the first layer formamide molecules with those forming the subsequent, outer molecular layers of the adsorption layer. Considering that first layer

formamides already complete their four possible hydrogen-bonds with surface waters and in-layer formamides in the lack of outer adsorbed molecular layers, as seen from the  $P(U_b)$  distributions in states I–III, we can conclude that their interaction with the subsequent layers in state IV most likely involves strong dipolar interactions, as without that the additional, roughly  $-20$  kJ/mol of the binding energy could not be explained.

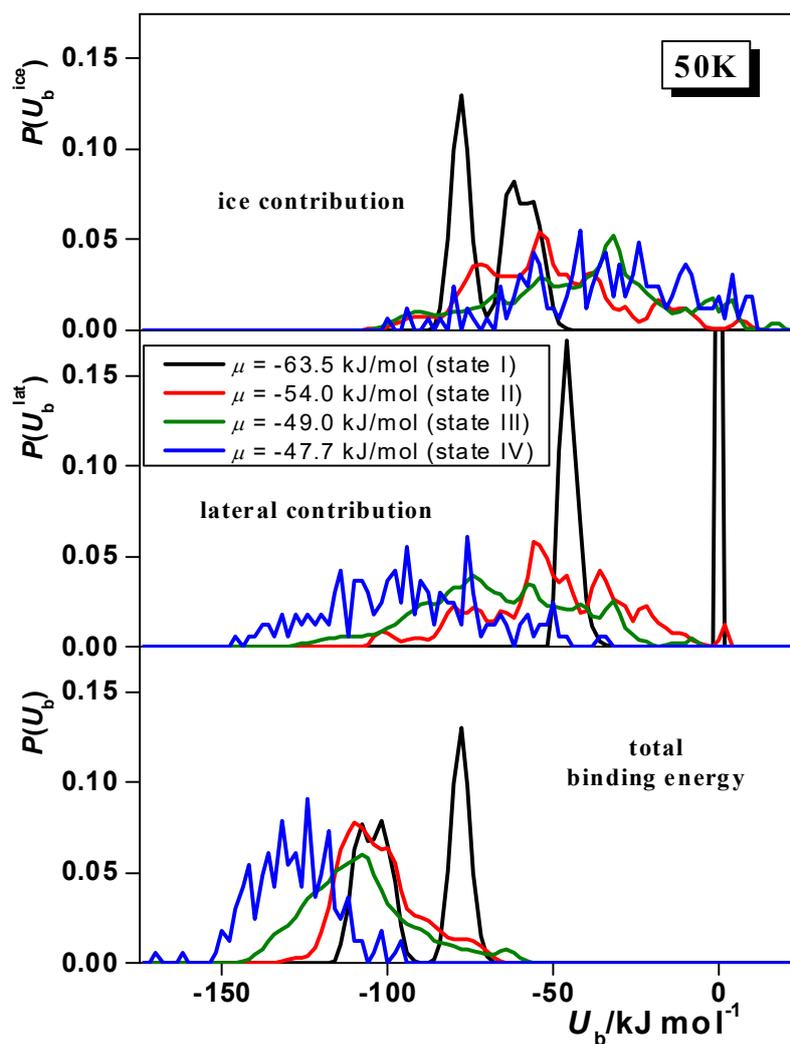


Figure 4.26: Distribution of the total binding energy ( $U_b$ ) of the first layer formamide molecules (bottom panels) as well as its contribution coming from the other formamide molecules in the system (middle panels) and from the ice phase (top panels), as obtained at 50 K at the surface of LDA ices in states I (black), II (red), III (green), and IV (blue).

The distributions obtained at lower temperatures (Figure 4.25 and 4.26) exhibit similar features, although they are affected by considerably larger noise with decreasing temperature due to the decreased mobility of the molecules. The most important difference from the 200 K data is that in state I the peak of the  $P(U_b^{\text{ice}})$  distribution occurs around  $-80$  kJ/mol (rather than  $-100$  kJ/mol) at 100 K, and it is split into two peaks, being at about  $-80$  kJ/mol and  $-60$  kJ/mol at 50 K. Correspondingly, at 100 K the  $P(U_b^{\text{lat}})$  distribution exhibits a much larger peak around  $-25$  kJ/mol than at 200 K, and a rather small peak occurs also close to  $-50$  kJ/mol. Further, at 50 K, the peak around  $-25$  kJ/mol completely vanishes, and that at  $-50$  kJ/mol becomes very large. These findings indicate that at these low temperatures adsorbed formamide molecules are not at all isolated from each other even at very low surface densities (state I corresponds, on average, to 3.5 and 1.5 formamide molecules per surface at 100 and 50 K, respectively; see Tables 3.5 and 3.6); instead, they are very strongly interacting with each other even at such low surface densities. This extremely strong affinity of the formamide molecules for lateral interaction is in a clear accordance with the observed non-Langmuir nature of the adsorption isotherms, as discussed previously.

Finally, the heat of adsorption at infinitely low surface coverage, an experimentally accessible quantity, can, in principle, be well estimated from the mean value of the  $P(U_b^{\text{ice}})$  distribution at low enough surface density, i.e., when the adsorbed molecules are completely isolated from each other. However, in the present case we can calculate its value only in the case of  $I_h$  ice, when, in state I, the adsorbed molecules are indeed isolated from each other (see the middle panel of Figure 4.24). For this system, the heat of adsorption at infinitely low surface density turns out to be  $-105.6$  kJ/mol. Unfortunately, we are not aware of any experiment concerning this quantity yet. However, its measurement in the future could provide a further test of the reliability of the present results. In the case of LDA, on the other hand, we do not have the situation of completely isolated adsorbed molecules even at the lowest surface density considered, and hence the mean value of  $P(U_b^{\text{ice}})$  in state I, which turned out to be  $-99$  kJ/mol,  $-80$  kJ/mol, and  $-68$  kJ/mol at 200, 100, and 50 K, respectively, can only serve as an upper estimate of the heat of adsorption at infinite dilution, which becomes increasingly crude with decreasing temperature. Besides the heat of adsorption at infinite dilution, the isosteric heat of the adsorption  $Q_{\text{st}}$  (i.e., the heat of adsorption corresponding to a given

finite surface coverage) can also be calculated, at least on LDA ice.  $Q_{\text{st}}$  can be calculated through the Clausius–Clapeyron equation as<sup>250</sup>

$$Q_{\text{st}} = R \frac{T_1 T_2}{T_2 - T_1} (\ln p_{\text{rel},1} - \ln p_{\text{rel},2}) \quad (4.12)$$

where the relative pressure values  $p_{\text{rel},1}$  and  $p_{\text{rel},2}$  correspond to the same surface density values along the  $T_1$  and  $T_2$  isotherms, respectively. We have selected points along the simulated isotherms in which the value of  $\Gamma$  agrees within 2% and calculated the isosteric heat of adsorption at these surface coverages. The obtained results are listed in Table 4.2. Any future measurement of the isosteric heat of adsorption can also provide an important test of the validity of the present results.

*Table 4.2: Isosteric Heat of the Adsorption of Formamide on LDA Ice at Various Surface Coverages*

|                                    |                 |                 |                 |                 |                 |                 |
|------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| $\Gamma/\mu\text{mol m}^{-2}$      | $0.54 \pm 0.01$ | $0.64 \pm 0.02$ | $1.46 \pm 0.03$ | $2.23 \pm 0.03$ | $3.48 \pm 0.07$ | $6.93 \pm 0.33$ |
| $Q_{\text{st}}/\text{kJ mol}^{-1}$ | -60.4           | -59.1           | -55.8           | -53.3           | -50.3           | -44.3           |

## 5. Summary

The main aim of this thesis was to investigate interface phenomena in formamide-water systems. Formamide is considered as one of the key precursor molecules in the prebiotic evolution theory. However, the enhancement of formamide concentration is necessary for its contribution to the synthesis of more complex biomolecules. This process could take place at different interfacial systems, such as the surface layer of aqueous formamide solutions, or in the adsorption layer of formamide at the surface of various ice forms. Hence, the vapor-liquid interface of the formamide-water mixtures and the formamide adsorption on crystalline and low-density amorphous ices under tropospheric and interstellar conditions have been studied in detail to test this hypothesis. To achieve these aims, an appropriate force fields combination is needed to represent accurately the experimental properties of the formamide-water systems. To choose the best available model pair, the thermodynamic changes occurring upon mixing formamide and water have been investigated by Monte Carlo simulation. The combination of five well-known formamide (OPLS, OPLS/AA\_mod, CHARMM, Cordeiro and MMPB) and three water models (SPC/E, TIP4P and TIP4P/2005) have been used to calculate the Helmholtz free energy, energy, and entropy of mixing along an appropriately chosen thermodynamic cycle using the method of thermodynamic integration. The performance of the studied model combinations has been evaluated according to the available experimental data concerning the energy of mixing, and also considering the well-known full miscibility of these compounds.

