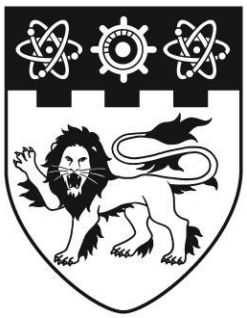
INSIGHTS INTO THE SOLVATION OF GLUCOSE IN WATER, DIMETHYL SULFOXIDE (DMSO), TETRAHYDROFURAN (THF) AND N,N-DIMETHYLFORMAMIDE (DMF) AND ITS POSSIBLE IMPLICATIONS ON GLUCOSE CHEMISTRY



NANYANG TECHNOLOGICAL UNIVERSITY

INSIGHTS INTO THE SOLVATION OF GLUCOSE IN WATER, DIMETHYL SULFOXIDE (DMSO), TETRAHYDROFURAN (THF) AND N,N-DIMETHYLFORMAMIDE (DMF) AND ITS POSSIBLE IMPLICATIONS ON GLUCOSE CHEMISTRY

VALLABH VASUDEVAN

School of Chemical and Biomedical Engineering

2015

2015

INSIGHTS INTO THE SOLVATION OF GLUCOSE IN WATER, DIMETHYL SULFOXIDE (DMSO), TETRAHYDROFURAN (THF) AND N,N-DIMETHYLFORMAMIDE (DMF) AND ITS POSSIBLE IMPLICATIONS ON GLUCOSE CHEMISTRY

VALLABH VASUDEVAN

School of Chemical and Biomedical Engineering

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of
Master of Engineering

2015

Abstract

Biomass is a carbon neutral renewable alternative to fossil fuels, and can help in reducing the demand for fossil fuels in the near future. This study is relevant to the liquid phase catalytic conversion of biomass to fuels, fuel additives and chemicals. Products obtained during the catalytic conversion can be used as platform chemicals for further upgradation to bulk chemicals, fuels and fuel additives. One such platform chemical is 5-hydroxymethylfurfural (HMF), which can be produced by the acid catalyzed dehydration of glucose, which is the largest monomeric component of lignocellulosic biomass. Recent research on the conversion of glucose to HMF has shown that the conversion and yield of the reaction are affected by the choice of solvents used. It has been noted that the solvents not only play the role of a benign solvation medium but also are involved in the conversion process and significantly affect the conversion and yield of the reaction. However, the exact role of solvents, as the reaction media, is not yet well understood. It would be extremely difficult to decouple physical solvation and chemical participation effects using experimental methods, so as to study them individually. To overcome this difficulty, computational modeling is used. The present work focuses on identifying key physical interactions between biomass molecules and solvents using molecular modeling. We perform molecular dynamics simulations, using OPLS/AA force field to investigate the physical solvation of glucose in four common solvents used in Biomass conversion studies, namely, water, Dimethylsulfoxide (DMSO), N,N-Dimethylformamide (DMF) and Tetrahydrofuran (THF). The parameters for the solvents were selected by careful consideration and were benchmarked with existing data from the literature to ensure accuracy. Force-field parameters for DMF, as reported in the literature with a negative partial charge on the amide nitrogen, resulted in poor prediction of the bulk properties like density and dielectric constant and showed reduced miscibility with water in a DMF-water mixed system. To address these issues, we proposed a new set of parameters with a small positive partial charge on the amide nitrogen. The new parameters were benchmarked with both, preexisting experimental data and using ab initio molecular dynamics, to ensure better accuracy of the DMF molecule. The local arrangement of solvents around the glucose molecule is analyzed using 2-dimensional radial pair distribution functions and 3-dimensional volumetric maps. Additionally, lifetimes and activation free energies of hydrogen bonds between solvents and glucose and the tendency of glucose molecules to agglomerate were studied. From our simulation

studies and analysis it was observed that i) all the aforementioned co-solvents compete with water to be in the first solvation shell of glucose and significant amount of water is pushed to the second coordination shell; ii) though fewer water molecules are directly coordinated with glucose in the presence of co-solvents, they are bound strongly to it; iii) DMSO, THF and DMF tend to localize more around the hydrogen atom of the hydroxyl groups of selected carbon atoms of glucose; iv) the preferential arrangement of co-solvents and water around glucose may play a role in facilitating the reaction pathway for the formation of HMF and levulinic acid and may reduce the likelihood of glucose' degradation to unwanted dehydration/rehydration products; v) Increasing the proportion of co-solvents also increases the hydrogen bond lifetimes between water and glucose and reduces the mobility of glucose molecules within the solvent. The reduced mobility of glucose molecules in the presence of co-solvents might be correlated to the experimentally observed reduction in the rate of formation of polymerization/condensation products and humins.

Acknowledgements

I would like to extend my heartfelt gratitude to my advisor, Prof. Samir Hemant Mushrif, for the constant encouragement and support, innumerable discussions and timely guidance throughout the last thirty months. My initiation into Computational Chemistry would have been extremely arduous without his support.

My gratitude goes to Dr. B.K. Chethana, Mr. Jithin John Varghese and Dr. Trinh Thang Quang for their encouragement and the general discussions on computational methods. I also thank Mr. Boddu Venkatesh and Ms. Sukriti Gupta for their encouragement.

My family and friends have always been a source of unconditional love and affection and inspiration which have been key motivations to pursue the doctoral program. My sincere gratitude to my parents, family and friends for being there to share my joys and to cheer me up in times of despair and frustration.

I also thank NTU and SCBE for the financial support, resources and facilities for carrying out my research.

I would like to acknowledge the financial support provided by Nanyang Technological University, Singapore under the Campus for Research Excellence and Technological Enterprise (CREATE) programme of the National Research Foundation (NRF), Prime Minister's Office, Singapore

Table of Contents

List of Figures and Schemes

Figures

Schemes

List of Tables

List of Abbreviations

HMF	5-Hydroxymethylfurfural
DMSO	Dimethylsulfoxide
DMF	N,N-Dimethylformamide
THF	Tetrahydrofuran
RDF	Radial Pair Distribution Functions
RMSF	Root Mean Squared Fluctuations
PEG	Polyethylene Glycol
GROMACS	Groningen Machine for Chemical Simulations
VMD	Visual Molecular Dynamics
OPLS/AA	Optimized Potentials for Liquid Solvents/ All Atom Model
NMR	Nuclear Magnetic Resonance
MD	Molecular Dynamics
NVT	Normal Volume and Temperature

Chapter 1: Introduction

With the increasing global demand to cut down on the usage of fossil fuels, the research focus has shifted towards renewable resources to slowly replace/supplement fossil fuels. As one of the possible resources, biomass has garnered significant attention in this area. Thus a lot of research in chemistry, engineering and agriculture are focused on utilizing this source of fuel and use it in a sustainable way. This has been mainly driven by a few factors such as (i) the requirement for a new carbon based fuel that can be incorporated into the existing machinery to compensate the dwindling fuel reserves and also a drop in petrol production; and (ii) the requirement of a fuel source that does not lead to a rapid increase in the production of greenhouse gases, without significant reduction in the amount of energy produced. These factors have led to biomass becoming a front-runner in this field[1].

Traditionally biomass has been classified into three types of feedstocks, based on the nature of the source. They are starchy feedstocks, triglyceride feedstocks and lignocellulosic feedstocks. Starchy feedstocks are comprised of glucose polysaccharides joined by α- glycosidic linkages, such as amylase and amylopectin, which can be readily hydrolyzed into the constituent sugar monomers, making them easy to process. Trigylderide feedstocks are comprised of fatty acids and glycerol derived from plant and animal sources. Sources of triglycerides for the production of fuels include various vegetable oils, waste oil products (e.g., yellow grease, trap grease), and algal sources.[1] Lignocellulosic biomass is the most abundant class of biomass. While starch and triglycerides are only present in some crops, lignocellulose contributes structural integrity to plants and is thus always present. We shall look at lignocellulosic biomass in detail in the next section.

The bio-fuels that can be produced from biomass can be classified into 3 distinct categories. They are

i. 1st generation bio-fuels: They are basically made up of bio-ethanol and bio-diesel and use conventional methods to process food crops to derive them
ii. 2nd generation bio-fuels: They use lignocellulosic biomass to produce bio-oils and platform chemicals that can be used for various purposes.

iii. 3rd generation bio-fuels: They are relatively new form of bio-fuels that are generated from triglyceride sources

1.1. Lignocellulosic Biomass

Lignocellulosic biomass is an attractive, cheap and abundantly available feedstock. It is the biomass obtained from plants. Lignocellulosic biomass consists of three components, namely, Lignin (15-20%), Cellulose (40-50%), and Hemicellulose (25-35%).[2] The major problem with the usage of Lignocellulosic biomass is that it must be broken down into its primary components before it can be converted to fuels. This is because the different components have different chemical compositions and they behave differently to the same reaction conditions. The cellulose and hemicellulose components are mostly made up of sugar groups and the lignin component is mostly made of aromatic monomers. The reaction chemistry for these different components is different and thus in the liquid phase catalytic processing, they need to be treated separately. The separation processes for separating lignin from cellulose and hemicellulose are expensive and can lead to inflated prices of fuels and chemicals from Lignocellulosic biomass.[3]

1.1.1. Lignin

Lignin is a complex amorphous polymer usually found in the cell wall of plants, and provides plants with structural rigidity and a hydrophobic vascular system for the transportation of water and solutes.[4] Though the structure for lignin is very complex, in general, it is a dendritic network of phenyl propene basic units.[5,6] The purpose of pretreatment is to depolymerize Lignin so that the cellulose and hemicellulose can be reached.[1] Though Lignin can be separated, the reaction chemistry for converting lignin into useful compounds is not easy and thus difficult to be used for production of fuels. The major usage of lignin still is by direct burning,[7] though recent experimental work has shown that aromatic compounds from lignin can be used as an effective solvent in the conversion of cellulose to bio-fuels.[8]

1.1.2. Hemicellulose

Hemicellulolse is an amorphic polymer in the plant cell wall.[9] It mostly consists of five different sugar monomers; D-xylose, L-arabinose, D-galactose, D-glucose, and D-mannose, with xylose

being the most abundant. Hemicellulose is bound to lignin and interlaced with cellulose.[1] It is usually removed during the pretreatment steps so that the glucose can be recovered more efficiently form cellulose in the subsequent hydrolysis steps. Hemicellulose extraction is done by both, physical and chemical methods, such that xylose is usually preserved. Xylose monomers are easily formed by simple acid hydrolysis of hemicellulose and this xylose can be used as a feedstock for the production of bio-ethanol by fermentation[10,11] or for the production of furfural.[12]

1.1.3. Cellulose

Cellulose is a polymer consisting of a linear chain of several hundred to over ten thousand β-glycosidic linked D-glucose units.[9] Cellulose is typically present along with the complex lignin/hemicellulose matrix, and is largely inaccessible for hydrolysis in untreated biomass. Biomass pretreatment via milling and physical/chemical treatments helps to break this matrix and extract hemicellulose, so that the subsequent hydrolysis steps to convert cellulose to the glucose monomers are more effective. This hydrolysis step to convert cellulose is considered to be more difficult than the similar conversion of hemicellulose to xylose. High yields of glucose are usually obtained by enzymatic hydrolysis of cellulose after the pretreatment of biomass. Non-enzymatic conversion can be achieved in the presence of mineral acids under harsher conditions at elevated temperatures. However, harsher conditions required for the conversion to glucose facilitate the formation of dehydration products such as 5-hydroxymethylfurfural, levulinic acid and humins.[1] This basic degradation of glucose in acidic conditions is one of the important ways for the production of bio-fuels/chemicals by catalytic conversion of biomass.

1.2. General Approaches for the conversion of Lignocellulosic Biomass to Bio-fuels

The main challenge in the conversion of biomass to fuels and chemicals is to remove the oxygen in it, without breaking the C-C bonds. There are three basic approaches for biomass conversion. They are (i) pyrolysis to produce bio-oil; and (ii) liquid phase catalytic process to produce platform chemicals which can then be upgraded; and (iii) biochemical conversion of biomass components by digesting them using enzymes to derive specific products.[1] The present work focusses on the liquid phase catalytic conversion method and the other two pathways are beyond the scope of this thesis.

1.2.1 Catalytic upgrading pathways for biomass conversion

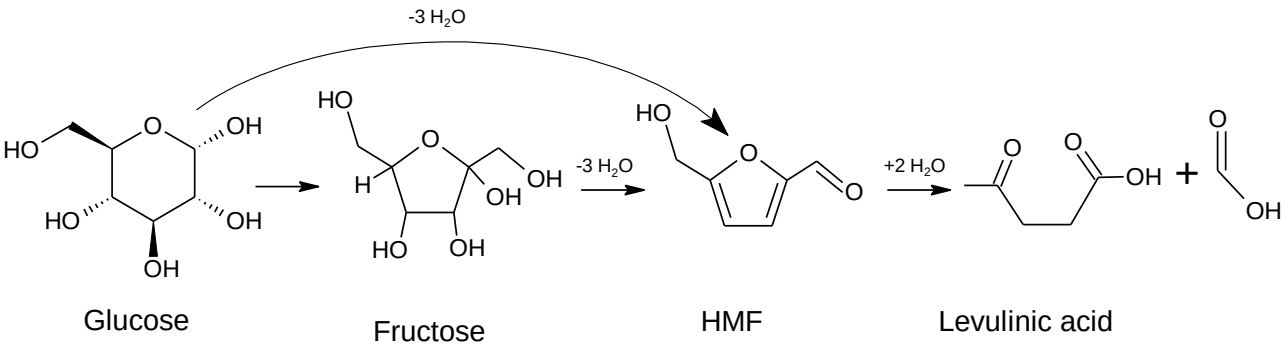
Liquid phase catalytic processing is the preferred method when selectivity towards a specific product, such as Hydroxymethylfurfural (HMF) is desired.[1] The process is usually carried out in the aqueous phase and this makes them very attractive as aqueous phase processes do not require concentration of the aqueous solution and generally yield a gas phase or hydrophobic product that can be easily separated from water, which reduces the cost of separation in the catalytic processing strategies. One reason to favor the catalytic steps is the reduction in oxygen content that is achieved by this process due to various reaction pathways such as hydrogenolysis, dehydration or hydrogenation.[13] Another advantage of this process is that the processing time is reduced.[14] The limitation of this pathway is that the feedstock needs to be pretreated before it can be treated.[14,15]

In the following section we will look into some of the existing work that has been done on the conversion of biomass molecules like glucose and fructose to specific platform chemicals like HMF and levulinic acid. We will look into some of the experimental and computational work done to study the role of solvents in the conversion of biomass molecules. In chapter 4 we will look at some basic introduction into classical mechanics based molecular dynamics and a look into some of the analytical methods used. Chapter 5 discusses about the benchmarking of parameters to ensure accuracy of simulations with an in depth look into the errors present when literature data was used to select parameters for N,N-Dimethylformamide and the solution proposed by us. Finally we shall look into the solvation of glucose in pure and mixed solvent systems. The work done in this thesis focusses on the conversion of glucose to HMF and the effects of solvents on the conversion mechanism. We will also have a brief overview of both experimental and simulation studies that have been done on the role of solvents in the conversion process.

