Investigating acid-catalyzed biomass reactions and solvent effects in humins formation, using multiscale molecular modeling

by

José Carlos Velasco Calderón

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

Department of Chemical and Materials Engineering
University of Alberta

Abstract

The conversion of lignocellulosic biomass to produce chemicals has gained significant interest recently. A broad spectrum of specialty chemicals can be obtained from lignocellulose-derived molecules. One of the most promising reactions in this category is the dehydration of cellulose-derived sugars such as glucose or fructose to 5-hydroxyl methyl furfural (5-HMF). 5-HMF is a valuable chemical platform from which polymers and biofuels can be produced. However, acid-catalyzed dehydration of fructose or glucose to 5-HMF is accompanied by side reactions, forming soluble polymers and insoluble humins. Humins are carbonaceous, polymeric by-products formed during acid-catalyzed condensed phase transformation of biomass-derived moieties. They are responsible for significant carbon loss and catalyst deactivation, and little is known about the chemistry and composition of humins. Additionally, condensed phase catalytic transformation of biomass is highly susceptible to solvent composition, where addition of aprotic solvents has shown to alter conversion and selectivity and can also avoid the formation of unwanted dehydration and rehydration by-products like humins and smaller molecular weight acids. Thus, revealing the mechanism of humins formation and elucidating the effect of solvent medium on humins formation chemistry are crucial knowledge gaps to fill to make quantum leaps in the development of biomass to chemicals technology.

In this thesis, for the first time, the reaction pathways and the activation and reaction-free energies of all the elementary reaction steps of the 5-HMF initiated acid-catalyzed humins formation are computed using density functional theory (DFT)-based calculations. The reaction pathway elucidated in this work explains the subsequent chemistry of further polymerization leading to humins. The computed mechanism and suggested step growth polymerization scheme are in excellent agreement with the experimental spectroscopic data in the literature. This work provides mechanistic details of humins chemistry for the subsequent higher-scale kinetic modelling of humins formation and optimizes the reaction conditions to minimize them.

The solvent effects are studied at two levels: physical interactions of the solvent with reacting species and preferential solvation as well as their chemical effects in altering reaction kinetics and thermodynamics. A computational approach of molecular mechanics-based molecular dynamics and well-tempered metadynamics methods is implemented to investigate physical effects. The effect of aprotic solvent on the

interaction of reacting species with the hydronium ions (acid catalyst) is studied for the fructose dehydration to 5-HMF as a simulation system model. Analogous to heterogeneous catalysis, the interaction between the reactant and the catalyst and the competition between reactant and solvent to interact with the catalyst surface is believed to govern the reactivity. We demonstrate a remarkable enhancement in the interaction between the hydronium ions and the fructose hydroxyl group as the DMSO concentration increased. Also, adding DMSO improves the stabilization of the hydronium ions in the first solvation shell of fructose compared to that in the bulk solvent. On the other hand, hydronium ions became less stable near 5-HMF as the concentration of DMSO increased. These observations help to better understand the improved conversion and selectivity during the acid dehydration of fructose to 5-HMF experimentally observed when adding aprotic polar solvent to the reacting system.

Three solvation modelling approaches are systematically compared to study the chemical effects of the solvent environment, i.e., implicit solvent, explicit, non-polarizable, equilibrated solvent molecules without dynamics and explicit solvent molecules with dynamics with polarizability. The model reaction is the acid-catalyzed protonation of 5-HMF, the first charge transfer step, leading to humins formation. The effect of the solvent environment on activation-free energy barriers was only captured when solvent and reaction dynamics were considered and when the solvent molecules were polarizable. The charge redistribution in the reacting species was observed to lead to the dynamic solvent reorganization; as the solvent composition changes, the Marcus reorganization energy changes, altering the reaction-free energy barriers. Additionally, the changes in the timescale of solvent reorganization led to non-equilibrium solvation.

This thesis investigates the molecular-level formation of humins from biomass and their derivatives, identifying key intermediates and modeling humin formation in condensed-phase reactions during acid-catalyzed biomass conversion. These findings guide the selection of catalysts, solvents, and operating conditions to enhance desired product selectivity and reduce humin formation in biomass-to-chemicals transformations.

Preface

This thesis is original work by Jose Carlos Velasco Calderón under the supervision of Dr. Samir H. Mushrif.

Chapter 2 of this thesis has been published as J. C. Velasco Calderón, J.S. Arora, and S. H. Mushrif "Mechanistic Investigation into the Formation of Humins in Acid-Catalyzed Biomass Reactions," ACS Omega 2022, 7, 49, 44786–44795. I was responsible for the conceptualization, methodology, validation, performing simulations, investigation, data analysis, writing, review, and editing. Dr. J.S. Arora provided guidance in simulation methods. Dr. S. H. Mushrif was the supervisory author in the conceptualization, investigation, review, and editing.