The results have revealed that the Helmholtz free energy of mixing is always negative in the whole composition range, showing that the mixtures are always more stable than the two neat components. The energy and entropy of mixing have exposed almost corresponding values to the ideal term of mixing. The OPLS/AA\_mod model of formamide has a negative energy of mixing in combination with all the three water models considered. This behavior is in contrast to both that of the other formamide models considered and also to the available experimental data.<sup>113,114</sup> The OPLS model of formamide has been found to exhibit miscibility gaps with all the three water models considered, while the model combinations involving the Cordeiro or MMPB formamide

models very closely approach the miscibility limit at a certain compositions, and hence full miscibility in these cases might also depend on the exact conditions under which the simulations are performed. The CHARMM model of formamide has shown full miscibility with the three tested water models. The experimental  $U^{\text{mix}}$  data are somewhat better reproduced when the SPC/E water model is used; therefore, the use of the CHARMM model of formamide together with the SPC/E model of water seems to be the best combination when describing the thermodynamics of mixing of formamide and water. Hence, this combination has been used for further investigation of the interfacial phenomenon.

After the identification of the best available force field combination, the intrinsic liquid surface of formamide-water mixtures of different compositions has been investigated in detail in the canonical  $(N, V, T)$  ensemble at 300 K by molecular dynamics simulations and ITIM analysis. The general picture that has emerged from the various results of this study is that the surface properties of these mixtures are governed by the strong and extensive hydrogen bonding between the surface molecules. Similarly to the bulk liquid phase,<sup>100</sup> water and formamide molecules have turned out to be able to substitute each other even in the two-dimensional hydrogen-bonding network at the surface of their mixtures. The orientational preferences has been found to be driven by the pursuit of the minimization of the free energy, hence the molecules taking positions to maximize the hydrogen bonds they are able to form. Therefore, both molecules prefer parallel alignment with the macroscopic plane of the liquid surface. The liquid surface of the formamide-water system is built up from a strong two-dimensional hydrogen bonding network, which indicates percolating properties at the liquid surface in the entire composition range. Hence, the surface layer is more compact than the subsequent molecular layers, which is in contrast to other aqueous solutions, such as hydrogen cyanide, acetonitrile and acetone.<sup>145,152,231</sup> The dynamical behavior of the surface molecules shows that the water molecules stabilize the stay of the formamide molecules at the liquid surface, while the formamide molecules immobilize the water molecules in the surface layer, preventing them from exhibiting considerable diffusion at the liquid surface during their entire stay there. Similarly to the bulk liquid phase, the formamide and water molecules have been found to be miscible with each other on the molecular level also at the surface of the liquid phase. Hence, no large self-association of the formamide molecules are found at the surface at any composition, only a slightly

adsorption of the formamide molecules has been detected in the surface layer. These findings mean from the prebiotic evolutionary point of view that under the examined conditions the enrichment of formamide is not significant at the surface of its aqueous solution. Therefore, its participation in biomolecular synthesis at the vapor-liquid interface of formamide-water systems is not justified.

Finally, the adsorption of formamide at the surface of  $I_h$  and LDA ices under both interstellar (in the temperature range of 50–200 K) and tropospheric (at 200 K) conditions have been investigated by grand canonical Monte Carlo simulations. The obtained adsorption isotherms have revealed multilayer adsorption of formamide and the presence of a stable, yet unsaturated monolayer at the beginning of the formation of the outer adsorption layers. The isotherms have turned out to be between types II and III according to the IUPAC convention. This finding indicates, in accordance with earlier results and what has been observed at the vapor-liquid interface,<sup>110,111,236</sup> that the adsorbate-adsorbent and lateral interactions are of roughly the same magnitude. Thus, adsorbed formamide molecules are not isolated from each other even at very low surface densities, and this effect is more marked at lower temperatures. As a consequence, the isotherms exhibit non-Langmuir shape even in the pressure range that precedes multilayer adsorption. The calculated surface density suggests strong lateral association of the formamide molecules in the case of all investigated conditions. The analysis of the first adsorbed molecular layer has revealed a similar behavior of the formamide molecules to that at the vapor-liquid interface. Formamide molecules strive to maximize the hydrogen-bonds they form with their neighbors. Hence, to achieve the maximum possible four hydrogen bonds to be formed with both surface waters and other formamides, the adsorbed formamide molecules in the first molecular layer prefer parallel and tilted alignments relative to the ice surface. The ratio of the water and formamide molecules among the established hydrogen bonded neighbors depends on the temperature, ice type, and surface coverage. We have found that the adsorption of formamide on LDA ice in the temperature range characteristic to the interstellar medium and also on  $I_h$  ice at the tropospheric temperature of 200 K is a strongly preferred process, which certainly has to be taken into account in studying chemical reactivity in these environments.

These results support the possible role of formamide in the prebiotic evolution theory by suggesting the surface of ice grains in the Earth's atmosphere as a possible location of formamide enrichment. The atmosphere of the early and modern Earth could

have thus a possible place of the synthesis of complex biomolecules. Furthermore, due to the adsorption of formamide on the LDA ice surface in the interstellar medium, the synthesis of new complex biomolecules could also occur outside the Earth if initiated by high energy particles (e.g., heat, electromagnetic radiation, etc.). Thus, these observations fit both into the prebiotic and the panspermia hypotheses. Conclusions in this work presented the possible locations of biomolecule synthesis due to the thick adsorption layer of formamide. However, the "how" is still not fully clarified. Studies in both experimental and computational chemistry have already been published exploring possible reaction pathways in the synthesis with the help of formamide.<sup>10,15,16,251,252</sup> Based on the identified locations, the range of possible interactions and catalysts can be further narrowed. The results of this thesis also show the importance of hydrogen bonds in the formation of surface phenomena. These secondary bonds proved to be determining in terms of energy and orientation properties of the surface molecules. Hence emphasizes the importance of hydrogen bonds in the formation of life. This raises further questions about what other effects they provide and what role can hydrogen bonds play in the synthesis of biomolecules. Furthermore, hydrogen bonds could promote the formation of asymmetric structures in life-essential macromolecules.

## Thesis Points

1. I compared the possible combinations of five formamide and three water models by Monte Carlo simulation and thermodynamic integration. Based on the miscibility of the components and the thermodynamic properties, such as the Helmholtz free energy, energy and entropy of their mixing, the most realistic representation of the formamide-water system is found to be the pair of the CHARMM formamide and SPC/E water models.
2. I found the existence of an extensive, percolating hydrogen bonding network at the liquid-vapor interface of formamide-water mixtures, in which the components are interchangeable with each other and the molecules show full miscibility on the molecular level. No strong adsorption or lateral self-association of any of the two components have been detected at the surface layer. Furthermore, formamide molecules have been found to prevent water from considerable diffusion at the liquid surface, while waters have been found to stabilize the stay of the formamide molecules at the surface of their liquid mixtures.
3. I showed that formamide and water molecules prefer parallel alignment with the macroscopic surface plane at the surface of their liquid mixtures. These preferences can be interpreted as the intention to minimize their free energy, hence they are oriented in such a way that the number of hydrogen-bonds they form is maximized.
4. I found that the adsorption of formamide on the surface of crystalline and low-density amorphous ices at atmospheric and interstellar conditions is a strongly favorable phenomenon, which thus leads to a strong enhancement of the formamide concentration. Adsorbed formamide molecules are not isolated from each other on such icy surfaces, in addition, multilayer adsorption begins before the first layer becomes saturated. Based on these findings, the adsorbate-adsorbent and lateral interactions turn out to be of rather similar strengths.
5. I determined the orientational preferences of the adsorbed formamide molecules in the first molecular layer at the surface of crystalline and low-density amorphous ices, which show parallel and tilted alignments relative to the ice surface. Similar to the vapor-liquid interface, formamide molecules try to minimize their free energy in such a way that they form the maximum possible hydrogen-bonds with their neighbors.

## **Publications Included in The Thesis**

1. Miscibility and Thermodynamics of Mixing of Different Models of Formamide and Water in Computer Simulation; B. Kiss, B. Fábián, A. Idrissi, M. Szőri and P. Jedlovszky; *J. Phys. Chem. B* 2017, 121, 7147–7155; DOI: 10.1021/acs.jpcc.7b04965; Erratum: DOI: 10.1021/acs.jpcc.7b09210
2. Investigation of the Liquid–Vapor Interface of Water–Formamide Mixtures by Computer Simulation and Intrinsic Surface Analysis; B. Kiss, B. Fábián, A. Idrissi, M. Szőri and P. Jedlovszky; *J. Phys. Chem. C* 2018, 122, 19639–19651; DOI: 10.1021/acs.jpcc.8b05874
3. Adsorption of Formamide at the Surface of Amorphous and Crystalline Ices under Interstellar and Tropospheric Conditions. A Grand Canonical Monte Carlo Simulation Study; B. Kiss, S. Picaud, M. Szőri and P. Jedlovszky; *J. Phys. Chem. A* 2019, 123, 2935–2948; DOI: 10.1021/acs.jpca.9b00850

## **Publications Not Included in The Thesis**

1. Bioimprinted lipases in PVA nanofibers as efficient immobilized biocatalysts; D. Weiser, P. L. Sóti, G. Bánóczy, V. Bódai, B. Kiss, Á. Gellért, Z. K. Nagy, B. Koczka, A. Szilágyi, G. Marosi, L. Poppe; *Tetrahedron* 72 (2016) 7335-7342; DOI: 10.1016/j.tet.2016.06.027
2. Adsorption of Propylene Oxide on Amorphous Ice under Interstellar Conditions. A Grand Canonical Monte Carlo Simulation Study; B. Kiss, M. Szőri and P. Jedlovszky; *J. Phys. Chem. C* 2020, 124, 30, 16402–16414; DOI: 10.1021/acs.jpcc.0c03722

## Short Summary of the Thesis

Formamide and its aqueous mixtures might have played a key role in the prebiotic evolution and the panspermia theory. To reach an effective formamide enrichment which is necessary for the biomolecule synthesis, accumulation of formamide molecules has occurred. This can be explained with the sometimes extreme and complex environments including interfaces. The aim of this doctoral dissertation is to interpret the phenomena occurring at interfaces of formamide -water binary systems using computer simulation methods at ambient and cold condition. To achieve this purpose, the thermodynamic changes that occur upon mixing five models of formamide and three models of water, including the miscibility of these model combinations itself, is studied by performing Monte Carlo computer simulations using an appropriately chosen thermodynamic cycle and the method of thermodynamic integration. The Helmholtz free energy, energy, and entropy of mixing of these model combinations in the entire composition range have been calculated. Concerning both the miscibility and the energy of mixing of these model combinations, the combination of the CHARMM formamide and SPC/E water models have been selected for the further simulations of formamide –water systems.