Chapter 2: Literature Review

The chemistry of converting lignocellulosic biomass to hydrocarbon fuels involves the removal of oxygen and the formation of C–C bonds, utilizing the least amount of hydrogen from an external source. This comprises of two steps: (i) conversion of the lignocellulosic biomass feedstock to a platform chemical such as 5-Hydroxymethylfurfural (HMF), Levulinic Acid, etc., involving partial removal of oxygen; and (ii) usage of these platform chemicals as feedstocks to produce hydrocarbons suitable for fuels/chemicals applications, by achieving controlled C–C coupling reactions while minimizing the undesirable branching processes and the removal of the remaining oxygen functionality.[1]

Of all the platform chemicals, the present work will focus on the conversion of glucose to HMF (*cf* scheme 1). Other major platform chemicals such as levulinic acid and gamma-valerolactone can be derived from HMF. HMF can be generated from both the cellulosic and hemicellulosic components of biomass and can thus utilize a larger fraction of the feedstock, hence increasing the interest in the chemical as the desired product from biomass conversion.



Scheme 1[16]: One of the possible mechanisms of the conversion of glucose to HMF

2.1. HMF synthesis and challenges associated with it:

HMF is synthesized mainly by the dehydration of glucose/fructose, requiring the removal of three water molecules. Antal et al.[17] reported that possible side products are formed by the decomposition of fructose in water at high temperatures, due to isomerization, dehydration,

fragmentation and condensation[16] Reducing water content is also critical to the selective preparation of HMF, because it is readily hydrated in water to form levulinic acid and formic acid.[18] Levulinic acid is thermodynamically more stable and thus HMF is readily converted into LA, thus reducing the possibility of HMF based chemicals such as 2,5-FDCA, 2,5-Diformylfuran (2,5-DFF), 2,5-bis(hydroxymethyl)furan (2,5-BHMF).[19]

In the present work glucose, the monomer of cellulose, is considered as the starting material for the conversion reaction. HMF can be produced with good selectivity (e.g., 90%) from xylose and fructose respectively in biphasic reactors,[20,21] whereas yields are lower for glucose (42% at low concentrations of 3 wt%).[18]

2.2. Glucose vs Fructose as starting material:

Reactivity of glucose (aldose) is lower than that of fructose (ketose), and this fact has been explained by the much lower abundance of acyclic glucose compared to acyclic fructose.[22,23] Glucose can form a very stable ring structure, so the enolisation rate in solution is lower than fructose, which forms less stable ring structures.[22] Since enolisation is the rate-determining step for HMF formation, fructose will react much faster than glucose. On the other hand, fructose forms equilibrium mixtures of difructose and dianhydrides, and thus the most reactive groups are internally blocked, forming smaller amounts of by-products.[22] Glucose forms true oligosaccharides which still contain reactive reducing groups, resulting in a greater risk of cross-polymerization with reactive intermediates and HMF.[22]

There are different conversion techniques for the catalytic conversion of glucose to HMF such as homogeneous and heterogeneous catalytic systems. The study here focuses on the homogeneous catalytic conversion of glucose to HMF with a specific interest in studying solvent effects in the reaction system.

2.3. Solvent effects on glucose to HMF conversion

In 1987 Musau et al.[24] demonstrated that fructose can be converted into HMF in the presence of DMSO as a solvent at 150 ◦C, with no catalyst added.[24] They tested different fructose/Dimethylsulfoxide (DMSO) molar ratios and found that an optimum conversion was

attained for a ratio of 0.8.The authors suggested that DMSO associates initially with only D-fructose at the start of the dehydration reaction, after which the generated water associates with DMSO, reducing the amount of DMSO available to D-fructose. Consequently, DMSO had to be sufficiently in excess to associate with all the water released at the end of the reaction. Amarasekara et al.[25] studied the mechanism of the dehydration of fructose to HMF in DMSO at 150 ◦C without any added mineral or Lewis acid catalyst by monitoring the reaction by NMR spectroscopy (Scheme 2). It was possible to identify an intermediate (4R,5R)-4-hydroxy-5-hydroxymethyl4,5-dihydrofuran-2-carbaldehyde using a combination of 1H and 13C NMR spectra. Ionic liquids have been promising solvents for carbohydrate transformations.[26] These solvents can dissolve carbohydrates, even at high concentrations,[27,28] and can be easily recycled. Furthermore, they have been shown not only to act as solvents, but also as reaction promoters for glucose dehydration reactions.[29,30] The problem with using DMSO as a solvent in high temperature systems is the possible formation of toxic sulfur based compounds at 165 °C and hence is difficult to scale up using this mechanism.

Alamillo et al.[31] showed that the conversion of glucose to HMF in a THF-water system showed similar conversion rates and yields to those got in the biphasic system made up of water and methyl-isobutyl ketone (MIBK) which was shown to have very high conversion and selectivities by Roman-Leshkov et al.[32] Mellmer et al[33] have suggested that the enhanced stability provided by water of the reactant state can increase the activation barrier for the conversion of glucose to HMF.

Both experimental and simulation studies so far have not been able to answer the question of the changes that occur in the solvent system as a co-solvent is added and how the co-solvent concentration affects glucose molecules. Roman-Leshkov et al[32] has studied different co-solvents and shown that the for glucose and fructose the conversion and yield increases as a co-solvent is added. The use of ionic liquids as co-solvents has been studied quite often recently using both experimental and computational methods, while the common co-solvents such as DMSO, N,N-Dimethylformamide (DMF), etc. and their effect on the reaction environment have not been studied.

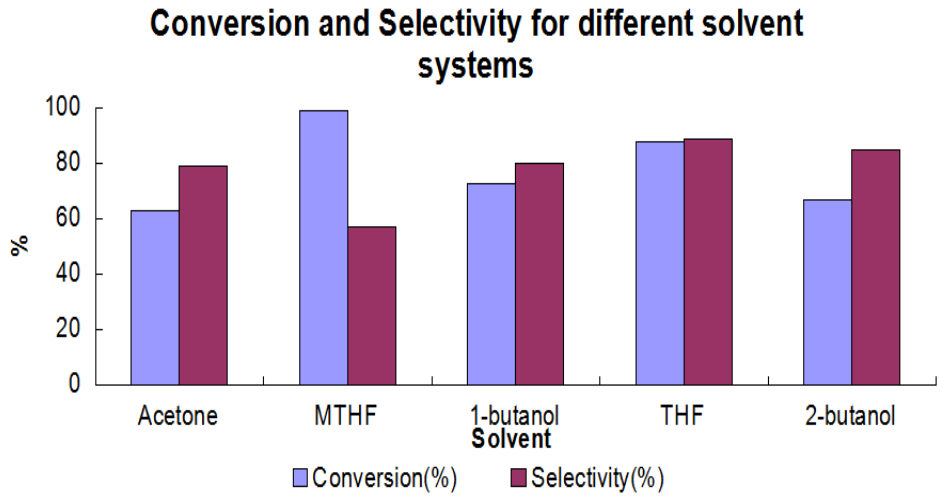


Fig.1 Conversion (blue) and selectivity in the conversion of glucose to HMF in a water co-solvent mixed system. The co-solvent used is shown on the x-axis[32]

The effects of the addition of a co-solvent can be classified into two types, viz., physical and chemical. Physical solvation effects would include preferential solvation of selected functional groups of biomass molecules and their derivatives, thereby protecting them from undesired side reactions.[34,35] Whereas, chemical effects of the co-solvent would include direct participation of the solvent in the reaction as a catalyst[25,36] or solvent dynamics/environment altering the key activation energy barriers in the reaction mechanism.[33,37-41] It would be extremely difficult to decouple physical solvation and chemical participation effects using experimental methods, so as to study them individually. However, it is possible to do it computationally by combining force-field/molecular mechanics based simulations with first–principles calculations. The force-field based molecular simulations can accurately capture physical interactions, but do not contain the physics to model chemical reactions; whereas the first–principles simulations can investigate both, physical and chemical interactions The increased accuracy in the first-principles simulations makes it extremely expensive and force-field based molecular simulations can lead to lower costs and high accuracy especially in the study of physical solavation. .

2.4. Investigating solvent effects in biomass conversion using molecular modeling

Solvent characteristics not only alter the yield and selectivity in liquid phase biomass reactions but also affect solvation and deconstruction of untreated lignocellulosic biomass and individual polymeric biomass species.[42-46] This section provides a brief summary of the recent molecular

level modeling efforts to understand the role of solvent in biomass processing and provides a perspective for future work. Very little effort has been done so far in this area.

2.4.1. Solvents for biomass dissolution and deconstruction

Bergenstrahle et al. (2010)[47] used molecular mechanics–MD simulations to show that cellulose oligomers preferred to pair with each other than with water due to the presence of hydrophobic parts in cellulose, thus preferring to stack on one another. This effect is shown to increase further as the length of the polymer chain increases.[48] MD studies by Cai et al. (2012)[49] suggested that hydrogen bonds between the cellulose molecule and an additive like urea, in a urea–water solvent mixture, increases the solubility of cellulose. The search for a more effective solvent for the dissolution of cellulose (and lignocellulose) has turned towards ionic liquids in the recent past and there has been a considerable interest in this area, leading to identification of many ionic liquids that can dissolve lignocellulose.[42,48,50,51]

2.4.2. Solvent effects on glucose and fructose conversion reactions

In order to investigate the effect of the addition of DMSO on fructose dehydration to HMF, Mushrif et al. (2012)[34] and Nikolakis et al. (2012)[35], performed molecular mechanics based MD simulations of fructose and HMF in water + DMSO mixtures. They observed that both solvents compete to be in the first solvation shell of the solute. The preferential arrangement of DMSO around HMF provides a shielding effect to the HMF molecule, which protects it from further rehydration to levulinic acid and formic acid and from humins formation. Analysis of the local 3-dimensional arrangement of DMSO molecules around fructose also suggested that DMSO protects fructose from side reactions that would lead to condensation or reversion products. However, the presence of DMSO molecules does not hamper water molecules from coming into contact with the oxygen atom of the hydroxyl groups of fructose, which is required for the proton transfer from water to fructose, to initiate the dehydration reaction to HMF.

Caratzoulas and Vlachos (2011)[52], investigated the chemistry of fructose dehydration to HMF in a condensed phase environment in which they treated solvent water molecules explicitly. They showed that proton transfer during the reaction gets accelerated if it is mediated by water and that the high barrier in the charge transfer step in the reaction could be due to the reorganization of the

polar solvent. It is thus hypothesized that the addition of DMSO, as a co-solvent to water, facilitates the solvent reorganization and thus reduces the barrier for the rate limiting step. This is also confirmed by Qian and Wei (2012)[53] when they performed CPMD-metadynamics simulation of uncatalyzed hydride transfer step in glucose isomerization. They observed that the barrier increases by almost 30 kcal/mol, when calculations are performed in condensed phase, with explicit water molecules. In the recent work done by the same group using fully condensed phase CPMD metadynamics calculations[54], they also demonstrated that the free energy barrier for glucose condensation reaction is much higher in DMSO than in water (*c.f.* Fig 2.). This suggested that DMSO, in addition to having a preferential solvation effect, also slows down the kinetics of side reactions and helps achieve better selectivity for the dehydration reaction. Recent investigations into $CrCl_3$ catalyzed glucose to fructose isomerization were performed by Mushrif et al. (2014)[39], using CPMD metadynamics, with explicit water molecules. They demonstrated that, contrary to popular belief that the hexahydrated Cr^{3+} is the active species, the partially hydrolyzed $[Cr(H_2O)_5OH]^{2+}$ complex is the active isomerization species. The protonation of the OH^- group on the metal center facilitates the rate limiting hydride transfer and ring opening steps.

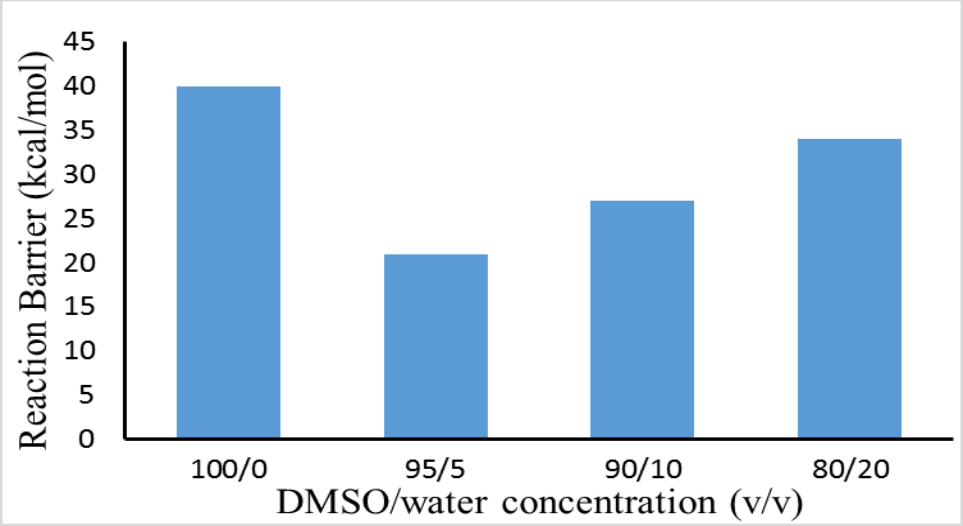


Fig.2 Reaction barrier for HMF production with change in DMSO and water concentrations in solvent media[54]

The work done by Mellmer et al[33] and by Mushrif et al[39] show that the arrangement of the solvents around glucose and the reaction environment created by the solvent molecules affect the kinetics of the conversion process and based on that the conversion yield and selectivity can be affected.

In this present work we try and answer this question of how the solvent environment changes as the water content in the system decreases from a pure water system to being mixed with co-solvents.