Chapter 3 of this thesis has been published as J. C. Velasco Calderón, S. Jiang, and S. H. Mushrif's "Understanding the Effect of Solvent Environment on the Interaction of Hydronium Ion with Biomass Derived Species: A Molecular Dynamics and Metadynamics Investigation," ChemPhysChem 2021, 22, 2222–2230. I was responsible for the conceptualization, methodology, validation, performing simulations, investigation, data analysis, writing - review and editing. S. Jiang assisted in the methodology and investigation. Dr. S. H. Mushrif was the supervisory author in the conceptualization, investigation, review, and editing.

Chapter 4 of this thesis has been submitted to a peer review journal as J. C. Velasco Calderón and S. H. Mushrif's "Energetics of acid-catalyzed biomass reactions: How and why does the solvent model matter?". I was responsible for the conceptualization, methodology, validation, performing simulations, investigation, data analysis, writing - review & editing. Dr. S. H. Mushrif was the supervisory author in the conceptualization, investigation, review, and editing.

Acknowledgements

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background

1.1.1 Biomass as a renewable source of energy

Biomass is the organic material in which the energy of the sunlight is stored in chemical bonds. This process occurs from the photosynthesis reaction between CO_2 in the atmosphere and sunlight to form carbohydrates as the building blocks of biomass. [1] When the chemical energy stored is released by combining with O_2, CO_2 is generated, leading to a carbon-neutral and cyclical process; hence, it is considered a renewable energy source. [2] Bioenergy is the largest renewable source, with 12 % of energy consumption globally. [3] More than 50 % of bioenergy currently corresponds to traditional uses such as cooking and heating. Other end uses such as electricity generation, and transport fuels share the rest of the total bioenergy consumption. [4]

1.1.2 Different biomass feedstocks

Biomass is a complex matrix with several feedstocks of different chemical composition and conversion technologies leading to a wide range of bio-products. The combination of feedstock and technology will determine the products that can be obtained for their end-use.[5,6] The biogenic feedstock can be divided into two broad categories: conventional and advanced. Oil crops fall under conventional bio-feedstock, such as palm, canola, and sunflower oils. Usually, these feedstocks are treated by transesterification to produce biodiesel or hydro processes HVO/HEFA to mainly produce renewable diesel (RD) and, more recently, sustainable aviation fuel (SAF). Additionally, sugar and starch crops such as sugarcane, corn, and cereals are fermented to produce bioethanol mainly. [7,8] On the other hand, biomass from waste is considered an advanced biogenic feedstock. Some examples are the biomass fractions of municipal solid waste or wet wastes from agriculture, which undergo anaerobic digestion to produce biogas or can be gasified to produce syngas. Finally, lignocellulosic biomass is the most promising advanced feedstock. Lignocellulose accounts for forestry, agricultural residues as well as energy crops. Lignocellulosic biomass has a wider range of production processes leading to different end-uses. Woody biomass can be pelletized and combusted for electricity and heat generation and converted into liquid or gas biofuels for transport. In addition, lignocellulose biomass-based materials and products can be obtained.

1.1.3 Environmental benefits of lignocellulosic biomass over other biogenic feedstocks

Among the biogenic sources, lignocellulosic biomass is the most abundant feedstock on the earth. [9] It is estimated that agricultural, forestry residues and energy crops (lignocellulosic feedstocks) have a substantially higher feedstock potential in terms of EJ per annum compared to the other biogenic sources. [10] In addition, the transformation of lignocellulosic biomass is less carbon intensive than other biogenic feedstocks. Lignocellulose-derived products can reduce up to 80 % GHG emission compared to the fossil-derived counterparts. [11] Lignocellulose conversion technologies are less carbon intense than biomass

farming feedstocks due to avoiding fertilizer, herbicides, and require less water consumption. [12] These advantages promote the imminent deployment of this technology, currently for biofuel production, such as renewable diesel and SAF.[13] However, since biomass is the only renewable source of carbon, lignocellulose is an appealing alternative to provide "green carbon" for chemical production to decarbonize the chemical industry. However, the main obstacle to scaling up this technology is the complex chemical composition and the challenging conversion chemistry associated with lignocellulose.