Molecular dynamics simulations of the liquid–vapor interface of formamide –water mixtures of different compositions, have been performed in the canonical ( $N, V, T$ ) ensemble at 300 K. The layer of the surface molecules has been identified and analyzed in detail in terms of the intrinsic identification of the truly interfacial molecules (ITIM) method. The obtained results reveal the surface molecules form a percolating lateral hydrogen-bonding network in every case. However, the molecules mix with each other even on the molecular scale. Neither strong adsorption, nor lateral self-association of any of the components in the surface layer has been observed, although formamide exhibits a slight ability of being accumulated in the surface layer. The orientational preferences of the surface molecules are also governed by the requirement of maximizing their hydrogen bonds. The main preference of both molecules is found to be the parallel alignment with the macroscopic plane of the liquid surface. The dynamics of the surface molecules is also dominated by the hydrogen bonding.

Furthermore, the adsorption of formamide is studied both at the surface of crystalline ( $I_h$ ) ice at 200 K and at the surface of low density amorphous (LDA) ice in the temperature range of 50–200 K by grand canonical Monte Carlo (GCMC) simulation. These systems are characteristic of the upper troposphere and of the interstellar medium

(ISM), respectively. The results reveal that the adsorption of formamide at the ice surface under these conditions is a very strongly preferred process, which has to be taken into account when studying the chemical reactivity in these environments. The adsorption is found to lead to the formation of multimolecular adsorption layer, the occurrence of which somewhat precedes the saturation of the first molecular layer. Adsorption is found to be slightly stronger on LDA than  $I_h$  ice under identical thermodynamic conditions, due to the larger surface area exposed to the adsorption. The first layer formamide molecules are very strongly bound to the ice surface, forming typically four hydrogen bonds with each other and the surface water molecules.

The results of this thesis support the role of formamide in the prebiotic evolution and the panspermia theory by suggesting the surface of ice grains in the Earth's atmosphere and in the interstellar medium as a possible location of formamide enrichment. Furthermore, they emphasize the importance of hydrogen bonds in the formation of life.

## Bref Résumé de la Thèse

Le formamide et ses mélanges aqueux pourraient avoir joué un rôle clé dans l'évolution des prébiotiques et la théorie de la panspermie. Pour parvenir à un enrichissement efficace du formamide, nécessaire à la synthèse des biomolécules, une accumulation de molécules de formamide s'est produite. Cela peut s'expliquer par les conditions extrêmes et complexes (y compris les interfaces) de son environnement environnements. L'objectif de cette thèse de doctorat est d'interpréter les phénomènes qui se produisent aux interfaces des systèmes binaires formamide-eau en utilisant des méthodes de modélisation moléculaire. Pour atteindre cet objectif, nous avons utilisé la méthode Monte Carlo pour analyser propriétés thermodynamiques de ce mélange. Ces propriétés telles que L'énergie libre de Helmholtz, l'énergie et l'entropie ont été calculées suivant un cycle thermodynamique choisi de manière appropriée et la méthode d'intégration thermodynamique. L'évolution de ces propriétés en fonction de la composition du mélange, ont été calculées en combinant cinq modèles de formamide à trois modèles d'eau. Nos résultats montrent que la combinaison du modèle de potentiel d'interaction CHARMM formamide et du modèle SPC/E de l'eau reproduisent correctement les propriétés thermodynamiques expérimentales de ce mélange et par conséquent nous les avons choisis pour analyser les propriétés d'interface de ce mélange.

Des simulations de dynamique moléculaire de l'interface liquide-vapeur de mélanges formamide-eau à différentes compositions, ont été réalisées dans l'ensemble canonique ( $N, V, T$ ) à 300 K. Les molécules à l'interface ont été identifiées par un algorithme d'identification intrinsèque des molécules véritablement à l'interface. Les résultats obtenus révèlent que les molécules de surface forment dans tous les cas un réseau de liaison hydrogène latéral percolant. Cependant, les molécules à l'interface se mélangent entre elles à l'échelle moléculaire. Ni une forte adsorption, ni une auto-association latérale de l'un des composants à l'interface n'ont été observées, bien que le formamide présente une légère capacité d'accumulation à l'interface. Les préférences d'orientation des molécules de surface sont également régies par l'exigence de maximiser leurs liaisons hydrogène. La dynamique des molécules à l'interface est également dominée par la liaison hydrogène.

En outre, Nous avons étudié par simulation de Monte Carlo canonique (GCMC), l'adsorption du formamide est étudiée à l'interface de la glace cristalline ( $I_h$ ) à 200 K et à celle de la glace amorphe de faible densité (LDA) dans la gamme de température de 50-

200 K. Ces systèmes sont caractéristiques de la haute troposphère et du milieu interstellaire (ISM), respectivement. Les résultats révèlent que l'adsorption du formamide à la surface de la glace dans ces conditions est un processus très fortement privilégié, qui doit être pris en compte lors de l'étude de la réactivité chimique dans ces environnements. Il s'avère que l'adsorption conduit à la formation d'une couche d'adsorption multimoléculaire, dont l'apparition précède quelque peu la saturation de l'interface. Dans des conditions thermodynamiques identiques, l'adsorption est légèrement plus forte sur la glace LDA que sur la glace I<sub>h</sub>, en raison de la plus grande surface exposée à l'adsorption. Les molécules de formamide de la première couche sont très fortement liées à la surface de la glace, formant typiquement quatre liaisons hydrogène entre elles et avec les molécules d'eau de surface.

Les résultats de cette thèse soutiennent le rôle du formamide dans l'évolution prébiotique et la théorie de la panspermie en suggérant que la surface, de particules de glaces dans l'atmosphère terrestre et dans le milieu interstellaire, comme un lieu possible d'enrichissement en formamide. En outre, ils soulignent l'importance des liaisons hydrogène dans la formation de la vie.

## References

1. Oparin, A. I. *The Origin of Life*. (Macmillan, 1938).
2. Urey, H. C. On the Early Chemical History of the Earth and the Origin of Life. *Proc. Natl. Acad. Sci.* **38**, 351–363 (1952).
3. Šponer, J. E. *et al.* Emergence of the First Catalytic Oligonucleotides in a Formamide-Based Origin Scenario. *Chem. - A Eur. J.* **22**, 3572–3586 (2016).
4. Menor-Salván, C. & Marín-Yaseli, M. R. Prebiotic chemistry in eutectic solutions at the water–ice matrix. *Chem. Soc. Rev.* **41**, 5404 (2012).
5. Szori, M. *et al.* Chemical evolution of biomolecule building blocks. Can thermodynamics explain the accumulation of glycine in the prebiotic ocean? *Phys. Chem. Chem. Phys.* **13**, 7449–7458 (2011).
6. McCollom, T. M. Miller-Urey and Beyond: What Have We Learned About Prebiotic Organic Synthesis Reactions in the Past 60 Years? *Annu. Rev. Earth Planet. Sci.* **41**, 207–229 (2013).
7. Cleaves, H. J., Chalmers, J. H., Lazcano, A., Miller, S. L. & Bada, J. L. A reassessment of prebiotic organic synthesis in neutral planetary atmospheres. *Orig. Life Evol. Biosph.* **38**, 105–115 (2008).
8. Saitta, A. M. & Saija, F. Miller experiments in atomistic computer simulations. *Proc. Natl. Acad. Sci.* **111**, 13768–13773 (2014).
9. Saladino, R., Botta, G., Delfino, M. & Di Mauro, E. Meteorites as catalysts for prebiotic chemistry. *Chem. - A Eur. J.* **19**, 16916–16922 (2013).
10. Ferus, M. *et al.* High-energy chemistry of formamide: A unified mechanism of nucleobase formation. *Proc. Natl. Acad. Sci.* **112**, 657–662 (2015).
11. Sutherland, J. D. The Origin of Life - Out of the Blue. *Angewandte Chemie - International Edition* **55**, 104–121 (2016).
12. Šponer, J. E. *et al.* Prebiotic synthesis of nucleic acids and their building blocks at the atomic level-merging models and mechanisms from advanced computations and experiments. *Phys. Chem. Chem. Phys.* **18**, 20047–20066 (2016).
13. Saladino, R. *et al.* First Evidence on the Role of Heavy Ion Irradiation of Meteorites and Formamide in the Origin of Biomolecules. *Orig. Life Evol. Biosph.* **46**, 515–521 (2016).
14. Saladino, R. *et al.* Meteorite-catalyzed syntheses of nucleosides and of other prebiotic compounds from formamide under proton irradiation. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E2746–E2755 (2015).
15. Ferus, M. *et al.* On the road from formamide ices to nucleobases: IR-spectroscopic observation of a direct reaction between cyano radicals and formamide in a high-energy impact event. *J. Am. Chem. Soc.* **134**, 20788–20796 (2012).