2.5. Key Questions that have not been answered in the existing literature so far:

a. It has previously been seen that water present near glucose is necessary for the conversion reactions and the presence of co-solvents has shown to improve reaction yields. Thus the question of how the glucose-water interaction changes as the co-solvent is added in different proportions, possibly altering the reaction outcome, has not been answered yet?
b. The possible role of physical interactions between glucose and co-solvents and amongst the solvent molecules in the system in enhancing the selectivity of towards the conversion to HMF has never been investigated.
c. The role of co-solvents in directing the reactions to specific pathways and the possible effect that they may have on the reaction outcomes. It has also to be seen how the solvent-solvent interactions and solvent-solute interactions change as the reaction proceeds.

Chapter 3: Objectives

In the present work, we investigate the solvation of glucose, a monomer of cellulose, in water, DMSO, DMF and THF, in different mixture proportions, using force-field based molecular simulations. Solvents selected in the present study have shown to increase the yield of HMF, when the reaction is catalyzed by inorganic catalysts like Cr, Sn, etc.[16,55] This is one of the most widely studied and an important reaction in biomass conversion technology, because HMF is considered as a platform chemical for the production of different chemicals that have further use as fuels, fuel additives, polymers, plasticizers, etc. Scheme 1 illustrates the pathway for the conversion of glucose to HMF and levulinic acid.[1] It has to be noted that computational work has also shown the possibility of converting glucose to HMF without forming fructose as an intermediate.[56]

The major objectives of the present study

a. Glucose has four ring hydroxyl groups as well as a ring oxygen. The protonation of the different oxygen groups will lead to different reaction pathways. Different reaction pathways lead to different products. The reaction pathway of focus in this study is the conversion of glucose to HMF which is started by the protonation of the carbon 2 hydroxyl group. Hence understanding the preferential solvation of the different groups of glucose by the different solvents namely DMSO, DMF and THF (in mixtures with water) will allow us to predict which pathway is preferred and also understand the shielding effect of the different co-solvents in enhancing the possibility of specific pathways.
b. The presence of large number of hydroxyl groups in the glucose molecule leads to the possibility of large number of hydrogen bonds between glucose and water molecules. The presence of co-solvents lead to competition for hydrogen bonding between solvent molecules and glucose. The increased competition can lead to reorientation and changes in the arrangement of water molecuels around glucose which will be studied to give us an idea of the changes in the interactions between glucose and water.
c. The changes in the local arrangement of the solvent molecules is also affected by the interactions between water and co-solvents. The interactions between the different solvent molecules leads affects the solvent environment around glucose. Studying the changes in the solvent environment can help in identifying how the co-solvents can improve the effectiveness

of the conversion mechanism.

d. The conversion mechanism of glucose to dehydration products is also affected by the interactions between glucose molecules themselves. Looking into the behaviour of glucose molecules as the solvent concentration of the solvent mixture changes can lead to some useful insights into some possible implications towards the selectivity of reaction mechanisms.
e. To identify and study some key interactions that might have an effect on the solvent behaviour and the changes in these interactions as the solvent concentrations changes.

.

Chapter 4: Computational Methods

The molecular dynamics method was first introduced by Alder and Wainwright in the late 1950's[57,58] to study the interactions of hard spheres. Many important insights concerning the behavior of simple liquids emerged from their studies. The next major advance was in 1964, when Rahman carried out the first simulation using a realistic potential for liquid argon.[59] The first molecular dynamics simulation of a realistic system was done by Rahman and Stillinger in their simulation of liquid water in 1974.[60] The first protein simulations appeared in 1977 with the simulation of the bovine pancreatic trypsin inhibitor (BPTI).[61] Today in the literature, one routinely finds molecular dynamics simulations of solvated proteins, protein-DNA complexes as well as lipid systems addressing a variety of issues including the thermodynamics of ligand binding and the folding of small proteins. The number of simulation techniques has greatly expanded; there exist now many specialized techniques for particular problems, including mixed quantum mechanical - classical simulations, that are being employed to study enzymatic reactions in the context of the full protein. Molecular dynamics simulation techniques are widely used in experimental procedures such as X-ray crystallography and NMR structure determination. There are also higher order quantum mechanical simulations that have been shown to accurately represent atoms and can be used to simulate different reactions in real time.

In this work we use a classical mechanics based simulation software to study the physical solvation interactions called GROMACS[62-65], about which we shall see later. We will also have a brief look into the OPLS/AA force fields which we use to represent the system.

4.1. GROMACS

GROMACS[62-65] (Groningen Machine for Chemical Simulations) was a molecular dynamics package developed by the Biophysical Chemistry Department of University of Groningen for simulations of proteins, lipids and nucleic acids which later branched out into different types of molecules. It is one of the fastest and most popular software packages around. It is a free, open source released under the GNU General Public License.

4.2. Molecular Dynamics

Molecular dynamics is a computer simulation of the movement of atoms and molecules under the influence of physical forces. The atoms and molecules are allowed to interact for a period of time to give a view of their motion in time. In the case of GROMACS, the trajectories are calculated using Newton's Equations of Motions with the forces and potential energies given by Forcefields about which we will discuss later. Molecular Dynamics are used as tools to study the properties of different compounds under different of conditions and to understand their behaviours in the presence of different forces and constraints.

The Molecular Mechanics used by GROMACS has some important principles to enhance the speed of calculations. These are:

1. Nuclei and electrons are clumped together into atom-like particles
2. The atom-like particles are spherical with radii derived from measurements or theory and has a charge
3. Interactions are based upon classical potentials and springs
4. The interactions must be preassigned to the specific set of atoms
5. The interactions affect the spatial distribution of the atom-like particles and their energies

4.3. Force Fields

The Molecular Dynamics are that used to study the behaviour of different molecules require a representation of the energy of the molecule as a function of the atomic coordinates. The states expected to be occupied at thermal equilibrium are the low-energy regions of this function and the forces can be derived from the gradient of this function. This is why such functions are called 'force fields'[66].

These force fields were initially developed to help simulate proteins in different systems to understand their structure and dynamics in various solvents. To improve the speed of the atomistic calculations for proteins, the average over the electronic motions are considered such that the energy surface on which the atoms move is the Born-Oppenheimer ground state energy. It is not yet feasible to calculate these energy surfaces for large macromolecules using quantum chemistry

electronic structure calculations. To calculate these most simulations use simple functions to define the various energies, adjusting a large number of parameters to optimise the systems to match macromolecular experimental data as well as quantum calculations for smaller molecules.

The design and parameterization of force fields is a complex problem involving many decisions concerning which data to use to fit for a system, transferability from the 'fit set and computational efficiency[66].

4.4. All Atom Optimised Potentials for Liquid Simulations (OPLS/AA) force field

The OPLS potentials were developed by Jorgensen and co-workers to simulate liquid state properties, initially for water and for more than 40 organic liquids. These were called OPLS (Optimized Potentials for Liquid Simulations) and placed a strong emphasis on deriving non-bonded interactions by comparison to liquid-state thermodynamics. Indeed, the earliest applications of OPLS potentials were to rigid-molecule Monte Carlo simulations of the structure and thermodynamics of liquid hydrogen fluoride.[67] The reproduction of densities and heats of vaporization provides some confidence in both the size of the molecules and in the strengths of their intermolecular interactions. These early models (now called OPLS-UA) treated hydrogens bonded to aliphatic carbons as part of an extended atom but represented all other hydrogens explicitly. The initial applications to proteins[68] (Jorgensen and Swenson, 1985; Jorgensen and Tirado-Rives, 1988; Tirado-Rives and Jorgensen, 1990) used a polar-hydrogen – only representation, taking the atom types and the valence (bond, angle, dihedral) parameters from the 1984 Amber force field. This was called the AMBER/OPLS force field, and for some time was reasonably popular. As with Amber and CHARMM, an all-atom version (OPLS-AA) was developed later, but with much the same philosophy for derivation of charges and van der Waals parameters from simulations on pure liquids.[69-71] The parameter choices were intended to be ''functional group friendly,'' so that they could be easily transferred to other molecules with similar chemical groupings. Although the parameters were principally derived with reference to condensed phase simulations, comparisons to gas-phase peptide energetics also show good results.[72]

The energy equations that represent the different potentials as defined earlier for OPLS/AA force field are given by

$$E_{bonds} = \sum K_r (r - r_0)^2$$

$$E_{angles} = \sum k_\theta (\theta - \theta_0)^2$$

$$E_{dihedrals} = \frac{V_1}{2}[1 + \cos(\varphi - \varphi_0)] + \frac{V_2}{2}[1 + \cos 2(\varphi - \varphi_0)] + \frac{V_3}{2}[1 + \cos 3(\varphi - \varphi_0)] + \frac{V_4}{2}[1 + \cos 4(\varphi - \varphi_0)]$$

$$E_{nonbonded} = \sum f_{ij} \left({A_{ij}}/{r_{ij}^{12}} - {C_{ij}}/{r_{ij}^{6}} + {q_i q_j e^2}/{4\pi\varepsilon_0 r_{ij}} \right)$$

4.5. Analytical techniques used for analysis of MD simulations:

4.5.1. Radial Pair Distribution Function

It is a measure of the probability of finding a particle at a distance of r away from a given reference particle, relative to that for a bulk system. The RDF is usually determined by calculating the distance between all particle pairs and binning them into a histogram.

We identify the first and second solvation shells from the radial pair distribution function g(r). The first co-ordination shell is defined as a spherical shell with radius bound between r_0 and r_1 around the central particle. Where r_0 is the first position where g(r) is no longer zero and r_1 is the location of the first minima indicated in the figure below. Similarly the second coordination shell is defined as the spherical shell with radius bound between r_1 and r_2 where r_2 is the location of the second minima. The First coordination shell is the volume which is directly influenced by the central particle, while the influence of the particle decreases as the order of the solvation shells increases due to shielding from the previous shells and larger distance between the shell and the central particle.

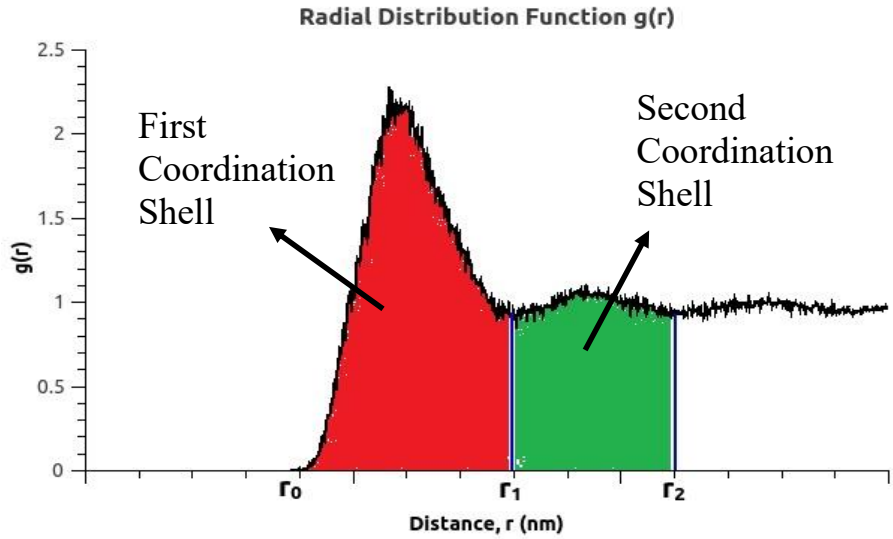


Fig.3 Radial pair distribution function curve where the first coordination shell between r_0 and r_1 is shown in red and second coordination shell between r_1 and r_2 is shown in green

4.5.2. Volumetric Density map

The volumetric density map is generated by mapping the weighted atomic density at each grid point in a three dimensional grid. This is done by replacing each atom in the selected space with a normalized Gaussian distribution of width (standard deviation) equal to its atomic radius. The Gaussian distribution for each atom is then weighted using an optional weight. The calculations done in this work assume a default weight of one (i.e, the number density). The various gaussians are then additively distributed on a grid and the different volumes are averaged out for the entire trajectory.

4.5.3 Hydrogen Bonding Analysis

The hydrogen bond lifetimes are calculated from the auto-correlation function derived from the simulation trajectory. The methodology used for the calculations use the luzar and chandler approach[73].

The instantaneous configuration of the system is defined by the function r(t) which denotes the position of all atoms at time t. Based on the configuration of the molecules the criteria for a hydrogen population in the system can be defined using the hydrogen bond population operator h[r(t)] = h(t). This operator has value 1 when two given molecules are hydrogen bonded and 0

when they are not. The average of the hydrogen bond population operator is <h>.

The auto-correlation function is defined as the probability that a hydrogen bond is intact at time t given that the hydrogen bond was intact at time zero.

$$c(t) = < h(0)h(t) > / < h >$$

The auto-correlation function is used to characterize the fluctuations in the hydrogen bond population lifetimes. This is computed from the simulation trajectory by recording the occurrence of two non-zero populations separated by time t.

The kinetics of the hydrogen bond breakage and reformation are defined by the forward rate constant for hydrogen bond breakage k and the reverse rate constant k' for hydrogen bond formation both of which follow first order kinetics. The values for these rate constants can be derived from the reverse flux correlation function K(t). This is also defined as the rate of relaxation to equilibrium given that the initial system has a non-equilibrium concentration of hydrogen bonds. Since at equilibrium, the probability that a given pair of hydrogen bonds will still be bonded in a very large system is negligibly small, the value of c(t) will tend to zero and and the reverse flux correlation can be calculated from the rate of relaxation of c(t)

$$K(t) = -dc(t)/dt = - < \dot{h}(0)[1 - h(t)] > / < h >$$

The rate of relaxation can be written in terms of the hydrogen bond forming and breaking rates .

$$K(t) = -kc(t) + k'n(t)$$

Here n(t) is defined as the probability that two molecules are initially hydrogen bonded but are unbonded at time t even though they are within the cut-off distance. The hydrogen bond cut-off distance is taken as 3.5 A. The two probabilities c(t) and n(t) correspond to local populations that can break and form hydrogen bonds in the first solvation shell. From equations (2) and (3) we get

$$dc(t)/dt = kc(t) - k'n(t)$$

The values of k and k' can be identified by the selecting the values for them so that the graph K(t) vs kc(t)-k'n(t) has a unit slope. Only one such combination for the two rate constants will give this result. Thus the k and k' that satisfy this criteria can be chosen as the rate constants for hydrogen bond breaking and hydrogen bond formation respectively.