1.1.4 Lignocellulose composition

The main components of lignocellulose are cellulose, hemicellulose, and lignin as shown in Figure 1.1. Depending on the feedstock (energy crops, grasses, and soft/hardwood), cellulose and hemicellulose comprise 70 % wt. of dry biomass, while lignin content represents 25 % wt. [14–16] Cellulose is a polysaccharide with a crystallinity structure, and its elemental constituent is D-glucose linked through glycosidic bonds with formula ($C_6H_{10}O_5$). [17] On the other hand, hemicellulose consists of different sugar units, mainly pentoses (such as xylose and arabinose) and hexose units (glucose, galactose, and mannose). Hemicellulose is crosslinked to cellulose and lignin. Naturally, the transformation of hemicellulose compared to cellulose is more difficult due to its diverse composition. [18] Finally, lignin is an amorphous branched biopolymer, made up of three main monomer components, which are 4-propenyl phenol, 4-propenyl-2-methoxyphenol, and 4-propenyl-2.5-dymethoxylphenol.
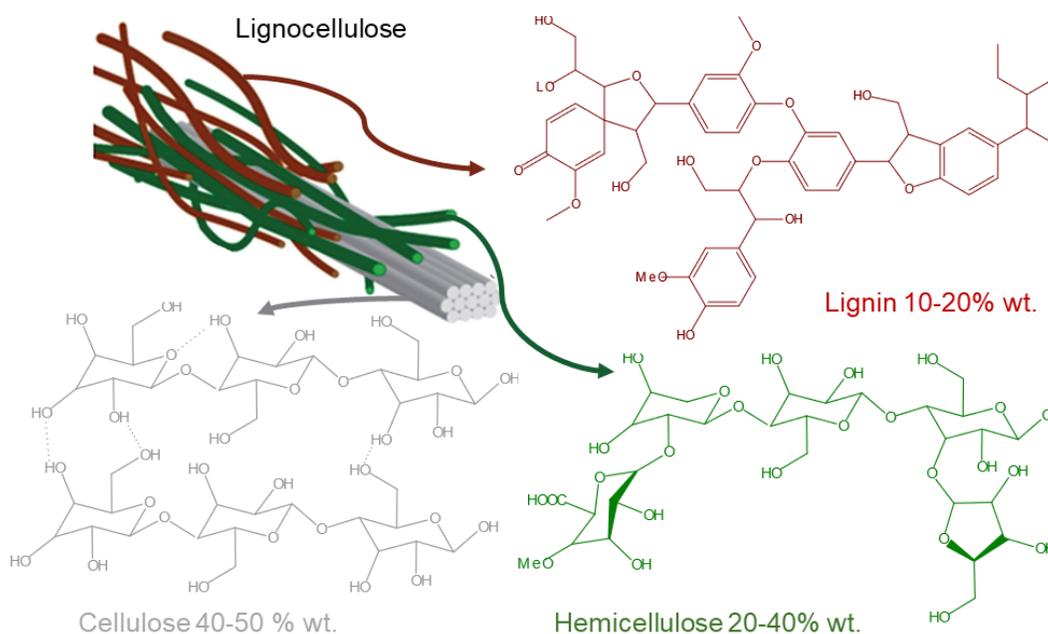
Figure 1.1 Composition of lignocellulosic biomass [14–16].

1.1.5 Pretreatment methods used for lignocellulosic biomass

Lignocellulose's recalcitrant structure makes it difficult to separate these three components to maximize their use adequately, and it must be pre-treated before its conversion to biofuels or value-added products. The most common deconstruction methods, such as mechanical, chemical, and biochemical, undergo extreme temperature and pressure conditions or aggressive use of chemicals, degrading hemicellulose and lignin.[,19] The mechanical category involves milling to reduce the biomass in size or ultrasonic to break down the biomass into smaller molecules.[20] Chemical methods englobe acid or alkaline hydrolysis to reduce lignin content and remove hemicellulose from cellulose or non-catalytic treatment at high temperatures (150-230 °C).[21,22] Organic solvents or oxidizing agents such as ozone, oxygen, or air are also used to delignify cellulose. Finally, biological hydrolysis using microorganisms removes hemicellulose and lignin from cellulose.[23] The choice of pretreatment ultimately depends on the conversion technology selected and the product targeted.

1.1.6 Technological routes for lignocellulose transformation

There are three main conversion categories for lignocellulose transformation, which are thermochemical, biochemical, and chemical conversion. [24]

1.1.6.1 Thermochemical conversion

Thermochemical processes use heat to produce either energy or liquid/ gaseous biofuels. Currently, there are four thermochemical production routes, i.e., combustion, pyrolysis, gasification, and liquefaction. Combustion processes target the production of electricity and heat, burning biomass. During the combustion process, the exothermic reaction combining biomass with oxygen produces CO_2, water, and heat.[25] The last one can be converted into mechanical and electrical energy.[26] The heating value of biogenic material is low compared to fossil feedstocks, i.e., coal, and this is not environmentally friendly because of the high emissions of CO, CO_2, NOx, and particulate materials. Lignocellulosic liquefaction is performed in water at lower temperatures (280 to 370 °C) and elevated pressures (10-25 MPa).[27] It targets a high H/C ratio liquid oil to be upgraded for biofuel production. Also, solid products are obtained and are mainly used for combustion or fertilizers.[28] On the other hand, pyrolysis is the thermal decomposition of biomass in the absence of air.[25] This process can be catalytic or non-catalytic and produce solids known as char, bio-oil, and not condensable gases consisting mainly of H_2, CO_2, CO, and light hydrocarbons.[29] Biochar qualifies for carbon sequestration and can be used as fertilizers, while bio-oil is upgraded to produce mainly renewable diesel and SAF.[30] Finally, gasification is the most mature technique in this category. It consists of the conversion of dry lignocellulose into syngas. Further, syngas can be used for electricity generation or biofuel production.[31,32] Downstream technologies for syngas conversion to biofuels, such as FT, syngas fermentation, and methanol-DME synthesis, are already commercialized.[33–35]