16. Ferus, M. *et al.* Formation of nucleobases in a Miller-Urey reducing atmosphere. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4306–4311 (2017).
17. Miller, S. L. A production of amino acids under possible primitive earth conditions. *Science (80-. )*. **117**, 528–529 (1953).
18. López-Sepulcre, A. *et al.* Interstellar Formamide (NH<sub>2</sub>CHO), a Key Prebiotic Precursor. *ACS Earth Sp. Chem.* **3**, 2122–2137 (2019).
19. Matthews, C. N. The HCN World. in *Origins* 121–135 (Kluwer Academic Publishers, 2006). doi:10.1007/1-4020-2522-x\_8
20. Matthews, C. N. & Moser, R. E. Peptide synthesis from hydrogen cyanide and water. *Nature* **215**, 1230–1234 (1967).
21. Arrhenius, S. & Borns, H. Worlds in the Making. The Evolution of the Universe. *Bull. Am. Geogr. Soc.* **41**, 123 (1909).
22. Oró, J. Comets and the formation of biochemical compounds on the primitive earth. *Nature* **190**, 389–390 (1961).
23. Oró, J., Mills, T. & Lazcano, A. Comets and the formation of biochemical compounds on the primitive Earth - A review. *Orig. Life Evol. Biosph.* **21**, 267–277 (1991).
24. Hoyle, F.; Wickramasinghe, C. *Lifecloud – The Origin of Life in the Universe*. (J.M. Dent and Sons, 1978).
25. Irvine, W. M. *et al.* The HNC/HCN ratio in comets. *Earth, Moon Planets* **78**, 29–35 (1997).
26. Chyba, C. F., Thomas, P. J., Brookshaw, L. & Sagan, C. Cometary delivery of organic molecules to the early earth. *Science (80-. )*. **249**, 366–373 (1990).
27. Brack, A. Life in the solar system. *Adv. Sp. Res.* **24**, 417–433 (1999).
28. Rubin, R. H. *et al.* Microwave Detection of Interstellar Formamide. *Astrophys. J.* **169**, L39 (1971).
29. Turner, B. E. A molecular line survey of Sagittarius B2 and Orion-KL from 70 to 115 GHz. I - The observational data. *Astrophys. J. Suppl. Ser.* **70**, 539 (1989).
30. Gibb, E., Nummelin, A., Irvine, W. M., Whittet, D. C. B. & Bergman, P. Chemistry of the Organic-rich Hot Core G327.3–0.6. *Astrophys. J.* **545**, 309–326 (2000).
31. Bisschop, S. E., Jørgensen, J. K., Van Dishoeck, E. F. & De Wachter, E. B. M. Testing grain-surface chemistry in massive hot-core regions. *Astron. Astrophys.* **465**, 913–929 (2007).
32. Feng, S. *et al.* Inferring the evolutionary stages of the internal structures of NGC 7538 S and IRS1 from chemistry. *Astron. Astrophys.* **593**, (2016).
33. Isokoski, K., Bottinelli, S. & Van Dishoeck, E. F. Chemistry of massive young stellar objects with a disk-like structure. *Astron. Astrophys.* **554**, (2013).
34. Suzuki, T. *et al.* The Difference in Abundance between N-bearing and O-bearing Species in High-mass Star-forming Regions. *Astrophys. J. Suppl. Ser.* **237**, 3 (2018).
35. Xu, J. L. & Wang, J. J. Kinematics and chemistry of the hot core in G20.08-0.14N.

- Mon. Not. R. Astron. Soc.* **431**, 2385–2396 (2013).
36. Allen, V., Van Der Tak, F. F. S., Sánchez-Monge, A., Cesaroni, R. & Beltrán, M. T. Chemical segregation in hot cores with disk candidates: An investigation with ALMA. *Astron. Astrophys.* **603**, (2017).
  37. Sakai, T. *et al.* ALMA Observations of the IRDC Clump G34.43+00.24 MM3: Complex Organic and Deuterated Molecules. *Astrophys. J.* **857**, 35 (2018).
  38. López-Sepulcre, A. *et al.* Shedding light on the formation of the pre-biotic molecule formamide with ASAI. *Mon. Not. R. Astron. Soc.* **449**, 2438–2458 (2015).
  39. Marcelino, N. *et al.* ALMA observations of the young protostellar system Barnard 1b: Signatures of an incipient hot corino in B1b-S. *Astron. Astrophys.* **620**, (2018).
  40. Kahane, C., Ceccarelli, C., Faure, A. & Caux, E. Detection of formamide, the simplest but crucial amide, in a solar-type protostar. *Astrophys. J. Lett.* **763**, (2013).
  41. Taquet, V. *et al.* Constraining the Abundances of Complex Organics in the Inner Regions of Solar-Type Protostars. *Astrophys. J.* **804**, (2015).
  42. Imai, M. *et al.* Discovery of a Hot Corino in the Bok Globule B335. *Astrophys. J.* **830**, L37 (2016).
  43. Oya, Y. *et al.* L483: Warm Carbon-chain Chemistry Source Harboring Hot Corino Activity. *Astrophys. J.* **837**, 174 (2017).
  44. Lee, C.-F. *et al.* Formation and Atmosphere of Complex Organic Molecules of the HH 212 Protostellar Disk. *Astrophys. J.* **843**, 27 (2017).
  45. Yamaguchi, T. *et al.* The 3 mm spectral line survey toward the lynds 1157B1 shocked region. I. Data. *Publications of the Astronomical Society of Japan* **64**, (2012).
  46. Mendoza, E. *et al.* Molecules with a peptide link in protostellar shocks: A comprehensive study of L1157. *Mon. Not. R. Astron. Soc.* **445**, 151–161 (2014).
  47. Ceccarelli, C. *et al.* Seeds Of Life In Space (SOLIS): The Organic Composition Diversity at 300–1000 au Scale in Solar-type Star-forming Regions. *Astrophys. J.* **850**, 176 (2017).
  48. Bockelée-Morvan, D. *et al.* New molecules found in comet C/1995 O1 (Hale-Bopp): Investigating the link between cometary and interstellar material. *Astron. Astrophys.* **353**, 1101–1114 (2000).
  49. Biver, N. *et al.* Complex organic molecules in comets C/2012 F6 (Lemmon) and C/2013 R1 (Lovejoy): Detection of ethylene glycol and formamide. *Astron. Astrophys.* **566**, (2014).
  50. Goesmann, F. *et al.* Organic compounds on comet 67P/Churyumov-Gerasimenko revealed by COSAC mass spectrometry. *Science (80-. )*. **349**, (2015).
  51. Muller, S. *et al.* A precise and accurate determination of the cosmic microwave background temperature at  $z = 0.89$ . *Astron. Astrophys.* **551**, (2013).
  52. Corby, J. F. *et al.* An ATCA survey of Sagittarius B2 at 7 mm: Chemical complexity

- meets broad-band interferometry. *Mon. Not. R. Astron. Soc.* **452**, 3969–3993 (2015).
53. Eyres, S. P. S. *et al.* ALMA reveals the aftermath of a white dwarf–brown dwarf merger in CK Vulpeculae. *Mon. Not. R. Astron. Soc.* **481**, 4931–4939 (2018).
  54. Gottlieb, C. A., Palmer, P., Rickard, L. J. & Zuckerman, B. Studies of Interstellar Formamide. *Astrophys. J.* **182**, 699 (1973).
  55. Adande, G. R., Woolf, N. J. & Ziurys, L. M. Observations of Interstellar Formamide: Availability of a Prebiotic Precursor in the Galactic Habitable Zone. *Astrobiology* **13**, 439–453 (2013).
  56. Schutte, W. A. *et al.* Weak ice absorption features at 7.24 and 7.41  $\mu\text{m}$  in the spectrum of the obscured young stellar object W 33A. *Astron. Astrophys.* **343**, 966–976 (1999).
  57. Solomon, P. M. Interstellar molecules. *Phys. Today* **26**, 32–40 (1973).
  58. Despois, D., Crovisier, J., Bockele-Morvan, D. & Biver, N. *Proceedings of the Second European Workshop on: Exo-Astrobiology*. (ESA Publications Division, 2002).
  59. Lis, D. C. *et al.* New molecular species in comet C/1995 O1 (Hale-Bopp) observed with the Caltech Submillimeter Observatory. in *Earth, Moon and Planets* **78**, 13–20 (1997).
  60. Hudson, R. L. & Moore, M. H. Reactions of nitriles in ices relevant to Titan, comets, and the interstellar medium: formation of cyanate ion, ketenimines, and isonitriles. *Icarus* **172**, 466–478 (2004).
  61. Rimola, A. *et al.* Can Formamide Be Formed on Interstellar Ice? An Atomistic Perspective. *ACS Earth Sp. Chem.* **2**, 720–734 (2018).
  62. Gao, Y., Qin, Z., Guan, L., Wang, X. & Chen, G. Z. Organoaqueous calcium chloride electrolytes for capacitive charge storage in carbon nanotubes at sub-zero-temperatures. *Chem. Commun.* **51**, 10819–10822 (2015).
  63. Carota, E., Botta, G., Rotelli, L., Di Mauro, E. & Saladino, R. Current Advances in Prebiotic Chemistry Under Space Conditions. *Curr. Org. Chem.* **19**, 1963–1979 (2015).
  64. Saladino, R., Botta, G., Pino, S., Costanzo, G. & Di Mauro, E. Genetics first or metabolism first? The formamide clue. *Chem. Soc. Rev.* **41**, 5526–5565 (2012).
  65. Botta, L. *et al.* Formamide-based prebiotic chemistry in the Phlegrean Fields. *Adv. Sp. Res.* **62**, 2372–2379 (2018).
  66. Mispelaer, F. *et al.* Diffusion measurements of CO, HNCO, H<sub>2</sub>CO, and NH<sub>3</sub> in amorphous water ice. *Astron. Astrophys.* **555**, A13 (2013).
  67. Cottin, H., Gazeau, M. C. & Raulin, F. Cometary organic chemistry: A review from observations, numerical and experimental simulations. *Planet. Space Sci.* **47**, 1141–1162 (1999).
  68. Wakelam, V., Loison, J. C., Mereau, R. & Ruaud, M. Binding energies: New values and impact on the efficiency of chemical desorption. *Mol. Astrophys.* **6**, 22–35