The rate of hydrogen bond breaking can be calculated by getting 1/k. The rate of hydrogen bond breaking derived from this is the average lifetime of the hydrogen bond.

$$\tau_{HB} = 1/k$$

If the process of hydrogen bond breakage can be assumed to be an Eyring process, then

$$\tau_{HB} = \frac{h}{k_B T} \exp(\frac{\Delta G}{k_B T})$$

Where h is the Planck's constant, kB is the Boltzmann's constant and T is the temperature ΔG is the Gibbs free energy of activation. The activation enthalpy of hydrogen bonding can then be calculated using the van't Hoff equation

$$\Delta H = \frac{\delta(\Delta G/T)}{\delta(1/T)}$$

4.6. System Description

Molecular dynamics (MD) simulations were performed using GROMACS[62-65] for pure and mixed solvents systems. In the case of mixed solvents systems, THF, DMSO and DMF were mixed with water in different proportions. Number of glucose and solvent molecules in the simulation systems and simulation box sizes are given in Table 1. Simulations were performed using the OPLS/AA force-field parameters[68,69]that were chosen and benchmarked against the published literature.[37,74,75]

For solvent molecules, macroscopic properties such as density and dielectric constant were also predicted and compared to experimental values. We would like to point out that applying a partial negative charge to the nitrogen atom on DMF, as suggested in the literature,[37] resulted in its reduced miscibility with water. Hence, DMF molecule was reparametrized and has been discussed in detail in this report. The TIP4P water model[76] was used to model water molecules in all the simulations. Simulation parameters for all the species are provided in Table 4. Each system was simulated for 4 ns in an NVT ensemble with an integration time step of 0.001 ps. This was preceded by a 100 ps of NPT simulation to get the correct density

The coordinates, forces, velocities and energies were recorded for every 0.1 ps along the MD trajectory. The output was visualized using VMD[77] and the analysis of MD trajectories was performed using standard tools within VMD and GROMACS. All the simulation systems were maintained at 298 K using the Nosé-Hoover thermostat.[78,79] A time constant of 0.2 ps was applied for the temperature coupling. No constrains on bonds and angles were applied. The cut-off distance of 1.2 nm was used for Lennard-Jones potential. The Coulomb potential was calculated using Particle Mesh Ewald[80] with a cut-off of 1.3 nm and Fourier grid spacing of 0.12. The neighbor list was updated every 0.01 ps within 1.3 nm. Periodic boundary conditions were applied in all directions. The hydrogen bond lifetimes between glucose and solvent molecules were calculated from the auto-correlation function derived from the simulation trajectory.

Ab initio MD simulations were performed using the CPMD package, version 3.15.1.[81]. It implements the Car–Parrinello scheme for *ab initio* MD calculations.[82] The first–principles calculations were performed using the planewave–pseudopotential implementation of Kohn–Sham density functional theory.[83] The Troullier–Martins pseudopotential[84] with the Perdew–Burke–Ernzerhof generalized gradient approximation[85] was used for all the atoms in the simulation system. Only the Γ-point was used for integration over the Brillouin zone in the reciprocal space. An energy cut–off of 90 Ryd. was used. Nosé-Hoover chain thermostat was used for controlling ionic and electronic temperatures. The frequency for the ionic thermostat was set to 1800 cm-1 and for the electron thermostat to 10000 cm-1. The fictitious electron mass parameter in CPMD was set to 600 a.u. Short molecular dynamics runs without the thermostat were performed to obtain an

approximate value about which the fictitious electronic kinetic energy oscillates. Based on the observations from these runs, a value of 0.07 a.u. was chosen for the electronic kinetic energy. Timestep of a single *ab initio* MD step was set to 0.0964 fs. Geometry optimization was performed on the system before molecular dynamics and the system was equilibrated for 1 ps. Energies, including the fictitious electronic kinetic energy, were monitored throughout the simulation (10 ps) to ascertain that the system did not deviate from the Born–Oppenheimer surface during the molecular dynamics simulation. Given the extremely high computational cost of CPMD simulations, a smaller simulation cell of 1.457 nm was used.

Table 1: Compositions of simulation systems and simulation box sizes for model systems studied in this work

Solvent System	Water Concentration in solvent (in wt%)	Number of glucose molecules	Number of Other solvent molecules	Number of water molecules	Cubic Box side length (nm)
Pure Water	**100**	**3**	**0**	**2984**	**4.47**
Pure DMSO	**0**	**3**	**865**	**0**	**4.69**
Pure DMF	**0**	**3**	**709**	**0**	**4.49**
Pure THF	**0**	**3**	**1003**	**0**	**5.23**
DMSO-Water Mixture	**90**	**3**	**60**	**1900**	**4**
	50	**3**	**300**	**1169**	**4.13**
	10	**3**	**530**	**270**	**4.13**
DMF-Water Mixture	**90**	**3**	**67**	**2430**	**4.32**
	50	**3**	**333**	**1350**	**4.31**
	10	**3**	**600**	**270**	**4.33**
THF-Water Mixture	**90**	**3**	**68**	**2430**	**4.33**
	50	**3**	**338**	**1350**	**4.37**
	10	**3**	**538**	**240**	**4.28**

Chapter 5: Benchmarking of force-field parameters for use in simulations

In the first section of the results and discussions section we shall discuss benchmarking of the chosen force-field parameters to ensure the accuracy of the predictions of our simulations. We shall then optimize the parameters for DMF to better represent the behaviour of DMF in a mixed solvent system. Following these we shall look into the solvation of glucose in pure and mixed solvent systems and analyze the effect of solvent molecules on other solvents and on glucose and also look into possible implications of these interactions on the conversion reaction.

5.1.1. THF Benchmarking

The Radial pair density functions for Tetrahydrofuran were calculated using GROMACS for a system that had been run at 298 K for a box of 3nm having 173 molecules of Tetrahydrofuran (THF). The system was initially run for 2100 energy minimization steps followed by 100000(100 ps) NPT steps using Berendsen thermostat to optimize the system. Then the NVT run was done for 2000000 (2ns) to generate a trajectory from which the radial pair distribution functions were calculated.

The RDFs were calculated between the centre of mass of the THF molecules using the g rdf command in GROMACS. The command allows GROMACS to calculate the co-ordinates of the centre of mass of the THF molecules after which it calculates the radial pair distribution functions assuming these points to represent the whole molecule. Figure 4 shows that the calculated RDFs are very similar to the ones available in literature[86] with a slight shift in the peak for the first solvation shell which can be accounted for due to the slight difference in density of the system. The actual density of THF is around 880 kg/m3 at 298 K while the density of the system that was used in the simulation was 872 kg/m3.

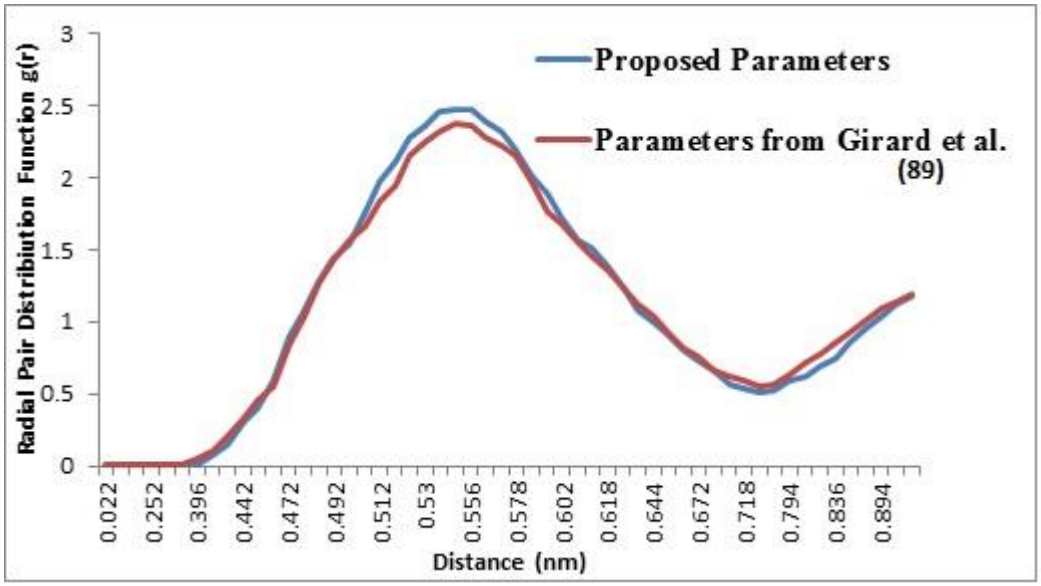


Figure 4. Radial pair distribution functions between the Centre of Mass of the bulk Tetrahydrofuran molecules.

5.1.2. DMSO Benchmarking

The Radial pair density functions for Dimethylsulfoxide were calculated using VMD for a system that had been run at 298 K for a box of 2nm having 56 molecules of Dimethylsulfoxide (DMSO). The system was initally run for 1900 energy minimiztion steps followed by 100000(100 ps) NPT steps using Berendsen thermostat to optimize the system. Then the NVT run was done for 2000000 (2ns) to generate a trajectory from which the radial pair distribution functions were calculated. The intermolecular RDFs were calculated between the Oxygens of the different DMSO molecues and also the Sulphur and oxygens of the different molecules. The calculated values and the literature[87] values are very similar and show a very good overlap as can be seen from Figure 5.

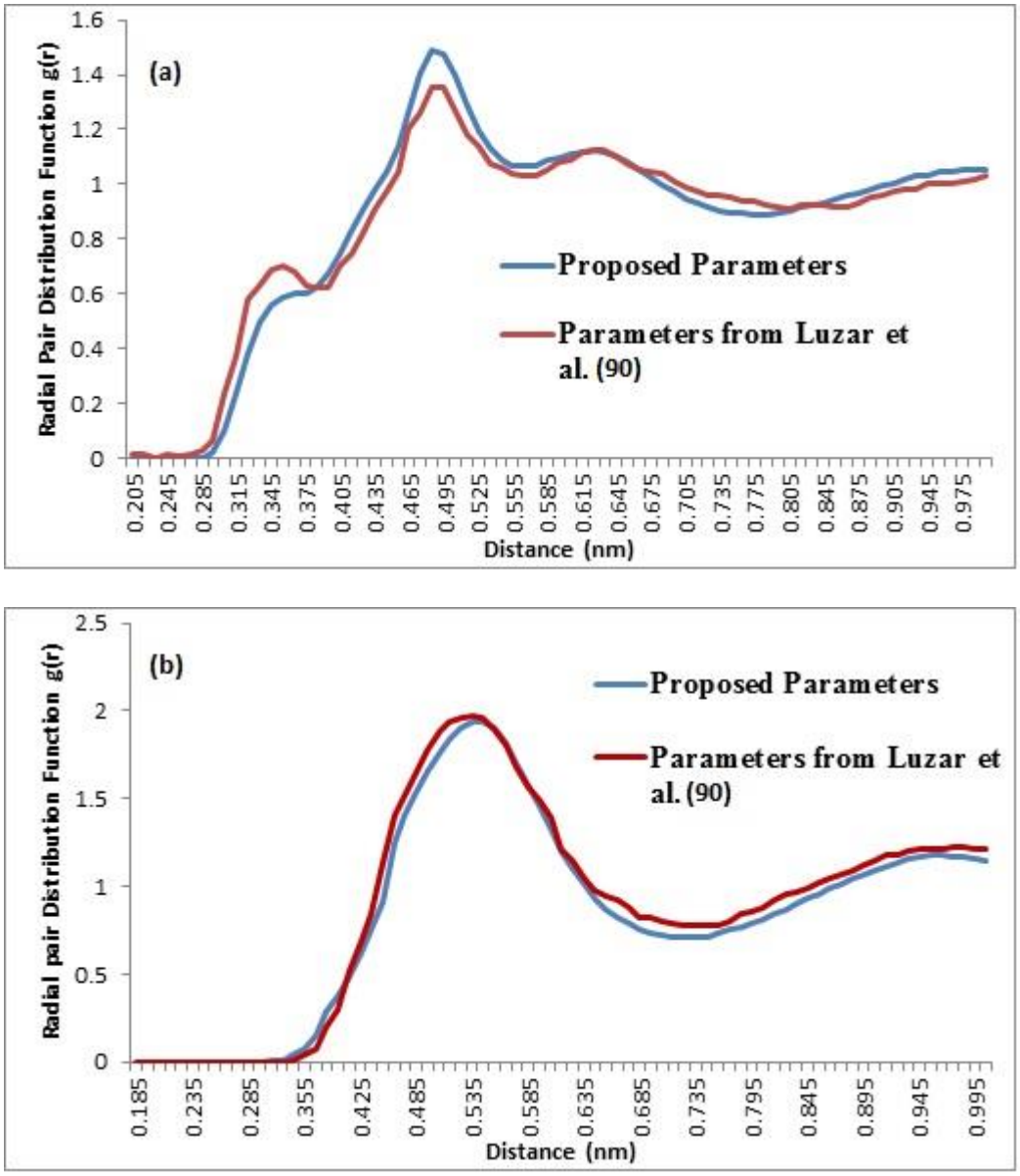


Figure 5: Intermolecular Radial pair distribution functions between a) oxygens from both literature and the calculations and the b) Sulphur and Oxygen

5.1.3. Glucose Parameters

The RDFs were calculated between the various Oxygen atoms on the Glucose molecule and the water oxygen using VMD for a system of box size 4 nm with 2217 TIP4P water molecules and 1 Glucose molecule at 298 K and 1 atm pressure. The system was initially run for 3700 energy minimization steps followed by100000 (100 ps) NPT steps for equilibration using a Berendsen thermostat. The NVT run was done for 4000000 (4 ns) steps at 298 K using a Nose-Hoover

thermostat to generate a trajectory from which the radial pair distribution functions were calculated. The intermolecular pair distribution functions as shown in Figure 4 show that the basic structure between both the calculated values as well as the literature values[88] has a lot of similarities, especially for the rdfs between Oxygen 1 and Water oxygen.