1.1.6.2 Biochemical conversion

Biochemical conversion involves the use of enzymes or biological agents to deconstruct the lignocellulosic biomass into carbohydrates, or to liquid or gaseous fuels.[36] The most representative technologies of this category are anaerobic digestion and fermentation. Municipal solid waste, agricultural residues, or sewage sludge are degraded during anaerobic digestion. Organic macromolecules are transformed into more simpler species through metabolic processes without oxygen to produce CH_4 and CO_2 mixed with traces of other gases, better known as biogas. Biogas is used mainly for electricity or heat generation. More recently, Fulcrum developed a process where the residual solid material is further gasified to produce biofuels through FT synthesis.[37] Besides, fermentation involves several reactions in series initiated by the acid and biological hydrolysis of lignocellulose to produce sugars in hexoses and pentoses.[38,39] Afterward, sugars are fermented in anaerobic conditions using biological agents to produce bioethanol and biogas. Ethanol has been the dominant biofuel in recent decades and is blended with fossil derived gasoline.[13] In addition, lignocellulose fermentation to produce ethanol is an emerging pathway with enhanced environmental benefits compared to conventional ethanol production.

1.1.6.3 Chemical conversion

In recent years, the chemical industry has used 13% of the oil consumed worldwide. [40,41] While renewable sources such as wind, hydro, solar and geothermal, among others, can cover a certain percentage of the global energy demand currently covered by fossil fuels, there are limited renewable carbon sources to replace petroleum-derived chemicals. [42] Carbon capture and utilization (CCU) is one of the popular renewable carbon sources in the energy industry; however, most of it comes from pre- or post-combustion processes within the natural gas or petroleum industry. [4,43] On the other hand, the technological alternative known as direct air capture (DAC) seems far in the future to mature and be considered a carbon source for the chemical industry. [44,45] Lignocellulosic biomass is an appealing alternative and renewable carbon source, even though the transformation of lignocellulose to produce chemicals has been relegated, prioritizing biofuel production deployment. [46]. Targeting the production of specialty and bulk chemicals and high-value materials and composites using biomass as the sole source of carbon, instead of producing only biofuels complete utilization of biomass carbon and intelligent strategies in selecting biomass fractions and components for refining into target chemicals and materials, minimizing the environmental footprint of the process. Furthermore, making high-value chemicals could be an economic driver for the biomass economy, boosting the entire bioenergy value chain. As shown in Figure 1.2, petroleum processes are commonly carried out at elevated temperatures. [47] In the same way, thermochemical treatments involve high temperatures that, in some cases, are above 1000 K. [48–50] However, the conversion processes of lignocellulose-derived carbohydrates are less energetically intense, treated at medium temperature conditions (usually below 400 K) in a liquid phase with catalysts.
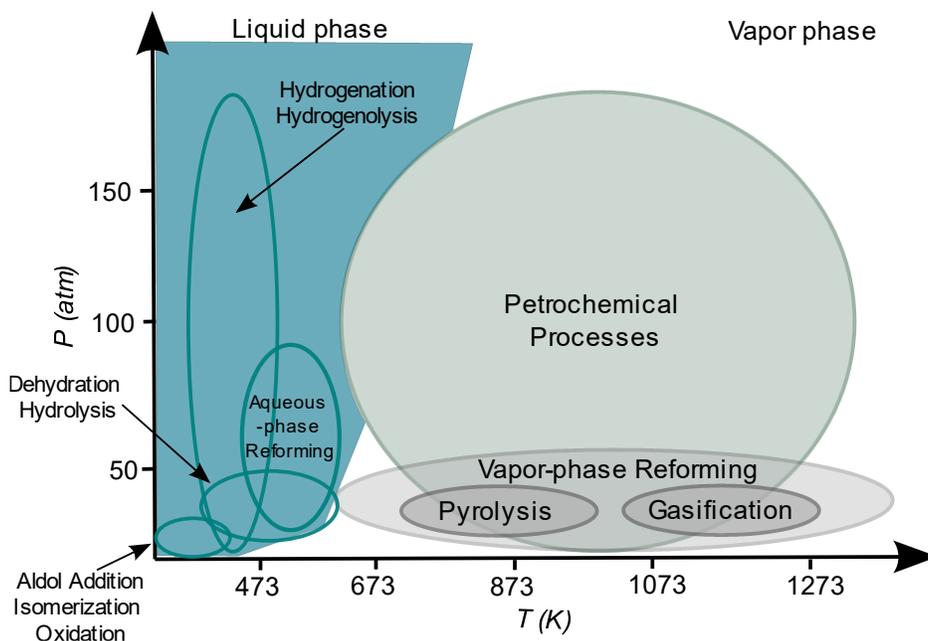
Figure 1.2 Diagram of approximate reaction conditions for the catalytic processing of petroleum and lignocellulosic biomass derived feedstocks. (Adapted from [51]) © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