- (2017).
69. Chaabouni, H., Diana, S., Nguyen, T. & Dulieu, F. Thermal desorption of formamide and methylamine from graphite and amorphous water ice surfaces. *Astron. Astrophys.* **612**, A47 (2018).
  70. Vichiatti, R. M., Spada, R. F. K., da Silva, A. B. F., Machado, F. B. C. & Haiduke, R. L. A. Could HCN Be Responsible for the Formamide Synthesis in Earth's Primitive Atmosphere? *Astrophys. J. Suppl. Ser.* **245**, 11 (2019).
  71. Ferus, M. *et al.* High-energy chemistry of formamide: A simpler way for nucleobase formation. *J. Phys. Chem. A* **118**, 719–736 (2014).
  72. Ferus, M. *et al.* HNCO-based synthesis of formamide in planetary atmospheres. *Astron. Astrophys.* **616**, A150 (2018).
  73. Adam, Z. R. *et al.* Estimating the capacity for production of formamide by radioactive minerals on the prebiotic Earth. *Sci. Rep.* **8**, 265 (2018).
  74. Bada, J. L., Chalmers, J. H. & Cleaves, H. J. Is formamide a geochemically plausible prebiotic solvent? *Phys. Chem. Chem. Phys.* **18**, 20085–20090 (2016).
  75. Stillinger, F. H. Water revisited. *Science (80- )*. **209**, 451–457 (1980).
  76. Malenkov, G. G. Structure and dynamics of liquid water. *J. Struct. Chem.* **47**, S1–S31 (2006).
  77. Fernández, D. P., Mulev, Y., Goodwin, A. R. H. & Sengers, J. M. H. L. A Database for the Static Dielectric Constant of Water and Steam. *J. Phys. Chem. Ref. Data* **24**, 33–70 (1995).
  78. Höhn, A. Formamide. in *Kirk-Othmer Encyclopedia of Chemical Technology* (John Wiley & Sons, Inc., 2000). doi:10.1002/0471238961.0615181308150814.a01
  79. Marechal, Y. Infrared-Spectra of a Poorly Known Species - Water .3. *J. Phys. Chem.* **97**, 2846–2850 (1993).
  80. Tolosa, S. Hidalgo, A., Sansón, J. A. Thermodynamic, Structural, and Dynamic Study of the N H...O=C Hydrogen Bond Association in Aqueous Solution. *Chem. Phys.* **255**, 73–84 (2000).
  81. Kritsana, P.; Reinhart, A. Test Particle Model Potential for Formamide and Molecular Dynamics Simulations of the Liquid. *J. Chem. Phys.* **86**, 5117–5145 (1987).
  82. Thompson, H. B.; LaPlanche, L. Electric Moments of N,N-Disubstituted Amides. *J. Phys. Chem.* **67**, 2230–2231 (1963).
  83. Puhovski, Y. P. & Rode, B. M. Structure and dynamics of liquid formamide. *Chem. Phys.* **190**, 61–82 (1995).
  84. Bakó, I. *et al.* Hydrogen bonded network properties in liquid formamide. *J. Chem. Phys.* **132**, 014506 (2010).
  85. DeSando, R. & Brown, G. An X-Ray Study of Formamide and Solutions of Potassium Iodide in Formamide. *J. Phys. Chem.* **72**, 1088–1091 (1968).
  86. Siegbarn, H. *et al.* ESCA applied to liquids. II. Valence and core electron spectra of

- formamide. *J. Electron Spectros. Relat. Phenomena* **5**, 1059–1079 (1974).
87. Faurkov Nielsen, O., Lund, P. -A. & Praestgaard, E. Hydrogen bonding in liquid formamide. A low frequency Raman study. *J. Chem. Phys.* **77**, 3878–3883 (1982).
  88. Kálmán, E. *et al.* The Molecular Structure and Hydrogen Bond Geometry in Liquid Formamide: Electron, Neutron, and X-Ray Diffraction Studies. *Zeitschrift fur Naturforsch. - Sect. A J. Phys. Sci.* **38**, 231–236 (1983).
  89. Ohtaki, H., Funaki, A., Rode, B. M. & Reibnegger, G. J. The Structure of Liquid Formamide Studied by Means of X-Ray Diffraction and *ab Initio* LCGO-MO-SCF Calculations. *Bull. Chem. Soc. Jpn.* **56**, 2116–2121 (1983).
  90. Radnai, T., Itoh, S. & Ohtaki, H. Liquid Structure of N , N -Dimethylformamide, Acetonitrile, and Their 1 : 1 Molar Mixture. *Bull. Chem. Soc. Jpn.* **61**, 3845–3852 (1988).
  91. Miyake, M., Kaji, O., Nakagawa, N. & Suzuki, T. Structural analysis of liquid formamide. *J. Chem. Soc. Faraday Trans. 2 Mol. Chem. Phys.* **81**, 277–281 (1985).
  92. Bellissent-Funel, M.-C., Nasr, S. & Bosio, L. X-ray and neutron scattering studies of the temperature and pressure dependence of the structure of liquid formamide. *J. Chem. Phys.* **106**, 7913–7919 (1997).
  93. Wiesmann, F.-J., Zeidler, M. D., Bertagnolli, H. & Chieux, P. A neutron diffraction study of liquid formamide. *Mol. Phys.* **57**, 275–285 (1986).
  94. Essex, J. W. & Jorgensen, W. L. Dielectric constants of formamide and dimethylformamide via computer simulation. *J. Phys. Chem.* **99**, 17956–17962 (1995).
  95. Puhovski, Y. P., Safonova, L. P. & Rode, B. M. Molecular dynamic simulations of a liquid formamide and N,N-dimethylformamide with new quantum mechanical potential. *J. Mol. Liq.* **103–104**, 15–31 (2003).
  96. Tsuchida, E. *Ab initio* molecular-dynamics study of liquid formamide. *J. Chem. Phys.* **121**, 4740–4746 (2004).
  97. Radnai, T. *et al.* Structure and dynamics of liquid formamide at high pressure and high temperature: comparison of X-ray diffraction and molecular dynamics results. *J. Mol. Liq.* **110**, 123–132 (2004).
  98. Cordeiro, J. M. M. C — H ... O and N — H ...O Hydrogen Bonds in Liquid Amides Investigated by Monte Carlo Simulation. *Int. J. Quantum Chem.* **66**, 709–727 (1997).
  99. Richardi, J., Krienke, H. & Fries, P. H. Dielectric constants of liquid formamide, N-methylformamide and dimethylformamide via molecular Ornstein-Zernike theory. *Chem. Phys. Lett.* **273**, 115–121 (1997).
  100. Elola, M. D. & Ladanyi, B. M. Computational study of structural and dynamical properties of formamide-water mixtures. *J. Chem. Phys.* **125**, 184506 (2006).
  101. Barthel, J., Buchner, R. & Wurm, B. The dynamics of liquid formamide, N-methylformamide, N,N-dimethylformamide, and N,N-dimethylacetamide. A dielectric relaxation study. *J. Mol. Liq.* **98–99**, 51–69 (2002).