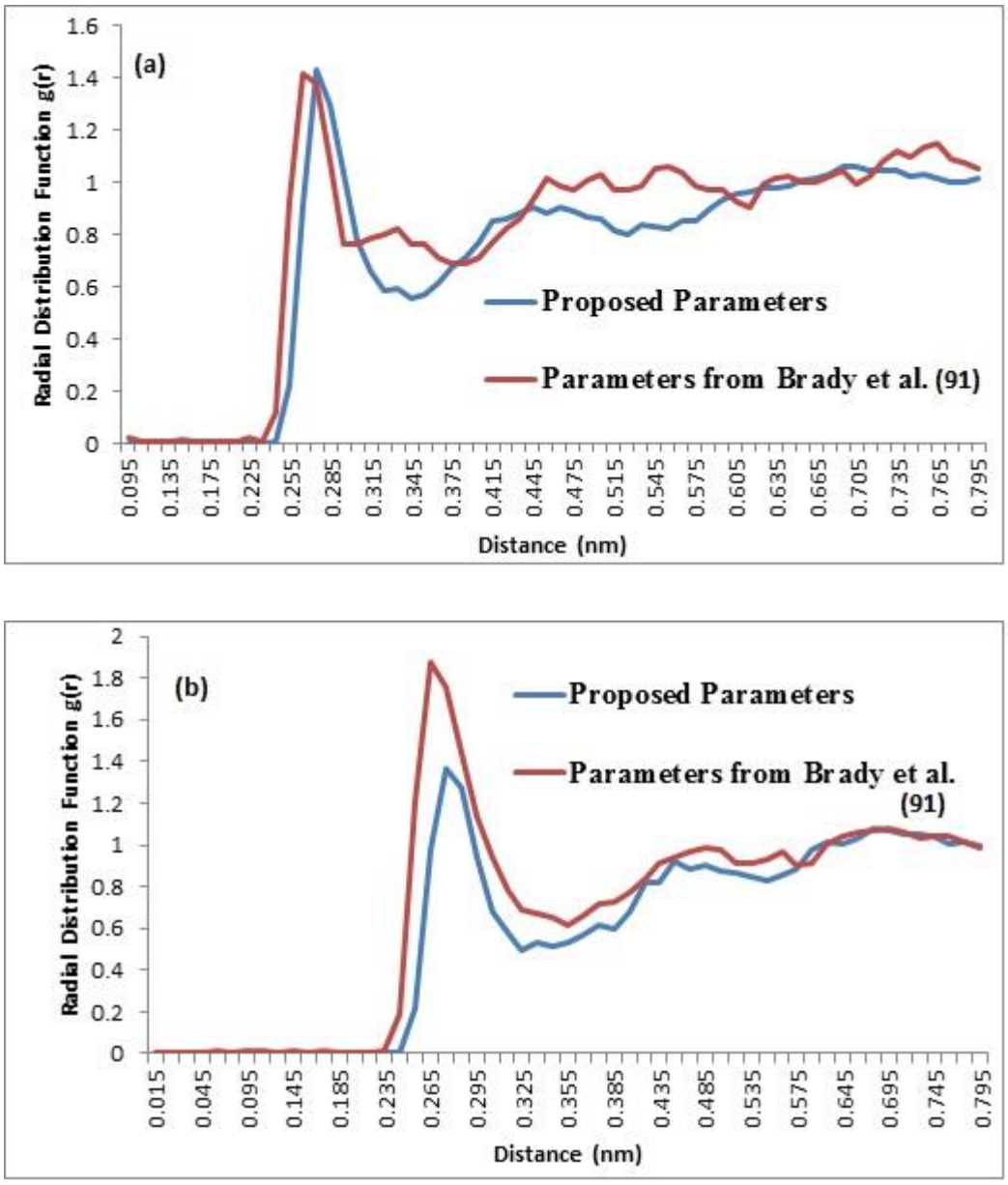
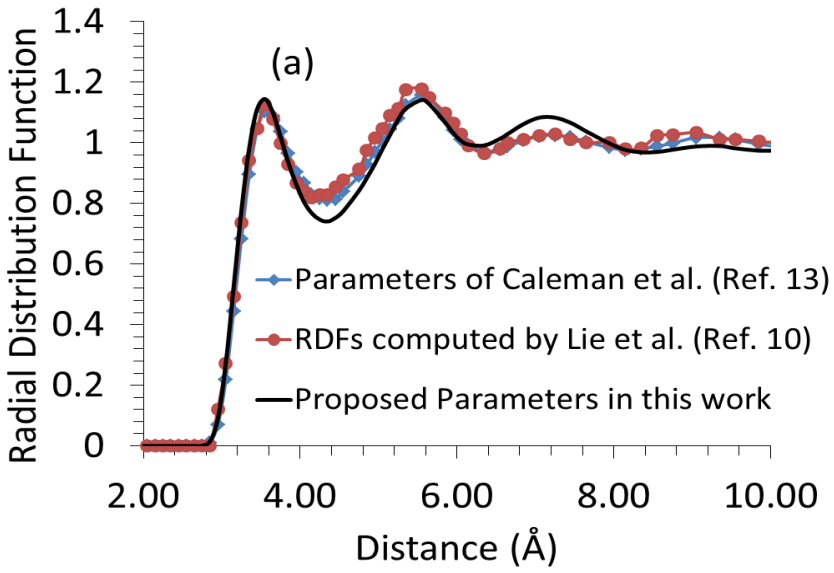


Figure 6: Intermolecular Radial pair distribution functions between a) oxygen attached to 1st carbon and water oxygen from both literature and the calculations and the b) oxygen attached to 3rd carbon and water oxygen

5.2. Optimizing DMF parameters

5.2.1. Pure DMF system. Bulk Density is a projection of the molecular level arrangement and interactions of any species on a macroscopic scale. The experimental bulk density of DMF is 944 kg/m3. Comparison of bulk densities predicted by the proposed set of force-filed parameters in this paper and by the parameters in the literature[37] is given in Table 3. The error in the predicted density using the parameters by Caleman et al.[37] was found to be 2.75%; whereas it was 1.45% when the proposed parameters were used.

The molecular structure of DMF was studied using the radial pair distribution functions (RDFs). RDFs were calculated between OA and HA and CA of DMF molecules (refer to Table 1 for nomenclature). RDFs computed using the proposed parameters were compared with the parameters by Caleman et al.[37] and with RDFs calculated by Lei et al.[89] As can be seen in Fig. 7, it was observed that the atom-atom RDFs do not show any significant difference. This is not surprising because, the change in the predicted bulk density is not large enough to show significantly different RDFs.



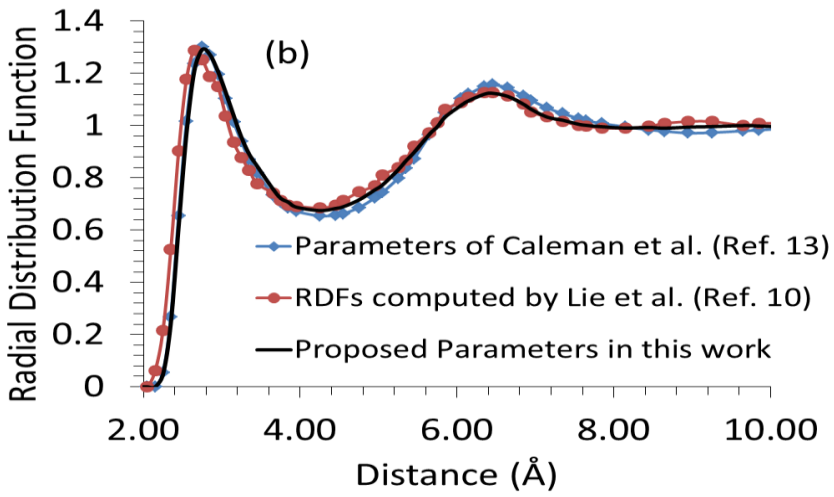


Figure 7. Atom-atom Radial Pair Distribution Functions calculated for Pure DMF systems between a) Acyl Oxygen (O_A) - Acyl Carbon (C_A) and b) Acyl Oxygen (O_A) – Acyl Hydrogen (H_A) using parameters from Caleman et al.[37] (blue), proposed parameters in this paper (black). RDFs computed by Lei et al.[89] are also shown in red.

The comparison of predicted dielectric constants is also shown in Table 2. It can be seen that the dielectric constant of DMF, predicted using the parameters suggested by Caleman et al.[37] has an error of ~50%. However, the proposed parameters show a marked improvement in the prediction and result in a reduced error of ~20%, as compared to the experimental value.[90]

Table 2.Comparison of simulation predicted Density and Dielectric constant of DMF with experimental values

Property	**Computed using Parameters from Caleman et al[37]**	**Computed using Proposed Parameters**	**Experimental[90]**	**Values reported by Caleman et al[37]**
Density (kg/m^3)	918	930	944	922
Dielectric Constant	15.68	30	38	15

The slight difference in the values of bulk density and dielectric constants predicted by Caleman et al in their paper[37] and the values predicted in the present work using the same parameters can be attributed to the innate characteristics of the different molecular dynamics algorithms and slight differences in the cut-off and other run parameters.

It has to be noted that prior to assigning a partial positive charge to the nitrogen atom in the proposed set of parameters, multiple empirical guesses were made to the charge and Lennard Jones parameters. The partial charge on the nitrogen atom was also changed between -0.14 to -0.57. The charges on carbon and hydrogen atoms were untouched due to the relatively higher electronegativity of Nitrogen and oxygen, and thus it was assumed that any major changes in the interactions were going to be affected by changes in the parameters of these two atoms. For different sets of parameters, the density and the dielectric constant were calculated first and were compared to the experimental data. In all the attempts, it was found that predicted densities were lower than the experimental value by 40 to 50 kg/m^3, and all the predicted dielectric constants were in the order of 12 to 15 (compared to an experimental value of 38). Hence, placing a higher negative partial charge on nitrogen did not show any improvement in predicting DMF properties. Therefore, a positive partial charge was assigned to the nitrogen atom.

5.2.2. DMF-Water intermixing. Figures 6a and 6b show snapshots of the DMF-water (50 wt%) system at the end of the 3 ns run, simulated using parameters of Caleman et al[37] and the proposed parameters in this paper, respectively. The phase separation of DMF and water can be clearly seen for the system simulated with a partial negative charge on the nitrogen atom (cf Fig 8a). Whereas, molecular level mixing can be seen for the system simulated using the proposed parameters, with a partial positive charge on the nitrogen atom (cf Fig. 8b). Visualization of the trajectory revealed that the phase separation shown in Fig 6a, had begun as early as 1 ns of the simulation time. Since DMF is expected to be completely miscible with water, the separation can be explained by the lower strength of interactions between DMF and water molecules in the simulation system. The DMF-DMF interactions, which are usually van der Waals interactions or dipole interactions, due to the weak hydrogen bonding capabilities of DMF, would have been much stronger than the hydrogen bond strength between water and DMF.

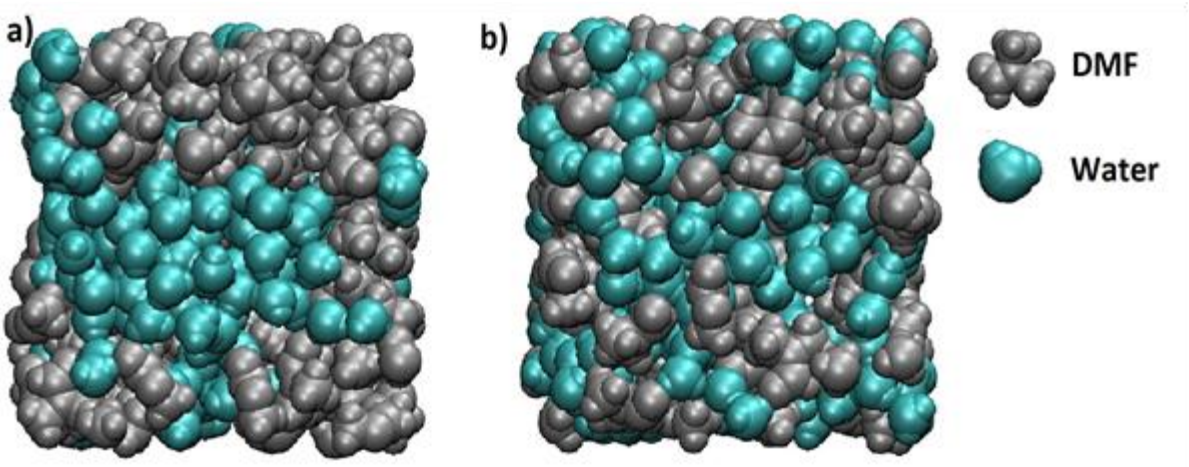
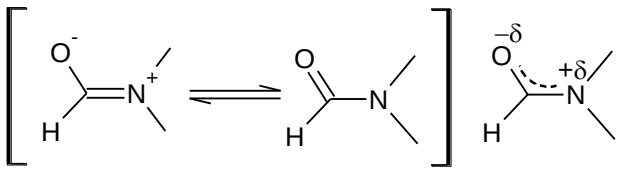


Figure 8. Snapshots of MD trajectories at 3ns for DMF-water systems. a) Separation of Water and DMF layers when the system was simulated with a partial negative charge on nitrogen, as suggested by Caleman et al[37] and b) molecular level mixing of water and DMF when the proposed parameters in this paper were used, with a partial positive charge on nitrogen.

5.2.3. Discussion

5.2.3.1. Why a positive partial charge on nitrogen? DMF is believed to exhibit a couple of resonance structures, as shown in Scheme 2.[91] In one of the structures, the nitrogen atom has a partial positive charge on it and a relatively higher partial negative charge on the acyl oxygen. The carbon-oxygen and carbon nitrogen bond distances were calculated for liquid DMF by X-ray diffraction studies and the C=O bond length was found to be 0.02 nm longer than the usual C=O bond, and the C-N bond length was found to be shorter by 0.04 nm than the usual C-N single bond[91].



Scheme 2. Predicted Resonance Structures of DMF using X-ray diffraction[91]

Additionally, alkyl carbon is believed to be capable of acting as an acceptor in the formation of a reverse dative bond with more electronegative elements like nitrogen and oxygen.[92] The formation

of the reverse dative bond transfers electron density from the element to carbon, and thus, changes (rather reverses) the charge on the carbon atom and on the atom in contact with this carbon atom in the molecule. Though this theory is inconsistent with the sp3 hybridization model, it can possibly explain the higher stabilization of N-C bond in N-CH3 than in N-C2H5. The presence of an additional carbon atom affects the capacity of the carbon atom in contact with the other element to accept the electron transfer in the formation of the dative bond. Hence, along the lines of the theory of formation of an N–C reverse dative bond in mind, we applied a slight partial positive charge of +0.04 to nitrogen. However, we retained the methyl carbon's charge as -0.11. To compensate the positive charge on the nitrogen atom, the negative charge on the oxygen atom was increased to -0.68, similar to that of alcoholic oxygen.

5.2.3.2. Increased density and change in the dielectric constant of pure DMF systems. The reversal of charge on the nitrogen atom leads to an attractive interaction between the nitrogen atom and the methyl carbon, and thus brings the DMF molecules closer. The calculated RDFs, however, showed very little difference when either set of parameters were used (cf Fig 5) and this is reflected in low errors in the predicted density of DMF. It has to be noted that the strength of the **H-C=O···H-C=O** hydrogen bond is higher than the **N-CH3···N** interaction, and hence dwarfs other interactions. Therefore, the local arrangement of DMF molecules is mainly influenced by the **H-C=O···H-C=O** hydrogen bond, thus, leading to very small differences when both sets of parameters were used.