The "know-how" for conversion methods of building block chemicals to final products is fully understood in the petrochemical industry.[52] For instance, the transformation of benzene obtained from crude oil to ethyl benzene and its subsequent conversion to styrene to obtain polystyrene has endless uses in producing textiles and in sectors such as transportation, health, and communications. In the case of the bio-based chemical industry, there is a similar but more complex conversion flow paths to that of the petrochemical.[52] Because biomass-derived constituents are densely functionalized and highly reactive, selective, energy-efficient transformation to chemicals is a very significant challenge. Technologies that have been tested for the transformation of cellulose-derived compounds to platform and commodity chemicals such as glucose, fructose, xylose, and arabinose, including some intermediate platforms, lead to C3, C4, C5, and C6 building blocks.[51] Subsequently, these building blocks are upgraded to secondary chemicals to obtain intermediate products equivalent to fossil sources, as shown in Figure 1.3 for glucose, since it is an important intermediate structurally similar to many cellulose-derived moieties. [52–54] This suggests that by transforming and upgrading lignocellulose derivatives; it is possible to substitute petroleum-based chemical products.[55] However, there is poor knowledge of fundamental reaction mechanisms and chemistries required for these transformations. Hence, it is necessary to understand the conversion processes and reaction mechanisms at fundamental level to overcome the technological challenges (explained in detail in following section) for deploying the bio-based chemical industry.
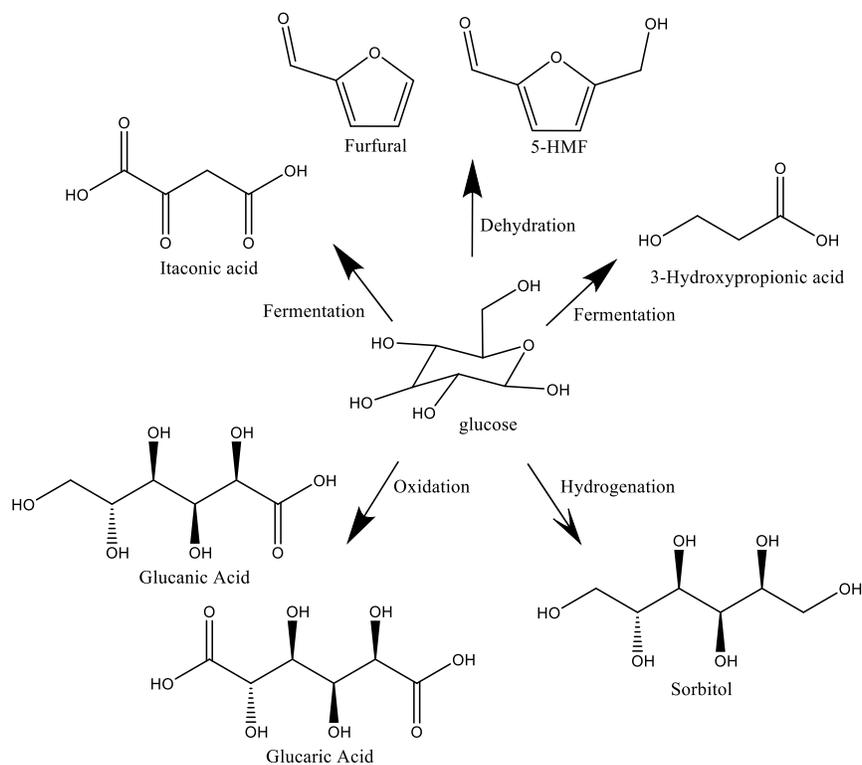
Figure 1.3 Valuable specialty chemicals derived from glucose.[52,53] **(Adapted from** [51]**). © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.**

1.1.6.4 Challenges in condensed phase biomass conversion

The main challenge in converting lignocellulose to specialized chemicals is that biomass moieties are highly functionalized and show high reactivity, as aforementioned. Due to the high biomass moieties functionality, controlling the selectivity by modulating catalyst technology and reaction conditions is difficult.[56] However, biomass molecules' high functionality allows access to a wide range of final products from a single biomass molecule. In a biorefinery treating lignocellulose, the cellulose and hemicellulose fractions undergo acid-catalyzed hydrolysis to be depolymerized, leading mainly to glucose (C6) and pentose units (C5).[52] Glucose, as an intermediate platform, exemplifies what is previously mentioned about the number of processes a single molecule can undergo to synthesize different building blocks, as shown in Figure 1.2.[53,57] The acid-catalyzed dehydration of cellulose-derived carbohydrates to furans has been extensively studied in the literature. [58] Specifically, the dehydration of glucose or fructose in acid media to 5-HMF is one of the most widely studied reactions in condensed-phase biomass to chemicals. [59] 5-HMF, a top valuable block chemical in condensed-phase biomass conversion, can undergo catalytic processes in acid conditions to rehydration, oxidation, or hydrogenation to obtain secondary chemicals for synthesizing polymers, such as polyhydroxy polyesters or polyurethanes, as shown in Figure 1.4. [60,61] However, 5-HMF synthesis is accompanied by side reactions, resulting in the formation of rehydration and condensation products such as humins. [62] The low selectivity in the case of 5-HMF production is mainly attributed to humins; up to 30