102. Jadżyn, J. & Świergiel, J. On similarity of hydrogen-bonded networks in liquid formamide and water as revealed in the static dielectric studies. *Phys. Chem. Chem. Phys.* **14**, 3170 (2012).
103. Lima, M., Chelli, R., Volkov, V. V. & Righini, R. Two-dimensional infrared spectroscopy of a structured liquid: Neat formamide. *J. Chem. Phys.* **130**, 204518 (2009).
104. Ludwig, R., Weinhold, F. & Farrar, T. C. Experimental and theoretical studies of hydrogen bonding in neat, liquid formamide. *J. Chem. Phys.* **102**, 5118–5125 (1995).
105. Macchiagodena, M. *et al.* Accurate prediction of bulk properties in hydrogen bonded liquids: amides as case studies. *Phys. Chem. Chem. Phys.* **18**, 25342–25354 (2016).
106. Jorgensen, W. L. & Swenson, C. J. Optimized intermolecular potential functions for amides and peptides. Structure and properties of liquid amides. *J. Am. Chem. Soc.* **107**, 569–578 (1985).
107. Wohlfarth, C. Static Dielectric Constants of Pure Liquids and Binary Liquid Mixtures. *Landolt-Barnstein Numer. Data Funct. Relationships Sci. Technol.* (1991). doi:10.1007/978-3-540-75506-7
108. Geiger, A., Stillinger, F. H. & Rahman, A. Aspects of the percolation process for hydrogen-bond networks in water. *J. Chem. Phys.* **70**, 4185–4193 (1979).
109. Stanley, H. E. & Teixeira, J. Interpretation of the unusual behavior of H<sub>2</sub>O and D<sub>2</sub>O at low temperatures: Tests of a percolation model. *J. Chem. Phys.* **73**, 3404 (1980).
110. Elola, M. D. & Ladanyi, B. M. Intermolecular polarizability dynamics of aqueous formamide liquid mixtures studied by molecular dynamics simulations. *J. Chem. Phys.* **126**, 084504 (2007).
111. Puhovski, Y. P. & Rode, B. M. Molecular dynamics simulations of aqueous formamide solution. II. Dynamics of solvent molecules. *J. Chem. Phys.* **102**, 2920–2927 (1995).
112. Coutant, C., Properties, P., David E. Mears<sup>1</sup>, A. D. E., Fumagalli, C. & Fruscella, W. Kirk-Othmer Encyclopedia of Chemical Technology. *Kirk-Othmer Encycl. Chem. Technol.* **3**, 596–624 (2000).
113. Zaichikov, A. M.; Golubinskii, O. E. Enthalpy of Mixing of Water with Some Primer and Secunder Amides. *Zh. Fiz. Khim* **70**, 1175–1179 (1996).
114. Zaichikov, A. M. Thermochemical study of the ternary system water - formamide - dimethylacetamide. *Russ. J. Gen. Chem.* **71**, 162–167 (2001).
115. Zoranić, L., Mazighi, R., Sokolić, F. & Perera, A. Concentration fluctuations and microheterogeneity in aqueous amide mixtures. *J. Chem. Phys.* **130**, 124315 (2009).
116. De La Luz, A. P., Méndez-Maldonado, G. A., Núñez-Rojas, E., Bresme, F. & Alejandre, J. A New Force Field of Formamide and the Effect of the Dielectric Constant on Miscibility. *J. Chem. Theory Comput.* **11**, 2792–2800 (2015).

117. Puhovski, Y. P. & Rode, B. M. Molecular Dynamics Simulations of Aqueous Formamide Solution. 1. Structure of Binary Mixtures. *J. Phys. Chem.* **99**, 1566–1576 (1995).
118. Cordeiro, M. A. M., Santana, W. P., Cusinato, R. & Cordeiro, J. M. M. Monte carlo investigations of intermolecular interactions in water-amide mixtures. *J. Mol. Struct. THEOCHEM* **759**, 159–164 (2006).
119. Biswas, S. & Mallik, B. S. Aqueous solvation of an amide molecule from first principles molecular simulations: Structure, hydrogen bond dynamics and spectral signature. *J. Mol. Liq.* **212**, 941–946 (2015).
120. Bakó, I. *et al.* Water-formamide mixtures: Topology of the hydrogen-bonded network. *J. Mol. Liq.* **228**, 25–31 (2017).
121. Jedlovszky, P., Idrissi, A. & Jancsó, G. Can existing models qualitatively describe the mixing behavior of acetone with water? *J. Chem. Phys.* **130**, 124516 (2009).
122. Pinke, A. & Jedlovszky, P. Modeling of mixing acetone and water: How can their full miscibility be reproduced in computer simulations? *J. Phys. Chem. B* **116**, 5977–5984 (2012).
123. Howard, P. H. *Handbook of Environmental Fate and Exposure Data for Organic Chemicals, Volume IV.* (Lewis, 1993).
124. Ge, X., Wexler, A. S. & Clegg, S. L. Atmospheric amines - Part I. A review. *Atmospheric Environment* **45**, 524–546 (2011).
125. Yi, Y., Zhou, X., Xue, L. & Wang, W. Air pollution: formation of brown, lighting-absorbing, secondary organic aerosols by reaction of hydroxyacetone and methylamine. *Environ. Chem. Lett.* **16**, 1083–1088 (2018).
126. Cheng, Y., Li, S.-M. M. & Leithead, A. Chemical Characteristics and Origins of Nitrogen-Containing Organic Compounds in PM 2.5 Aerosols in the Lower Fraser Valley. *Environ. Sci. Technol.* **40**, 5846–5852 (2006).
127. Laskin, A., Smith, J. S. & Laskin, J. Molecular Characterization of Nitrogen-Containing Organic Compounds in Biomass Burning Aerosols Using High-Resolution Mass Spectrometry. *Environ. Sci. Technol.* **43**, 3764–3771 (2009).
128. Raja, S. *et al.* Organic composition of fogwater in the Texas-Louisiana gulf coast corridor. *Atmos. Environ.* **43**, 4214–4222 (2009).
129. Zhu, L., Schade, G. W. & Nielsen, C. J. Real-time monitoring of emissions from monoethanolamine-based industrial scale carbon capture facilities. *Environ. Sci. Technol.* **47**, 14306–14314 (2013).
130. Borduas, N., da Silva, G., Murphy, J. G. & Abbatt, J. P. D. Experimental and Theoretical Understanding of the Gas Phase Oxidation of Atmospheric Amides with OH Radicals: Kinetics, Products, and Mechanisms. *J. Phys. Chem. A* **119**, 4298–4308 (2015).
131. Chakir, A., Solognac, G., Mellouki, A. & Daumont, D. Gas phase UV absorption cross-sections for a series of amides. *Chem. Phys. Lett.* **404**, 74–78 (2005).
132. Barnes, I., Solognac, G., Mellouki, A. & Becker, K. H. Aspects of the Atmospheric

- Chemistry of Amides. *ChemPhysChem* **11**, 3844–3857 (2010).
133. Bunkan, A. J. C., Mikoviny, T., Nielsen, C. J., Wisthaler, A. & Zhu, L. Experimental and Theoretical Study of the OH-Initiated Photo-oxidation of Formamide. *J. Phys. Chem. A* **120**, 1222–1230 (2016).
  134. Borduas, N., Abbatt, J. P. D., Murphy, J. G., So, S. & da Silva, G. Gas-Phase Mechanisms of the Reactions of Reduced Organic Nitrogen Compounds with OH Radicals. *Environ. Sci. Technol.* **50**, 11723–11734 (2016).
  135. Stribling, R. & Miller, S. L. Energy yields for hydrogen cyanide and formaldehyde syntheses: The hcn and amino acid concentrations in the primitive ocean. *Orig. life Evol. Biosph.* **17**, 261–273 (1987).
  136. Szori, M. & Jedlovszky, P. Adsorption of HCN at the surface of ice: A grand canonical Monte Carlo simulation study. *J. Phys. Chem. C* **118**, 3599–3609 (2014).
  137. Privat, R. & Jaubert, J.-N. Classification of global fluid-phase equilibrium behaviors in binary systems. *Chem. Eng. Res. Des.* **91**, 1807–1839 (2013).
  138. Allen, M. P. & Tildesley, D. J. Computer simulation of liquids. *Oxford Univ. Press. New York* (1987). doi:10.2307/2938686
  139. Harwood, D. B., Peters, C. J. & Siepman, J. I. A Monte Carlo simulation study of the liquid–liquid equilibria for binary dodecane/ethanol and ternary dodecane/ethanol/water mixtures. *Fluid Phase Equilib.* **407**, 269–279 (2016).
  140. Makihara, T., Akiyama, T., Nakamura, K. & Ito, T. Systematic Theoretical Investigations for Miscibility of GaN<sub>x</sub>As<sub>1-x-y</sub>Bi<sub>y</sub> Thin Films. *e-Journal Surf. Sci. Nanotechnol.* **12**, 171–174 (2014).
  141. Shi, T., Wen, G., Jiang, W., An, L. & Li, B. Monte Carlo simulation of miscibility of polymer blends with repulsive interactions: effect of chain structure. *Eur. Polym. J.* **39**, 551–560 (2003).
  142. Li, Q., Liu, D., Song, L., Wu, P. & Yan, Z. Prediction on miscibility of silicone and gasoline components by Monte Carlo simulation. *J. Mol. Model.* **20**, 2244 (2014).
  143. Cifra, P., Nies, E. & Broersma, J. Equation of State and Miscibility Behavior of Compressible Binary Lattice Polymers. A Monte Carlo Study and Comparison with Partition Function Theories. *Macromolecules* **29**, 6634–6644 (1996).
  144. Horváth, R. A., Horvai, G., Idrissi, A. & Jedlovszky, P. Thermodynamics of mixing methanol with supercritical CO<sub>2</sub> as seen from computer simulations and thermodynamic integration. *Phys. Chem. Chem. Phys.* **22**, 11652–11662 (2020).
  145. Fábrián, B., Szöri, M. & Jedlovszky, P. Floating patches of HCN at the surface of their aqueous solutions - Can they make 'hCN World' plausible? *J. Phys. Chem. C* **118**, 21469–21482 (2014).
  146. Adams, D. J. Grand canonical ensemble Monte Carlo for a Lennard-Jones fluid. *Mol. Phys.* **29**, 307–311 (1975).
  147. Dawley, M. M., Pirim, C. & Orlando, T. M. Thermal processing of formamide ices on silicate grain analogue. *J. Phys. Chem. A* **118**, 1220–1227 (2014).
  148. Darvas, M., Jedlovszky, P. & Jancsó, G. Free energy of mixing of pyridine and its