The dielectric constant is dependent on the dipole moment of the system which is affected by the charge distribution, symmetry and orientation of the molecules in the system. The change in the charge distribution coupled with the slight increase in the nitrogen-methyl carbon interactions, would affect the orientation of the molecules, leading to a considerable change in the dipole moment of the system. This change is reflected in the improved prediction of the dielectric constant, as seen in Table 3.

5.2.3.3. Improved mixing between DMF and water. The improved mixing of water and DMF, predicted by the proposed parameters in this paper, not only results in a change in the arrangement of water and DMF molecules around each other but also affects the water-water and DMF-DMF arrangement. Hence, the RDF between oxygen atoms of water was calculated using the proposed

parameters (positively charged nitrogen) and was compared to that calculated using parameters by Caleman et al (negatively charged nitrogen),[37] as shown in Fig 5. The location of the first solvation shell in both cases are the same, but the intensity of the peak representing the first solvation shell was lower for a positively charged nitrogen. The same trend follows for the second solvation shell and the difference in the intensity is more pronounced. Significantly higher water-water coordination numbers for the RDFs calculated using a negatively charged nitrogen reveal higher localization of water molecules around each other. This suggests that when the proposed parameters, with a partial positive charge on the nitrogen atom, are used there is lesser water, and hence more DMF molecules around a water molecule, as is expected when the mixing between DMF and water improves.

As mentioned before, the DMF-water mixing was also studied using CPMD and the water oxygen RDF was calculated and compared to the force-field based molecular dynamics results, as shown in Fig 9. The co-ordination numbers for the first and second solvation shells computed using CPMD were found to be in excellent agreement with the values calculated with the proposed parameters, with a positively charged nitrogen. This further validates the enhancement of mixing that is observed when the new parameters were used. Additionally, the CPMD calculated RDFs are in a much better agreement with the RDFs calculated from the proposed parameters, as can be seen in Fig. 9.

The 3–dimensional DMF-water arrangement was also visualized by calculating a volumetric density map of water and DMF around a random DMF molecule in the system, and is shown in Fig 10. The 3-D volumetric map is averaged over the entire simulation timescale. The volumetric map of system, calculated for the parameters by Caleman et al (negatively charged nitrogen)[37] (Fig 10a), shows that there is a large presence of DMF around itself while the probability density of water is very small. This was also confirmed by the higher DMF-DMF co-ordination number and lower DMF-water co-ordination number in DMF-DMF RDFs and DMF-water RDFs. This 3-D arrangement of water and DMF around a DMF molecule is expected when there is no proper intermixing between two species and DMF is usually surrounded by DMF only. In comparison, for the system simulated using proposed parameters (positively charged nitrogen) (Fig. 10b), the probability density of DMF and water is almost equivalent which is an expected arrangement for a system where the two species are completely mixed.

The DMF-water hydrogen bonding analysis was also performed using the Luzar and Chandler[73] method. The rate of breaking of hydrogen bonds and the lifetime of hydrogen bonds were calculated. Furthermore, the free energy of hydrogen bonds was also computed using the hydrogen bond auto correlation function. The results are given in Table 3. It can be seen that the rate of breaking of water-DMF hydrogen bonds is smaller and the average lifetime of hydrogen bonds is higher for the proposed parameters with positively charged nitrogen. The hydrogen bond free energy is also greater. From these results it can be said that the DMF and water molecules would tend to stay attached for longer periods and more strongly when the proposed parameters in this paper are used, thus reducing the chances of phase separation.

Table 3. Forward Hydrogen bond Lifetime and Hydrogen bond free energy

Non-bonded Parameters	**Rate (1/ps)**	**Time (ps)**	**ΔG (kJ/mol)**
From Caleman et al[37]	0.044	22.668	12.265
Proposed	0.010	99.456	15.930

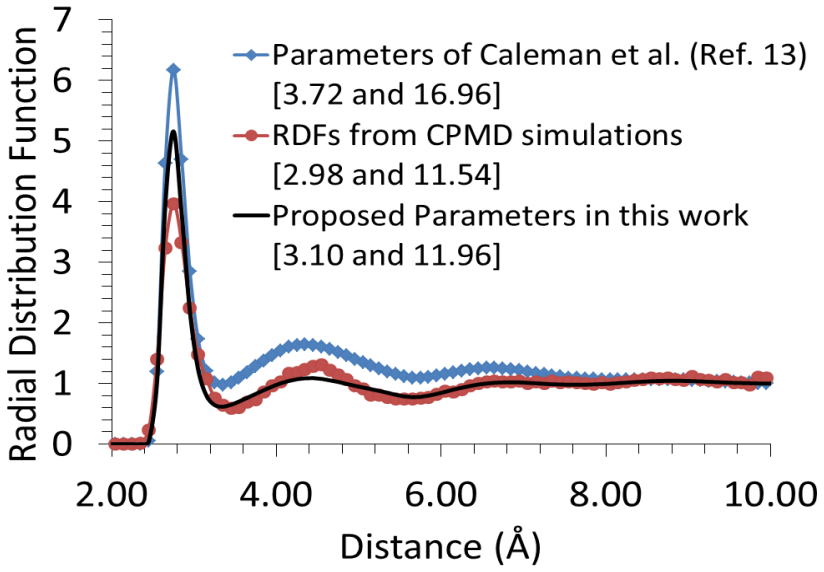


Figure 9. Radial Pair Distribution Function between oxygen atoms of water, computed using proposed parameters with a positively charged nitrogen (black), parameters by Caleman et al[37] with a negatively charged nitrogen (blue) and from CPMD calculations (red). Co-ordination Numbers of 1st and 2nd solvation shells are given in the brackets in the respective order.

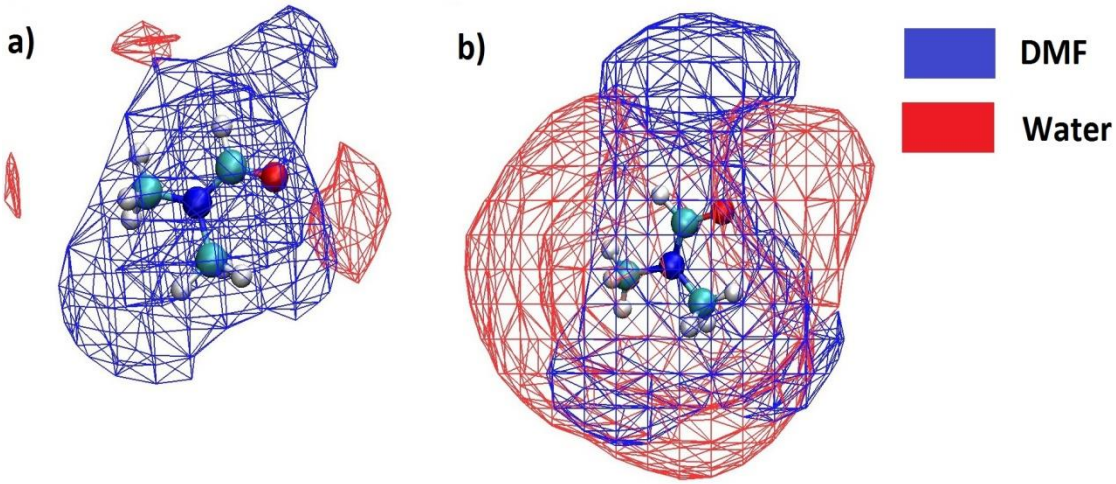
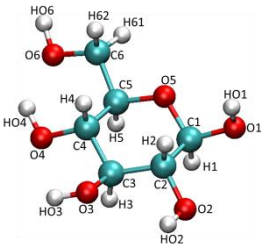
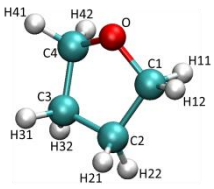
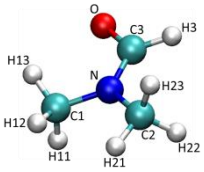
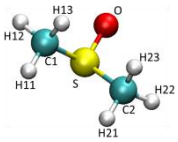


Figure 10. Volumetric Density Map of DMF (blue) and Water (red) around a given DMF molecule for a) parameters from Caleman et al[37] and b) proposed parameters, at an isovalue of 0.1

5.3. Final selected parameters for the different molecules in the study

Table 4: Non-bonded force-field parameters for glucose and co-solvents

Molecule	Atom Type	q€	σ (nm)	ε (kJ/mol)
Glucose 	C1/C5	0.26	0.35	0.28
	C2/C3/C4/C6	0.2	0.35	0.28
	O1/O2	-0.69	0.31	0.71
	O3/O4/O6	-0.68	0.31	0.71
	O5	-0.4	0.296	0.59
	HO1/H02/H03/ H04/HO6	0.42	0	0
	H1/H2/H3/H4/ H5/ H61/H62	0.06	0.25	0.126
Tetrahydrofuran (THF) 	C1/C4	0.17	0.35	0.28
	C2/C3	0	0.35	0.28
	O	-0.4	0.296	0.59
	H11/H12/H41/ H42	0.02	0.250	0.126
	H21/H22/H31/ H32	0	0.250	0.126
	C1/C2	-0.11	0.350	0.276
	C3	0.50	0.375	0.439
N,N – Dimethylformamide (DMF) 	O	-0.68	0.296	0.878
	H11/H12/H13 H21/H22/H23	0.00	0.25	0.126
	H3	0.06	0.250	0.126
	N	0.04	0.325	0.71
	C1/C2	-0.02	0.35	0.28
Dimethylsulfoxide (DMSO) 	H11/H12/H13/ H21/H22/H23	0.06	0.25	0.126
	O	-0.46	0.29	1.17
	S	0.14	0.36	1.65

Chapter 6: Solvation of glucose in pure and mixed solvent systems

6.1. Study of Glucose Solvation in selected solvents

In this section, we analyze the solvation of glucose in pure solvents (as a reference) and in mixed solvents systems. Since, the solvent medium in which the reactions in scheme 1 occur has an effect on the conversion and on the selectivity, we investigate the local organization and bonding of solvent molecules in the immediate vicinity of glucose by calculating radial distribution functions (RDF) between solute–solvent pairs. RDFs, however, may not be able to differentiate the high and low probability solvent regions which are at the same distance from a solute molecule and the local structure gets hidden in the computed radial averaging over the angular coordinates. Hence, we present and discuss the three-dimensional arrangement of solvent around the solutes by constructing volumetric maps of the time averaged solvent densities. Additionally, we also analyze the hydrogen bonding strengths and lifetimes and the mobility of glucose molecules in different solvent mixtures since that could have an effect on glucose condensation and polymerization reactions.

6.1.1 Glucose in pure solvents

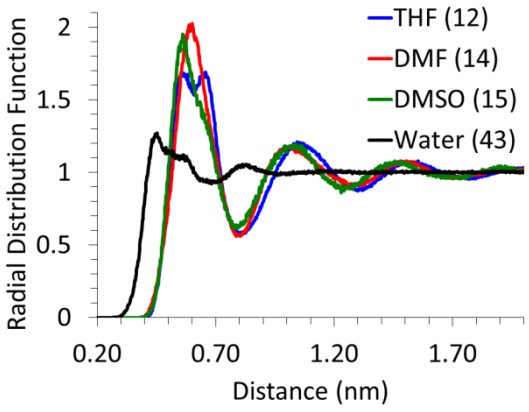
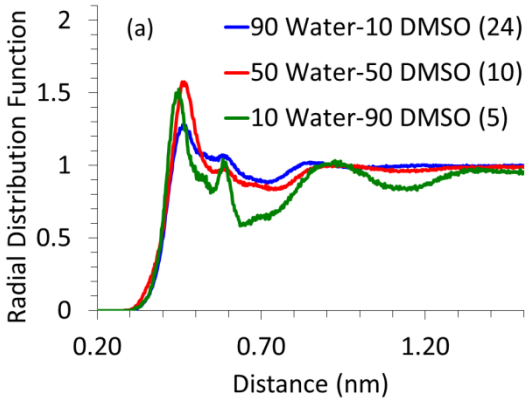


Figure 11: Center of mass radial pair distribution functions of glucose with water (black), DMSO (red), THF (green), and DMF (blue). The numbers in parentheses are first solvation shell coordination numbers.

The local arrangement of the solvent molecules around glucose was analyzed by calculating the center of mass (of the molecule) RDFs, as shown in Fig 11. The first solvation peak for water is at 0.41 nm and that for other solvents is at ~0.57 nm. Given the significantly smaller size of water molecules, as compared to other solvents, the number of water molecules in the first solvation shell is much higher than that of other solvents. The absence of an intense first solvation shell peak for glucose and water RDF can be attributed to the point that the interactions between hydroxyl groups of glucose and water would be similar to those amongst water molecules. However, the RDF curves between glucose and DMF, DMSO and THF exhibit an intense first solvation peak, indicating that the co-solvents' molecules are more densely packed around glucose than those in the bulk solvent system. This indicates that, in the mixed solvents systems, DMSO, DMF and THF molecules would compete with water to be in the first solvation shell. Analysis of MD simulations of glucose in mixed solvents' systems is in the following section.

6.1.2 Glucose in mixed solvents' system



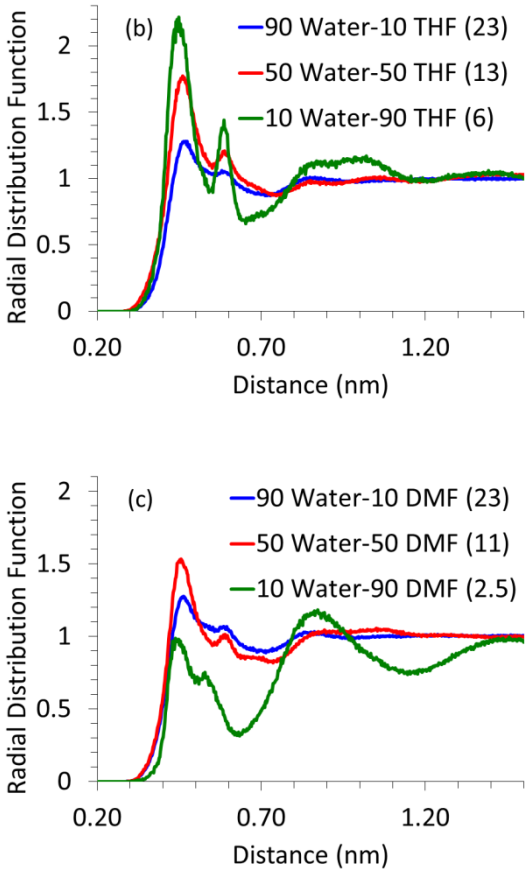


Figure 12: Center of mass Radial pair distribution function of glucose with water in a) DMSO-water-glucose system; b) THF-water-glucose system; c) DMF-water-glucose system. The numbers in parentheses are first solvation shell coordination numbers.