wt. % of the initial carbon content is lost in humins formation. [60] The process economics of biorefineries can be significantly improved by inhibiting the formation of humins, and this requires a thorough understanding of their formation-process and molecular structure. [63]
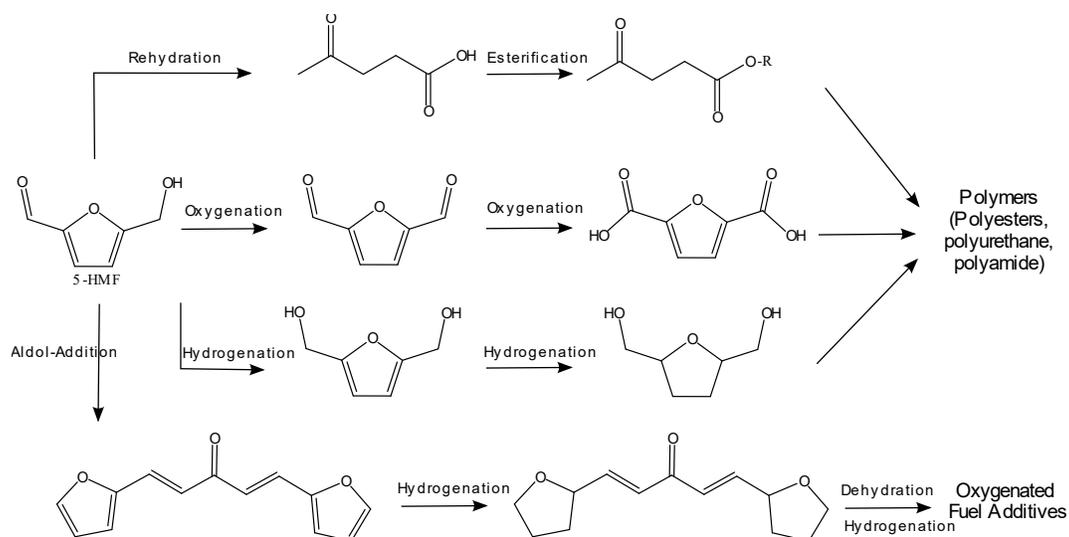
Figure 1.4 Sample reactions involved in the conversion of 5-HMF into chemicals. [52,53] (Adapted from [51]) © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

1.2 Implications of humins formation

Humins are polymerized, carbonaceous species formed during cellulosic biomass moieties' catalytic deconstruction and conversion reactions. Their formation lowers the selectivity towards desired platform chemicals, causes significant carbon loss, deactivates catalysts, and may cause severe processing and separation problems. Deposition of humins may also damage the reactor and its components. [64] Two examples of pilot and commercial scale technologies are Avantium and Biofine deploying building block chemicals from the acid-catalyzed conversion of lignocellulose.[65,66] However, their processes deal with the fact that the acid catalyzed conversion of cellulose and hemicellulose-derived molecules to platform chemicals are unavoidable, accompanied by the formation of humins. The operating temperature, the addition of co-solvents, and the strength of Lewis and Brønsted acidity of the catalysts are known to affect the formation of humins.[67] Most of the focus in the biomass community has been on improving the selectivity of desired products, with little attention to the chemistry of the formation of the undesired product, humins. Therefore, minimal knowledge of the chemistry of humins formation and their structure precludes systematic tailoring and optimization of reaction conditions and catalyst functionalities.

1.2.1 Experimental characterizations of humins composition

During the acid-catalyzed biomass conversion process, multiple reactions take place simultaneously. Hence, identifying and isolating the reactions leading to humins formation becomes experimentally challenging. Experimental techniques like FTIR can determine key functional groups present in humins.

Techniques like XPS can provide the elemental composition and propose possible reaction mechanisms leading to humins formation. A better understanding of humins chemical composition may positively influence biorefinery processes transforming lignocellulosic biomass to specialty chemicals in condensed phase as follows (i) Based on humins composition, possible reaction mechanisms leading to humins formation can be proposed, and by tailoring operating conditions or catalyst functionalities, humins formation can be suppressed. (ii) The understanding of the chemical composition of humins may lead to its valorization as to lignin, which structural characterization of lignin has guided the development of catalytic lignin conversion strategies or analogous to biochar, that influences a range of soil chemical properties and rapid changes to nutrient availability. [30] Hence, humins may also be integrated as a commodity, favoring the entire biomass to bio-based chemicals value chain.