- methyl-substituted derivatives with water, As seen from computer simulations. *J. Phys. Chem. B* **113**, 7615–7620 (2009).
149. MEZEI, M. & BEVERIDGE, D. L. Free Energy Simulations. *Ann. N. Y. Acad. Sci.* **482**, 1–23 (1986).
  150. Leach, A. R. *Molecular Modelling*. (Longman, 1996).
  151. Pártay, L. B., Jedlovsky, P. & Horvai, G. Structure of the liquid-vapor interface of water-acetonitrile mixtures as seen from molecular dynamics simulations and identification of truly interfacial molecules analysis. *J. Phys. Chem. C* **113**, 18173–18183 (2009).
  152. Fábrián, B., Jójárt, B., Horvai, G. & Jedlovsky, P. Properties of the Liquid-Vapor Interface of Acetone-Water Mixtures. A Computer Simulation and ITIM Analysis Study. *J. Phys. Chem. C* **119**, 12473–12487 (2015).
  153. Mixing in Analytical Chemistry. in *IUPAC Compendium of Chemical Terminology* (IUPAC, 2008). doi:10.1351/goldbook.m03945
  154. Atkin, P. & Paula, J. *Physical chemistry*. W.H. Freeman (2006). doi:10.1039/C1CS15191F
  155. Kiss, L. & Láng, G. *Elektrokémia*. (Semmelweis Kiadó, 2011).
  156. Zrínyi, M. *A fizikai kémia alapjai I.* (Műszaki Könyvkiadó, 2004).
  157. Miscibility. in *IUPAC Compendium of Chemical Terminology* (IUPAC, 2008). doi:10.1351/goldbook.mt07230
  158. Smith, J. M., Van Ness, H. C. & Abbott, M. M. *Introduction to Chemical Engineering Thermodynamics*. *Chemical Engineering* **27**, (2001).
  159. George Kaptay. *Anyagegyensúlyok Makro-, Mikro-, és Nano- Méretű Rendszerekben*. (2011).
  160. Mezei, M.; Swaminathan, S.; Beveridge, D. L. Ab Initio Calculation of the Free Energy of Liquid Water. *J. Am. Chem. Soc.* **100**, 3255–3256 (1978).
  161. Mezei, M. Direct calculation of the excess free energy of the dense Lennard-Jones fluid. *Mol. Simul.* **2**, 201–207 (1989).
  162. Mezei, M. Polynomial path for the calculation of liquid state free energies from computer simulations tested on liquid water. *J. Comput. Chem.* **13**, 651–656 (1992).
  163. Mináry, P., Jedlovsky, P., Mezei, M. & Turi, L. A comprehensive liquid simulation study of neat formic acid. *J. Phys. Chem. B* **104**, 8287–8294 (2000).
  164. Jedlovsky, P. *et al.* Structural and thermodynamic properties of different phases of supercooled liquid water. *J. Chem. Phys.* **128**, 244503 (2008).
  165. Rowlinson, J. S. *Liquids and Liquid Mixtures*. (Butterworth, 1969).
  166. T. M. Reed, K. E. G. *Gázok és folyadékok statisztikus termodinamikája*. (Műszaki Könyvkiadó, 1978).
  167. N. Ashcroft, N. M. *Solid State Physics*. (Holt, Rinehart and Winston, 1976).

168. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
169. Barker, J. A. & Watts, R. O. Monte carlo studies of the dielectric properties of water-like models. *Mol. Phys.* **26**, 789–792 (1973).
170. Barker, J. A. & Watts, R. O. Structure of water; A Monte Carlo calculation. *Chem. Phys. Lett.* **3**, 144–145 (1969).
171. Watts, R. O. Monte carlo studies of liquid water. *Mol. Phys.* **28**, 1069–1083 (1974).
172. Neumann, M. The dielectric constant of water. Computer simulations with the MCY potential. *J. Chem. Phys.* **82**, 5663–5672 (1985).
173. Cooley, J. W. & Tukey, J. W. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* **19**, 297–301 (1965).
174. García Fernández, R., Abascal, J. L. F. & Vega, C. The melting point of ice Ih for common water models calculated from direct coexistence of the solid-liquid interface. *J. Chem. Phys.* **124**, (2006).
175. Hantal, G., Jedlovsky, P., Hoang, P. N. M. & Picaud, S. Investigation of the adsorption behaviour of acetone at the surface of ice. A grand canonical Monte Carlo simulation study. *Phys. Chem. Chem. Phys.* **10**, 6369–6380 (2008).
176. Jedlovsky, P., Vincze, Á. & Horvai, G. New insight into the orientational order of water molecules at the water/1,2-dichloroethane interface: A Monte Carlo simulation study. *J. Chem. Phys.* **117**, 2271–2280 (2002).
177. Toxvaerd, S. & Stecki, J. Density profiles at a planar liquid-liquid interface. *J. Chem. Phys.* **102**, 7163–7168 (1995).
178. Chapela, G. A., Saville, G., Thompson, S. M. & Rowlinson, J. S. Computer simulation of a gas-liquid surface. Part 1. *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics* **73**, 1133–1144 (1977).
179. Holcomb, C. D., Clancy, P. & Zollweg, J. A. A critical study of the simulation of the liquid-vapour interface of a Lennard-Jones fluid. *Mol. Phys.* **78**, 437–459 (1993).
180. Chen, L. J. Area dependence of the surface tension of a Lennard-Jones fluid from molecular dynamics simulations. *J. Chem. Phys.* **103**, 10214–10216 (1995).
181. Jedlovsky, P. Rendezetlen kondenzált fázisok tulajdonságainak vizsgálata számítógépes szimulációs módszerekkel. (Eötvös Loránd Tudományegyetem, 2006).
182. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092 (1953).
183. Daan Frenkel, B. S. *Understanding Molecular Simulation*. (Academic Press, 2002).
184. Mezei, M. A cavity-biased (T, V,  $\mu$ ) monte carlo method for the computer simulation of fluids. *Mol. Phys.* **40**, 901–906 (1980).
185. Mezei, M. Grand-canonical ensemble monte carlo study of dense liquid Lennard-Jones, soft spheres and water. *Mol. Phys.* **61**, 565–582 (1987).

186. Verlet, L. Computer 'experiments' on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **159**, 98–103 (1967).
187. Penfold, J. The structure of the surface of pure liquids. *Reports Prog. Phys.* **64**, 777–814 (2001).
188. Linse, P. Monte Carlo simulation of liquid-liquid benzene-water interface. *J. Chem. Phys.* **86**, 4177–4187 (1987).
189. Benjamin, I. Theoretical study of the water/1,2-dichloroethane interface: Structure, dynamics, and conformational equilibria at the liquid-liquid interface. *J. Chem. Phys.* **97**, 1432–1445 (1992).
190. Reed, T. M. & Gubbins, K. E. *Applied Statistical Mechanics: Thermodynamic and Transport Properties of Fluids*. (McGraw-Hill, 1973).
191. Pártay, L. B., Hantal, G., Jedlovszky, P., Vincze, Á. & Horvai, G. A new method for determining the interfacial molecules and characterizing the surface roughness in computer simulations. Application to the liquid-vapor interface of water. *J. Comput. Chem.* **29**, 945–956 (2008).
192. Hantal, G., Darvas, M., Pártay, L. B., Horvai, G. & Jedlovszky, P. Molecular level properties of the free water surface and different organic liquid/water interfaces, as seen from ITIM analysis of computer simulation results. *J. Phys. Condens. Matter* **22**, (2010).
193. Pártay, L. B., Horvai, G. & Jedlovszky, P. Temperature and pressure dependence of the properties of the liquid-liquid interface. A computer simulation and identification of the truly interfacial molecules investigation of the Water-Benzene System. *J. Phys. Chem. C* **114**, 21681–21693 (2010).
194. Rowlinson, J. S. & Widom, B. *Molecular Theory of Capillarity*. **88**, (Clarendon Press, 1982).
195. Sega, M., Kantorovich, S. S., Jedlovszky, P. & Jorge, M. The generalized identification of truly interfacial molecules (ITIM) algorithm for nonplanar interfaces. *J. Chem. Phys.* **138**, 044110 (2013).
196. Chacón, E. & Tarazona, P. Intrinsic profiles beyond the capillary wave theory: A monte carlo study. *Phys. Rev. Lett.* **91**, 166103 (2003).
197. Chowdhary, J. & Ladanyi, B. M. Water/hydrocarbon interfaces: Effect of hydrocarbon branching on single-molecule relaxation. *J. Phys. Chem. B* **112**, 6259–6273 (2008).
198. Jorge, M. & Cordeiro, M. N. D. S. Intrinsic structure and dynamics of the water/nitrobenzene interface. *J. Phys. Chem. C* **111**, 17612–17626 (2007).
199. Jorge, M., Jedlovszky, P. & Cordeiro, M. N. D. S. A critical assessment of methods for the intrinsic analysis of liquid interfaces. 1. surface site distributions. *J. Phys. Chem. C* **114**, 11169–11179 (2010).
200. Sega, M. The role of a small-scale cutoff in determining molecular layers at fluid interfaces. *Phys. Chem. Chem. Phys.* **18**, 23354–23357 (2016).
201. Pártay, L. B., Jedlovszky, P., Vincze, Á. & Horvai, G. Properties of free surface of