As mentioned before, simulations were also performed for systems containing glucose in aqueous mixtures of DMSO, DMF and THF. Since the dehydration of glucose, leading to the formation of HMF, and the undesired condensation reactions, leading to the formation of humins, are acid catalyzed reactions[93-96], the presence of water in the first solvation shell of glucose is crucial. It has been shown that excessive water is detrimental towards the selectivity of desired products like HMF and levulinic acid and it enhances the formation of undesired products like humins.[97,98] Thus, the local arrangement of the water around glucose, in the mixed solvents systems, was analyzed using the RDF of water around glucose and is shown in Fig 10. It can be seen for all the solvents that as the proportion of water decreases and that of the co-solvent increases, the glucose–water RDF peak height increases. The number of water molecules in the first solvation shell however goes down.

This suggests that all the co-solvents compete with water to be in the first solvation shell of glucose. In the absence of a co-solvent, there are 43 water molecules in the first solvation shell (*cf* Fig. 9); however, even with the addition of 10 wt% of a co-solvent, almost half of the water molecules are replaced by the co-solvent. Though the amount of water in the first solvation shell decreases with an increase in the proportion of the co-solvent, the coordination between glucose and water becomes stronger. Even when 90% of water is replaced by the co-solvent, few (2 to 6) water molecules are left in the first solvation shell of glucose and those water molecules are coordinated very strongly. It has been shown in the literature[35,99] that excessive water leads to the formation of undesired side products and very little (less than 10%) water is required to obtain high yields of HMF. Our findings suggest that the small amount of water added to the reaction mixture strongly coordinates with glucose and hence could be responsible for its dehydration to HMF. Another significant aspect to be noted in Fig.10 is the presence of a prominent second solvation shell for glucose-water RDF for high co-solvent proportions. When DMF is used as a co-solvent, the first solvation shell loses its significance at high DMF proportions as the competition from DMF for interactions with glucose pushes water molecules away to the second solvation shell. The location of the second solvation shell is very close to the location of the peak of the first solvation shell of the glucose-co-solvent RDF curve. Hence, we believe that the second solvation shell arises due to the interaction between the co-solvent and water.

6.1.3 Atomic details of the positioning of solvent around glucose

Experimental and molecular dynamics studies have shown that the reaction output for glucose depends on the location at which the initial protonation occurs.[34,95,96,98,100-104] They also suggested that the selectivities of different reaction pathways are very similar. Feng et al. showed that there is an equal probability for the protonation of different oxygens in the glucose molecule.[100] Thus, preferential solvation or shielding of selected oxygen atoms of glucose by co-solvents can reduce the probability of certain reactions since, there is a lesser chance of the co-solvent shielded oxygen atom to get protonated. Liu et al. have used Car-Parrinello molecular dynamics to study the reactions that occur as different oxygens in glucose are protonated.[102] It has been shown that protonation of glucose at C_2 oxygen (carbon atom numbering same as shown in Table 2) leads to formation of HMF while

protonation of oxygen attached to C_3 and C_4 lead to other dehydration/rehydration products. It was suggested that protonation of C_1 did not lead to any significant dehydration products. Qian et al. have done a series of studies using Car-Parrinello molecular dynamics to study the conversion of glucose to different dehydration and condensation pathways by protonation of different oxygens.[53,54,98,103] They also proposed that it was the protonation of the C_2 oxygen which led to a dehydration reaction forming a 5-membered ring intermediate, which was a precursor to HMF. IR spectroscopy studies by Patil et al. indicated that the protonation of the C_1 oxygen in glucose leads to aldol condensation products which include unwanted humins.[95,96] Yang et al. performed DFT calculations on the different reaction pathways to confirm the different pathways generated due to protonation of different oxygen atoms.[104] Protonation of glucose is water mediated.[32,39] For both condensation and dehydration products to be formed, water molecules need to be present in the immediate vicinity of the specific oxygen atom so that proton can be transferred from water to that oxygen atom.

To identify which oxygen atoms of the glucose molecule are accessible to water and which ones are preferentially shielded by the co-solvent, the volumetric spatial density map of the solvent oxygen atoms around glucose was calculated. The spatial density maps were calculated for different co-solvent proportions to see if the change in the proportion of the co-solvent could affect the location and orientation of hydrogen bonded waters around glucose. We observed that the co-solvent molecules formed hydrogen bonds with the hydrogen atoms of the hydroxyl group of glucose. On the contrary, water molecules can form hydrogen bond with both, the oxygen and hydrogen atoms of the hydroxyl groups of the glucose molecule. Fig 11 shows the volumetric density map of solvent and water molecules around glucose for the all three co-solvents, in a mixture containing 10% by weight water. Simulation systems containing high water proportions (90%) show similarity with pure water solvated glucose, and hence not shown here. In those systems, water hydrogen bonds and solvates most of the hydroxyl groups. Since the protonation is not selective, presence of water around all hydroxyl groups can lead to a higher number of dehydration products and a reduction in yield and selectivity towards HMF. At low water proportions (10% water by weight), it can be seen that that C_1-OH group of glucose is preferentially covered by water for the DMSO-water system. The THF-water and DMF-

water systems also have little co-solvent near C_1-OH which could compete with the water molecules present. Additionally, for all systems, there is equal localization of water and co-solvent around the C_2 hydroxyl group of the glucose molecule. However, there is higher localization of co-solvent molecules in the vicinity of C_3 and C_4 hydroxyl groups of the glucose molecule for all three co-solvents. Reduced probability of finding water molecules near C_3 and C_4 hydroxyl groups suggest that, in the presence of co-solvents, there is a reduced probability of water being hydrogen bonded to these groups strongly, thus reducing the chances of these hydroxyl groups getting protonated. Combining these results with the aforementioned findings about the site specific protonation leading to a specific product makes us suggest that undesired products which are formed due to the protonation of glucose at C_3 and C_4 positions could be prevented because of the preferential solvation of these positions by co-solvent molecules.

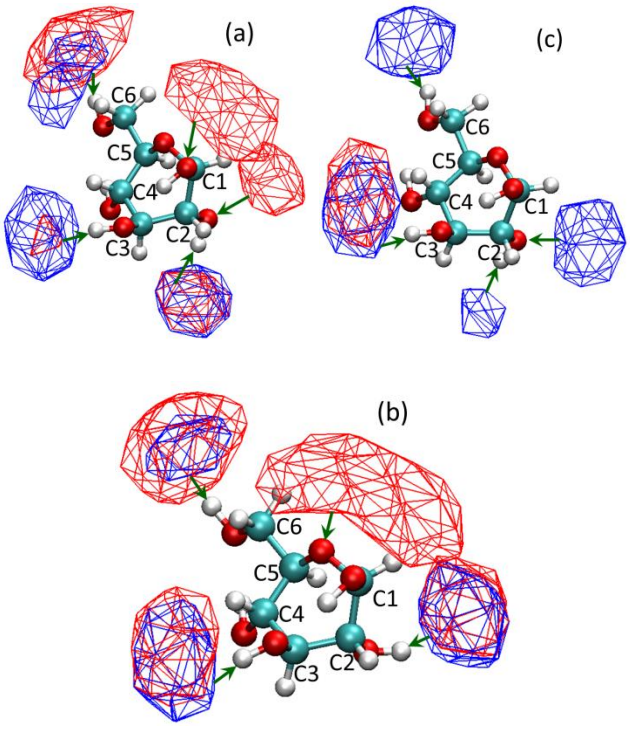


Figure 13: Volumetric spatial density maps of co–solvent (blue) and water (red) around glucose at 10% by wt. water (isovalue water:0.010; isovalue co–solvent:0.02). The large difference in isovalues is due to the large difference in solvent molecule numbers. (a) DMSO, (b) THF, and (c) DMF.

6.1.4 Mobility of Glucose molecules in solvent mixtures and hydrogen bonding analysis

The root mean square fluctuation (RMSF) of glucose molecules was calculated for all the mixed solvent systems. The initial state at time $t = 0$ in the NVT simulation was taken as the reference. The RMSF is shown in Fig 9. For all the co-solvents, it can be seen that the RMSF systematically decreases as water proportion decreases. This suggests that the mobility of glucose molecules is reduced as the water proportion decreases and co-solvent increases. Patil et al. have shown that humins are not only formed as condensation products of HMF molecules but also from condensation reactions between HMF and glucose and fructose.[95,96] For condensation products to be formed, the different components, namely glucose and HMF, or two glucose molecules have to come in the immediate vicinity of each other. Higher glucose concentration increases the visibility of one glucose molecule to the other. However, at lower glucose concentrations, the visibility of glucose molecules to each other is affected by the mobility of glucose molecules. Lower mobility would lead to lesser visibility and lower formation of unwanted condensation products. As can be inferred from Fig. 14, addition of co-solvents reduces the mobility of glucose molecules.

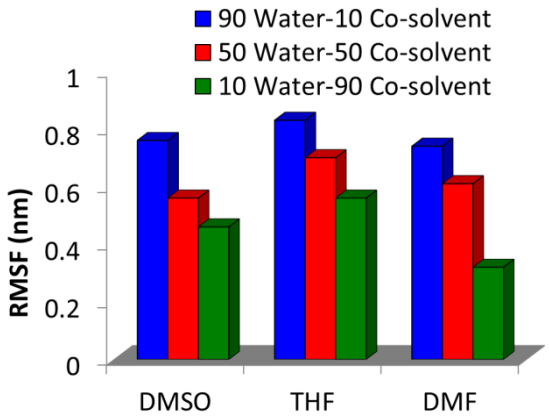


Figure 14: RMSF of two glucose molecules with respect to the third glucose molecule in a mixture of a) DMSO-water; b) THF-water; c) DMF-water at different concentrations of the solvents. (The legend specifies the different concentrations as wt. % of water in solvent mixture).

Table 5: Hydrogen bond Lifetimes in picoseconds and hydrogen bond free energies in KJ/mol. Free energies are in parentheses.

System	Hydrogen bond	Water proportion		
		90% by wt.	**50% by wt.**	**10% by wt.**
DMSO-Water Mixture	Glucose-DMSO	23.38 (12.34)	60.46 (14.7)	410.24 (19.44)
	Glucose-Water	6.67 (9.23)	13.73 (11.02)	61.77 (14.75)
DMF-Water Mixture	Glucose-DMF	34.83 (13.33)	154.94 (17.03)	1096 (21.8)
	Glucose-Water	7.11 (9.39)	32.36 (13.15)	66.36 (14.93)
THF-Water Mixture	Glucose-THF	27.3 (12.78)	75.56 (15.25)	157.37 (17.08)
	Glucose-Water	7.53 (9.53)	19.68 (11.91)	25.16 (12.53)

Hydrogen bonds are key interactions that can affect the mobility and orientation of molecules. The hydrogen bond lifetimes and the hydrogen bond free energies between glucose and the solvent molecules were calculated and are given in Table 5. It can be seen that the hydrogen bond lifetimes of glucose with water and with co-solvents increase with an increase in the proportion of co-solvents. All the hydrogen bond calculations were performed using the Luzar and Chandler method.[73] This observation is in agreement with the RDF results in Fig. 10, which suggested stronger glucose–water interaction in the presence of co-solvents. The increase in hydrogen bonding strength of glucose with water and with co-solvents with increase in the proportion of the co-solvent would also be the reason behind reduced mobility of glucose molecules in the presence of co-solvents.

Chapter 7: Conclusions and Future Work

7.1. Conclusion

The solvation of glucose in water, dimethyl sulfoxide (DMSO), tetrahydrofuran (THF) and N,N-Dimethylformamide (DMF) and in mixtures of aforementioned solvents and water was studied using molecular dynamics simulations. We had initially benchmarked our chosen parameters with data from both experimental and simulation studies done previously to ensure the accuracy of the system as well of the results. It was inferred that the preexisting notion of applying and negative partial on the amide nitrogen did not best represent the entire functionality of DMF molecule. Our analysis suggested that a small positive partial charge better represented the DMF molecule. We had used ab initio molecular modeling to confirm our approach and to attest to the accuracy of the proposed parameters. Using the benchmarked parameters, simulations of glucose in both pure and mixed solvent systems were run. Our simulation results suggest that DMSO, DMF and THF form a strong first solvation shell around glucose and they compete with water to be in the first solvation shell. Even with the addition of a small amount of co–solvent, almost 50% of water molecules are moved out of the first solvation shell. Though the number of water molecules in the first solvation decreases upon the addition of these co–solvents, the interaction between glucose and water becomes stronger. Additionally, analysis of the 3-dimensional arrangement of co-solvent and water molecules around glucose shows that there is a pronounced localization of co-solvent near C3 and C4 oxygen atoms of glucose. These observations, in conjunction with the previous experimental and computational literature on the mechanism of HMF formation, suggest that that the removal of excess water from the immediate vicinity of glucose, stronger water-glucose interaction due to the presence of co-solvents and preferential arrangement of co-solvents around glucose would promote the HMF formation reaction and deter the formation of other dehydration/rehydration products. The analysis of the root mean square fluctuations of glucose molecules along the molecular dynamics trajectory suggest that the mobility of glucose molecules decreases with increasing the co-solvent proportion. This decrease in the mobility is attributed to stronger and increased lifetimes of hydrogen bonds between

glucose and solvent molecules. This would have a direct implication on condensation reactions leading to the formation of unwanted products like humins. The reduced mobility of glucose molecules and stronger interaction with solvents would minimize the probability of two glucose molecules or a glucose molecule and a sugar derivative like HMF to come in close proximity to undergo intermolecular condensation reactions. Thus, co-solvents may also indirectly promote the unimolecular dehydration reaction of glucose to form HMF and levulinic acid by hindering the formation of humins and other condensation products. We believe that these simulation studies provide a unique insight into the role of physical (and preferential) solvation of glucose by co-solvents like DMSO, DMF and THF in altering/affecting the chemistry of glucose transformation, without directly taking part into the reactions.

7.2. Future Work

The current investigation was focused on the solvation of glucose in selected commonly used solvents like DMF, THF and DMSO. To the best of our knowledge there isn't significant literature on the molecular level study of glucose in solvents and the effect of the co-solvents on glucose and their possible influences of the reaction outcomes. This investigation offers an insight into some of the possible interactions between the different solvent molecules and glucose. To identify and specify certain key interactions, a more expanded and wider range of analysis needs to be performed.