Since humins are large, polymeric, and insoluble molecules and it is extremely difficult to determine their exact chemical structure, it becomes difficult to accurately determine the intermediates/precursors that result in the formation of humins. Compositional analysis was carried out for humins obtained from the acid hydrolysis and dehydration of hexoses (D-glucose, D-mannose, D-galactose) and pentoses (D-arabinose, D-xylose) as shown in Figure 1.5. Nuclear magnetic resonance (NMR) and infrared (IR) spectra results revealed that [62]
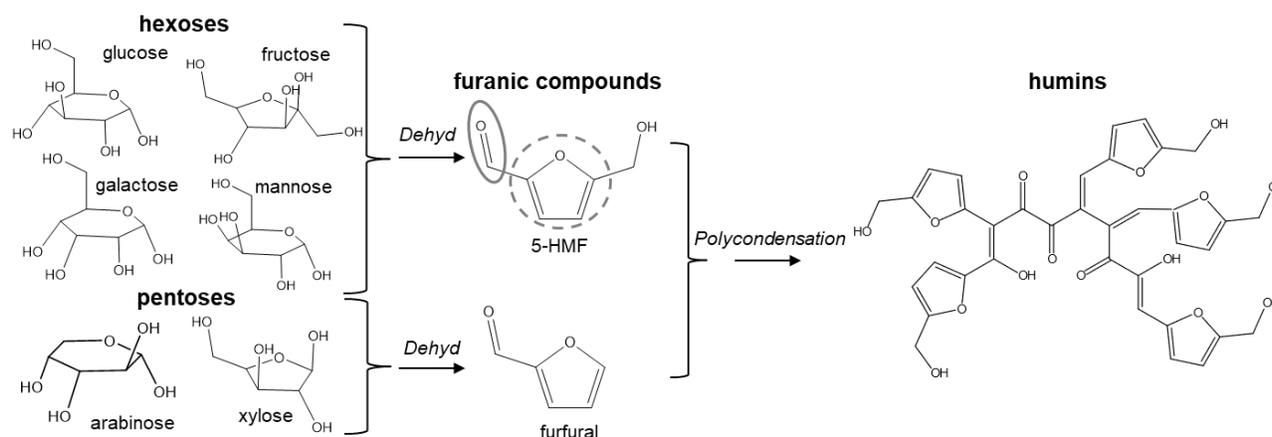
Figure 1.5 Humins derived from different hexoses, pentoses and furanic compounds

1. Humins structure contains a large amount of C-O-C bonds.
2. Humins consist of 60 % of furan rings and 20 % aliphatic fragments.

Hexoses and pentoses undergo dehydration to 5-HMF, and furfural, respectively. These in turn, can decompose to humins. IR spectra of humins from 5-HMF showed the absence of the carbonyl group of 5-HMF in contrast to the presence of the peaks corresponding to the furan ring and the hydroxymethyl group. These peaks suggest that 5-HMF is incorporated into the humins structure by aldol-condensation through its aldehyde group. Comparison between IR spectra of humins derived from acid-catalyzed dehydration of fructose and glucose highly resembled IR spectra for humins obtained from 5-HMF shown in Figure 1.6. However, there is a peak range where the intensities of the peaks differ between humins from glucose,

fructose, and 5-HMF. The set of peaks attributed to the furanic ring of 5-HMF (750 to 850, 1030, and 1525 cm-1) is stronger for humins derived from fructose and 5-HMF than from glucose.[68] It is suggested that the difference in peak intensity is because of the low 5-HMF concentration during the glucose dehydration, in contrast to the higher 5-HMF concentration in the reacting system using fructose or 5-HMF as a reactant. [69] The infrared spectrum comparison proposes that glucose and fructose first undergo dehydration to 5-HMF instead of direct sugar polymerization leading to humins. [70]
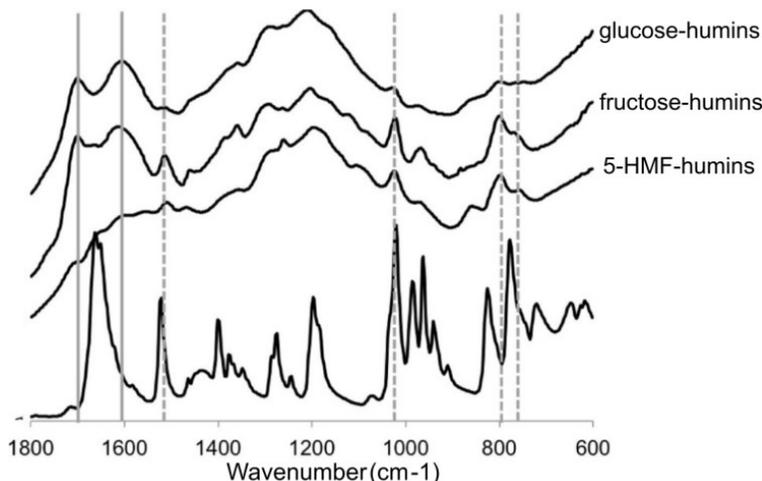
Figure 1.6 IR spectra of (a) HMF and of humins formed using 0.1 M H2SO4 at 408 K from (b) 0.1 M HMF at 90% conversion, (c) 0.1 M fructose at 88% conversion, and (d) 0.1 M glucose at 75% conversion. [68] Copyright © 2011, American Chemical Society.

The composition of the reaction system was measured via HPLC during the acid degradation of 5-HMF to humins. By analyzing the kinetic data, it is determined that levunilic and formic acid, which are also by-products of the rehydration of 5-HMF, are not involved in humins formation. Experiments also showed that when benzaldehyde is added during the acid catalyzed conversion of glucose to humins, benzyl groups are incorporated into the humins structure, proving that humins can be artificially functionalized. [71] Figure 1.4 shows the suggested humins structure from different carbohydrates as feedstock from different studies.