- water-methanol mixtures. Analysis of the truly interfacial molecular layer in computer simulation. *J. Phys. Chem. B* **112**, 5428–5438 (2008).
202. Pojják, K., Darvas, M., Horvai, G. & Jedlovsky, P. Properties of the liquid-vapor interface of water-dimethyl sulfoxide mixtures. A molecular dynamics simulation and ITIM analysis study. *J. Phys. Chem. C* **114**, 12207–12220 (2010).
  203. Voronoi, G. F. Recherches sur le Paralléloèders Primitives. *J. Reine Angew. Math.* **134**, 198–287 (1908).
  204. Medvedev, N. N. *The Voronoi-Delaunay Method in the Structural Investigation of Non-Crystalline Systems*. (SB RAS: Novosibirsk, 2000).
  205. Okabe, A.; Boots, B.; Sugihara, K.; Chiu, S. N. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. (John Wiley: Chichester, 2000).
  206. Zaninetti, L. The Voronoi tessellation generated from different distributions of seeds. *Phys. Lett. A* **165**, 143–147 (1992).
  207. Davis, P. J.; Polonsky, I. *Numerical Interpolation, Differentiation and Integration*. In *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. (National Bureau of Standards, 1972).
  208. García, B., Alcalde, R., Leal, J. M. & Matos, J. S. Solute-solvent interactions in amide-water mixed solvents. *J. Phys. Chem. B* **101**, 7991–7997 (1997).
  209. Mezei, M. MMC: Monte Carlo program for simulation of molecular assemblies. *MMC: Monte Carlo Program for Simulation of Molecular Assemblies* URL: <http://inka.mssm.edu/~mezei/mmc> (2004).
  210. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. 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 (1996).
  211. Lindahl, E., Bjelkmar, P., Larsson, P., Cuendet, M. A. & Hess, B. Implementation of the charmm force field in GROMACS: Analysis of protein stability effects from correction maps, virtual interaction sites, and water models. *J. Chem. Theory Comput.* **6**, 459–466 (2010).
  212. Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269–6271 (1987).
  213. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
  214. Abascal, J. L. & Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **123**, 234505 (2005).
  215. Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).
  216. Pronk, S. *et al.* GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
  217. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).

218. Hoover, W. G. W. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
219. Petitjean, M. *et al.* Adsorption of benzaldehyde at the surface of ice, studied by experimental method and computer simulation. *Langmuir* **26**, 9596–9606 (2010).
220. Hantal, G., Jedlovsky, P., Hoang, P. N. M. & Picaud, S. Calculation of the adsorption isotherm of formaldehyde on ice by grand canonical Monte Carlo simulation. *J. Phys. Chem. C* **111**, 14170–14178 (2007).
221. Jedlovsky, P. *et al.* Determination of the adsorption isotherm of methanol on the surface of ice. An experimental and Grand Canonical Monte Carlo simulation study. *J. Am. Chem. Soc.* **128**, 15300–15309 (2006).
222. Mészár, Z. E., Hantal, G., Picaud, S. & Jedlovsky, P. Adsorption of aromatic hydrocarbon molecules at the surface of ice, as seen by grand canonical monte carlo simulation. *J. Phys. Chem. C* **117**, 6719–6729 (2013).
223. Sumi, I., Fábíán, B., Picaud, S. & Jedlovsky, P. Adsorption of Fluorinated Methane Derivatives at the Surface of Ice under Tropospheric Conditions, As Seen from Grand Canonical Monte Carlo Simulations. *J. Phys. Chem. C* **120**, 17386–17399 (2016).
224. Fu, Z. *et al.* Adsorption of Nitrobenzene on the Surface of Ice: A Grand Canonical Monte Carlo Simulation Study. *J. Phys. Chem. C* **121**, 15746–15755 (2017).
225. Szentirmai, V., Szöri, M., Picaud, S. & Jedlovsky, P. Adsorption of Methylamine at the Surface of Ice. A Grand Canonical Monte Carlo Simulation Study. *J. Phys. Chem. C* **120**, 23480–23489 (2016).
226. Idrissi, A. *et al.* Free energy of mixing of acetone and methanol: A computer simulation investigation. *J. Phys. Chem. B* **117**, 16157–16164 (2013).
227. Villamañán, M. A. & Van Ness, H. C. Excess Thermodynamic Properties for Water/Acetone. *J. Chem. Eng. Data* **29**, 429–431 (1984).
228. Perera, A. & Sokolić, F. Modeling nonionic aqueous solutions: The acetone-water mixture. *J. Chem. Phys.* **121**, 11272–11282 (2004).
229. Segá, M., Fábíán, B. & Jedlovsky, P. Nonzero Ideal Gas Contribution to the Surface Tension of Water. *J. Phys. Chem. Lett.* **8**, 2608–2612 (2017).
230. Chowdhary, J. & Ladanyi, B. M. Surface fluctuations at the liquid-liquid interface. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **77**, 031609 (2008).
231. Idrissi, A., Hantal, G. & Jedlovsky, P. Properties of the liquid-vapor interface of acetone-methanol mixtures, as seen from computer simulation and ITIM surface analysis. *Phys. Chem. Chem. Phys.* **17**, 8913–8926 (2015).
232. Fábíán, B., Segá, M., Horvai, G. & Jedlovsky, P. Single Particle Dynamics at the Intrinsic Surface of Various Apolar, Aprotic Dipolar, and Hydrogen Bonding Liquids As Seen from Computer Simulations. *J. Phys. Chem. B* **121**, 5582–5594 (2017).
233. Kiang, T. Random Fragmentation in Two and Three Dimensions. *Z. Astrophys.* **64**, 433–439 (1966).
234. Pineda, E., Bruna, P. & Crespo, D. Cell size distribution in random tessellations of

- space. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* **70**, 8 (2004).
235. Idrissi, A., Damay, P., Yukichi, K. & Jedlovszky, P. Self-association of urea in aqueous solutions: A Voronoi polyhedron analysis study. *J. Chem. Phys.* **129**, 164512 (2008).
236. Kiss, B., Fábrián, B., Idrissi, A., Szöri, M. & Jedlovszky, P. Miscibility and Thermodynamics of Mixing of Different Models of Formamide and Water in Computer Simulation. *J. Phys. Chem. B* **121**, 7147–7155 (2017).
237. Sega, M., Horvai, G. & Jedlovszky, P. Microscopic origin of the surface tension anomaly of water. *Langmuir* **30**, 2969–2972 (2014).
238. Sega, M., Horvai, G. & Jedlovszky, P. Two-dimensional percolation at the free water surface and its relation with the surface tension anomaly of water. *J. Chem. Phys.* **141**, 054707 (2014).
239. Stauffer, D. *Introduction to Percolation Theory*. (Taylor and Francis: London, 1985).
240. Jedlovszky, P., Vincze, Á. & Horvai, G. Full description of the orientational statistics of molecules near to interfaces. Water at the interface with CCl<sub>4</sub>. in *Physical Chemistry Chemical Physics* **6**, 1874–1879 (2004).
241. Rideg, N. A., Darvas, M., Varga, I. & Jedlovszky, P. Lateral dynamics of surfactants at the free water surface: A computer simulation study. *Langmuir* **28**, 14944–14953 (2012).
242. Daub, C. D., Patey, G. N., Jack, D. B. & Sallabi, A. K. Monte Carlo simulations of the adsorption of CO<sub>2</sub> on the MgO(100) surface. *J. Chem. Phys.* **124**, 114706 (2006).
243. Langmuir, I. The constitution and fundamental properties of solids and liquids. Part I. Solids. *J. Am. Chem. Soc.* **38**, 2221–2295 (1916).
244. Shaw, D. J. *Introduction to Colloid and Surface Chemistry*. (Butterworths, 1980).
245. Horváth, R. A., Hantal, G., Picaud, S., Szöri, M. & Jedlovszky, P. Adsorption of Methylamine on Amorphous Ice under Interstellar Conditions. A Grand Canonical Monte Carlo Simulation Study. *J. Phys. Chem. A* **122**, 3398–3412 (2018).
246. Thomson, G. W. The antoine equation for vapor-pressure data. *Chem. Rev.* **38**, 1–39 (1946).
247. Stull, D. R. Vapor Pressure of Pure Substances. Organic and Inorganic Compounds. *Ind. Eng. Chem.* **39**, 517–540 (1947).
248. NIST Chemistry WebBook SRD 69. Available at: <http://webbook.nist.gov/%0Acgi/cbook.cgi?ID=C75127&Mask=4&Type=ANTOINE&Plot=on>. (Accessed: 6th December 2018)
249. Sarkar, C. *et al.* Overview of VOC emissions and chemistry from PTR-TOF-MS measurements during the SusKat-ABC campaign: High acetaldehyde, isoprene and isocyanic acid in wintertime air of the Kathmandu Valley. *Atmos. Chem. Phys.* **16**, 3979–4003 (2016).
250. Cimino, R. T., Kowalczyk, P., Ravikovitch, P. I. & Neimark, A. V. Determination of

Isosteric Heat of Adsorption by Quenched Solid Density Functional Theory. *Langmuir* **33**, 1769–1779 (2017).

251. Corazzi, M. A., Fedele, D., Poggiali, G. & Brucato, J. R. Photoprocessing of formamide ice: route towards prebiotic chemistry in space. *Astron. Astrophys.* **636**, A63 (2020).
252. Rotelli, L. *et al.* The key role of meteorites in the formation of relevant prebiotic molecules in a formamide/water environment. *Sci. Rep.* **6**, (2016).