Solvation studies have been done on the dissolution of biomass molecules from the solutes perspectives involving both experimental and simulation studies. There has not been an extensive study into the solvation studies of biomass molecules from the solvents perspectives. IN the present work we had undertaken an initial step into the solvation of glucose in certain solvents. We propose that in order to get a clear understanding of the solvent rearrangement and the changes in the physical interactions as the reaction proceeds from glucose towards HMF, the scope of the investigation needs to move onto different solutes like fructose, HMF, levulinic acid, etc. to identify and compare the different interactions. Akien et al.[105] have used in situ nmr to map the acid catalysed dehydration of fructose and have identified some key intermediary compounds in the reaction mixture. The identified components can be added to the system and a interaction

landscape can be generated from glucose to HMF in the same set of solvents.

The analysis so far has not involved the usage of other solvents and a comprehensive study of the same can help in developing an optimal set of solvents that can help in improving the reaction outcomes.

References

(1) Alonso, D. M.; Bond, J. Q.; Dumesic, J. A. *Green Chemistry* **2010**, *12*, 1493.

(2) Wyman, C. E.; Dale, B. E.; Elander, R. T.; Holtzapple, M.; Ladisch, M. R.; Lee, Y. Y. *Bioresource Technology* **2005**, *96*, 1959.

(3) Regalbuto, J. R. *Science* **2009**, *325*, 822.

(4) Vanholme, R.; Morreel, K.; Ralph, J.; Boerjan, W. *Current Opinion in Plant Biology* **2008**, *11*, 278.

(5) Chakar, F. S.; Ragauskas, A. J. *Industrial Crops and Products* **2004**, *20*, 131.

(6) Huber, G. W.; Iborra, S.; Corma, A. *Chemical Reviews* **2006**, *106*, 4044.

(7) Evans, R. J.; Milne, T. A.; Soltys, M. N. *Journal of Analytical and Applied Pyrolysis* **1986**, *9*, 207.

(8) Azadi, P.; Carrasquillo-Flores, R.; Pagan-Torres, Y. J.; Gurbuz, E. I.; Farnood, R.; Dumesic, J. A. *Green Chemistry* **2012**, *14*, 1573.

(9) LYND, L. R.; CUSHMAN, J. H.; NICHOLS, R. J.; WYMAN, C. E. *Science* **1991**, *251*, 1318.

(10) Saha, B. *J IND MICROBIOL BIOTECHNOL* **2003**, *30*, 279.

(11) Zaldivar, J.; Nielsen, J.; Olsson, L. *Appl Microbiol Biotechnol* **2001**, *56*, 17.

(12) Mamman, A. S.; Lee, J.-M.; Kim, Y.-C.; Hwang, I. T.; Park, N.-J.; Hwang, Y. K.; Chang, J.-S.; Hwang, J.-S. *Biofuels, Bioproducts and Biorefining* **2008**, *2*, 438.

(13) Fatih Demirbas, M. *Applied Energy* **2009**, *86, Supplement 1*, S151.

(14) Hayes, D. J. *Catalysis Today* **2009**, *145*, 138.

(15) Kumar, P.; Barrett, D. M.; Delwiche, M. J.; Stroeve, P. *Industrial & Engineering Chemistry Research* **2009**, *48*, 3713.

(16) Rosatella, A. A.; Simeonov, S. P.; Frade, R. F. M.; Afonso, C. A. M. *Green Chemistry* **2011**, *13*, 754.

(17) Antal, M. J., Jr.; Mok, W. S.; Richards, G. N. *Carbohydrate research* **1990**, *199*, 91.

(18) Rivalier, P.; Duhamet, J.; Moreau, C.; Durand, R. *Catalysis Today* **1995**, *24*, 165.

(19) Moreau, C.; Belgacem, M.; Gandini, A. *Topics in Catalysis* **2004**, *27*, 11.

(20) Zhou, C.-H.; Beltramini, J. N.; Fan, Y.-X.; Lu, G. Q. *Chemical Society Reviews* **2008**, *37*, 527.

(21) Hansen, T. S.; Woodley, J. M.; Riisager, A. *Carbohydrate research* **2009**, *344*, 2568.

(22) Kuster, B. F. M. *Starch - Stärke* **1990**, *42*, 314.

(23) van Dam, H. E.; Kieboom, A. P. G.; van Bekkum, H. *Starch - Stärke* **1986**, *38*, 95.

(24) Musau, R. M.; Munavu, R. M. *Biomass* **1987**, *13*, 67.

(25) Amarasekara, A. S.; Williams, L. D.; Ebede, C. C. *Carbohydrate research* **2008**, *343*, 3021.

(26) Murugesan, S.; Mousa, S.; Vijayaraghavan, A.; Ajayan, P. M.; Linhardt, R. J. *Journal of Biomedical Materials Research Part B: Applied Biomaterials* **2006**, *79B*, 298.

(27) Liu, Q.; Janssen, M. H. A.; van Rantwijk, F.; Sheldon, R. A. *Green Chemistry* **2005**, *7*, 39.

(28) Rosatella, A. A.; Branco, L. C.; Afonso, C. A. M. *Green Chemistry* **2009**, *11*, 1406.

(29) Moreau, C.; Finiels, A.; Vanoye, L. *Journal of Molecular Catalysis A: Chemical* **2006**, *253*, 165.

(30) Zhao, H.; Holladay, J. E.; Brown, H.; Zhang, Z. C. *Science* **2007**, *316*, 1597.

(31) Alamillo, R.; Tucker, M.; Chia, M.; Pagan-Torres, Y.; Dumesic, J. *Green Chemistry* **2012**, *14*, 1413.
(32) Román-Leshkov, Y.; Moliner, M.; Labinger, J. A.; Davis, M. E. *Angewandte Chemie International Edition* **2010**, *49*, 8954.
(33) Mellmer, M. A.; Sener, C.; Gallo, J. M. R.; Luterbacher, J. S.; Alonso, D. M.; Dumesic, J. A. *Angewandte Chemie International Edition* **2014**, *53*, 11872.
(34) Mushrif, S. H.; Caratzoulas, S.; Vlachos, D. G. *Physical Chemistry Chemical Physics* **2012**, *14*, 2637.
(35) Nikolakis, V.; Mushrif, S. H.; Herbert, B.; Booksh, K. S.; Vlachos, D. G. *The Journal of Physical Chemistry B* **2012**, *116*, 11274.
(36) Li, G.; Pidko, E. A.; Hensen, E. J. M. *Catalysis Science & Technology* **2014**, *4*, 2241.
(37) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. *Journal of Chemical Theory and Computation* **2012**, *8*, 61.
(38) Mushrif, S. H.; Varghese, J. J.; Krishnamurthy, C. B. *Physical Chemistry Chemical Physics* **2015**, *17*, 4961.
(39) Mushrif, S. H.; Varghese, J. J.; Vlachos, D. G. *Physical Chemistry Chemical Physics* **2014**, *16*, 19564.
(40) Mushrif, S. H.; Vasudevan, V.; Krishnamurthy, C. B.; Venkatesh, B. *Chemical Engineering Science* **2015**, *121*, 217.
(41) Nikbin, N.; Caratzoulas, S.; Vlachos, D. G. *ChemCatChem* **2012**, *4*, 504.
(42) Brandt, A.; Grasvik, J.; Hallett, J. P.; Welton, T. *Green Chemistry* **2013**, *15*, 550.
(43) Fort, D. A.; Remsing, R. C.; Swatloski, R. P.; Moyna, P.; Moyna, G.; Rogers, R. D. *Green Chemistry* **2007**, *9*, 63.
(44) Kosan, B.; Michels, C.; Meister, F. *Cellulose* **2008**, *15*, 59.
(45) Lateef, H.; Grimes, S.; Kewcharoenwong, P.; Feinberg, B. *Journal of Chemical Technology & Biotechnology* **2009**, *84*, 1818.
(46) Swatloski, R. P.; Spear, S. K.; Holbrey, J. D.; Rogers, R. D. *Journal of the American Chemical Society* **2002**, *124*, 4974.
(47) Bergenstråhle, M.; Wohlert, J.; Himmel, M. E.; Brady, J. W. *Carbohydrate research* **2010**, *345*, 2060.
(48) Westman, C. O. a. G. In *Cellulose - Fundamental Aspects*; Ven, D. T. G. M. V. D., Ed.; InTech: 2013.
(49) Cai, L.; Liu, Y.; Liang, H. *Polymer* **2012**, *53*, 1124.
(50) Domínguez de María, P. *Journal of Chemical Technology & Biotechnology* **2014**, *89*, 11.
(51) Liu, Z.; Zhang, F.-S. *Energy Conversion and Management* **2008**, *49*, 3498.
(52) Caratzoulas, S.; Vlachos, D. G. *Carbohydrate research* **2011**, *346*, 664.
(53) Qian, X.; Wei, X. *The Journal of Physical Chemistry B* **2012**, *116*, 10898.
(54) Qian, X.; Liu, D. *Carbohydrate research* **2014**, *388*, 50.
(55) Gallezot, P. *Chemical Society Reviews* **2012**, *41*, 1538.
(56) Assary, R. S.; Kim, T.; Low, J. J.; Greeley, J.; Curtiss, L. A. *Physical Chemistry Chemical Physics* **2012**, *14*, 16603.
(57) Alder, B. J.; Wainwright, T. E. *The Journal of Chemical Physics* **1957**, *27*, 1208.
(58) Alder, B. J.; Wainwright, T. E. *The Journal of Chemical Physics* **1959**, *31*, 459.
(59) Rahman, A. *Physical Review* **1964**, *136*, A405.

(60) Rahman, A.; Stillinger, F. H. *Physical Review A* **1974**, *10*, 368.
(61) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature* **1977**, *267*, 585.
(62) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Computer Physics Communications* **1995**, *91*, 43.
(63) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *Journal of Chemical Theory and Computation* **2008**, *4*, 435.
(64) Lindahl, E.; Hess, B.; van der Spoel, D. *J Mol Model* **2001**, *7*, 306.
(65) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *Journal of Computational Chemistry* **2005**, *26*, 1701.
(66) Ponder, J. W.; Case, D. A. *Advances in protein chemistry* **2003**, *66*, 27.
(67) Jorgensen, W. L. *Journal of the American Chemical Society* **1981**, *103*, 335.
(68) Jorgensen, W. L.; Swenson, C. J. *Journal of the American Chemical Society* **1985**, *107*, 569.
(69) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *Journal of the American Chemical Society* **1996**, *118*, 11225.
(70) Kaminski, G.; Jorgensen, W. L. *The Journal of Physical Chemistry* **1996**, *100*, 18010.
(71) Rizzo, R. C.; Jorgensen, W. L. *Journal of the American Chemical Society* **1999**, *121*, 4827.
(72) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *Journal of the American Chemical Society* **1997**, *119*, 5908.
(73) Luzar, A.; Chandler, D. *Nature* **1996**, *379*, 55.
(74) Bowron, D. T.; Finney, J. L.; Soper, A. K. *Journal of the American Chemical Society* **2006**, *128*, 5119.
(75) Chheda, J. N.; Roman-Leshkov, Y.; Dumesic, J. A. *Green Chemistry* **2007**, *9*, 342.
(76) Jorgensen, W. L.; Madura, J. D. *Journal of the American Chemical Society* **1983**, *105*, 1407.
(77) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996**, *14*, 33.
(78) Hoover, W. G. *Physical Review A* **1985**, *31*, 1695.
(79) Nosé, S. *The Journal of Chemical Physics* **1984**, *81*, 511.
(80) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *The Journal of Chemical Physics* **1995**, *103*, 8577.
(81) IBM Corp 1990-2008.
(82) Car, R.; Parrinello, M. *Physical Review Letters* **1985**, *55*, 2471.
(83) Kohn, W.; Sham, L. J. *Physical Review* **1965**, *140*, A1133.
(84) Troullier, N.; Martins, J. L. *Physical Review B* **1991**, *43*, 1993.
(85) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical Review Letters* **1996**, *77*, 3865.
(86) Girard, S.; MÜLler-Plathe, F. *Molecular Physics* **2003**, *101*, 779.
(87) Luzar, A.; Chandler, D. *The Journal of Chemical Physics* **1993**, *98*, 8160.
(88) Brady, J. W. *Journal of the American Chemical Society* **1989**, *111*, 5155.
(89) Lei, Y.; Li, H.; Pan, H.; Han, S. *The Journal of Physical Chemistry A* **2003**, *107*, 1574.
(90) Marcus, Y. *The properties of solvents*; Wiley, 1998.
(91) Park, S.-K.; Min, K.-C.; Lee, C.; Hong, S. K.; Kim, Y.; Lee, N.-S. *Bull. Korean Chem. Soc* **2009**, *30*, 2595.

(92) Baev, A. K.; Springer: 2012.
(93) Alonso, D. M.; Wettstein, S. G.; Dumesic, J. A. *Green Chemistry* **2013**, *15*, 584.
(94) Ohara, M.; Takagaki, A.; Nishimura, S.; Ebitani, K. *Applied Catalysis A: General* **2010**, *383*, 149.
(95) Patil, S. K. R.; Heltzel, J.; Lund, C. R. F. *Energy & Fuels* **2012**, *26*, 5281.
(96) Patil, S. K. R.; Lund, C. R. F. *Energy & Fuels* **2011**, *25*, 4745.
(97) Qi, X.; Watanabe, M.; Aida, T. M.; Smith, R. L. *Bioresource Technology* **2012**, *109*, 224.
(98) Qian, X. *The Journal of Physical Chemistry A* **2011**, *115*, 11740.
(99) Gallo, J. M. R.; Alonso, D. M.; Mellmer, M. A.; Dumesic, J. A. *Green Chemistry* **2013**, *15*, 85.
(100) Feng, S.; Bagia, C.; Mpourmpakis, G. *The Journal of Physical Chemistry A* **2013**, *117*, 5211.
(101) Huo, F.; Liu, Z.; Wang, W. *The Journal of Physical Chemistry B* **2013**, *117*, 11780.
(102) Liu, D.; Nimlos, M. R.; Johnson, D. K.; Himmel, M. E.; Qian, X. *The Journal of Physical Chemistry A* **2010**, *114*, 12936.
(103) Qian, X.; Nimlos, M. R.; Davis, M.; Johnson, D. K.; Himmel, M. E. *Carbohydrate research* **2005**, *340*, 2319.
(104) Yang, G.; Pidko, E. A.; Hensen, E. J. M. *Journal of Catalysis* **2012**, *295*, 122.
(105) Akien, G. R.; Qi, L.; Horvath, I. T. *Chemical Communications* **2012**, *48*, 5850.