1.2.2 Implications of feedstock and process conditions in humins formation

As previously mentioned, the feedstock and process parameters greatly influence the acid-catalyzed formation of humins. It was observed between glucose, xylose, and fructose that fructose yielded more humins than other reactants (30, 32, and 39 % wt., respectively, at 180 °C). The difference is attributed to the higher 5-HMF concentration during fructose dehydration in solution than other carbohydrates. Additionally, during the glucose conversion to humins in acid conditions at a constant reaction time, it was observed that humins yield mainly depends on temperature and acid concentration and a lesser extent to sugar concentration. [72] Finally, IR spectra of humins recovered were compared for different glucose and fructose conversion. The results disclosed that the IR spectra did not change appreciably as a function of conversion. [70] Adding DMSO to the reacting system leads to almost no solid formation (soluble humins) in contrast to particle precipitation found in pure water. This observation suggests that humins are partially

soluble in DMSO or that the particle size growth is deterred in the presence of DMSO. IR spectra of humins formed in the presence of DMSO indicated higher intensity for the 5-HMF aldehyde group peak. This suggests that DMSO inhibits humins oligomerization; and it is postulated that 5-HMF is incorporated into humins structure through its carbonyl group. [69]
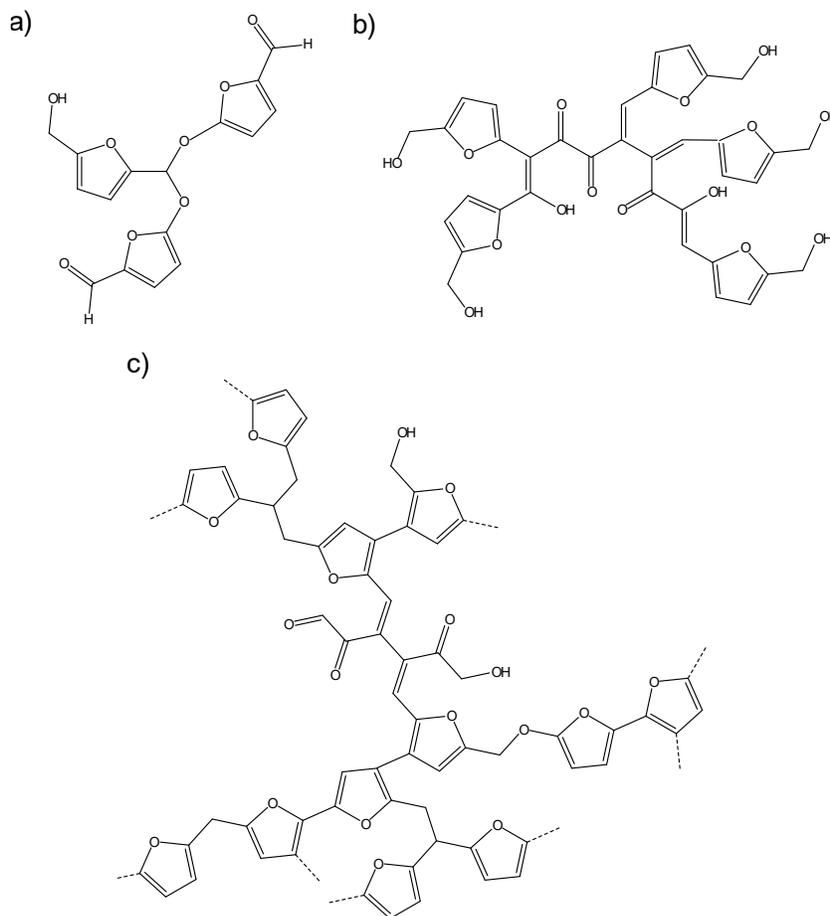
Figure 1.7 Suggested structure of (a) fructose-derived humins proposed by Sumerskii et al [62] (b) HMF-derived humins proposed by Patil et al.,[71] and (c) glucose-derived humins proposed by van Zandvoort et al.[64,72] Reproduced/Adapted from ref. 73 with permission from The Royal Society of Chemistry.

1.2.3 Humins morphology

Scanning electron microscopy (SEM) technique was used to determine particle size, shape, and texture for humins derived from different biomass moieties as shown in Fig 1.8. In the case of humins derived from 5-HMF, it was found that the average diameter increased as the reaction progressed. Particle growth involved both, an increase in single particle diameter and particle coalescence. When humins are obtained through a non-catalytic process (hydrothermal humins), humins morphology becomes amorphous (non-spherical). Changes in humins' morphology may result from different chemical compositions when humins are obtained under non-catalytic conditions. Moreover, hydrothermal humins require significantly longer reaction time

than acid-catalyzed humins. Under hydrothermal conditions, 15 % 5-HMF conversion happens after 120 h, while in acid-catalyzed humins formation, 5-HMF reaches complete transformation in just 3 h. [71] Similarly, in fructose, at 37 and 88 % conversion, humins morphology changes from spherical particles to amorphous shaped particles. Nevertheless, it has been indicated that changes in morphology are mainly a consequence of reaction time and not a function of conversion. [70] Comparing the morphology of humins derived from hexose sugars with those derived from xylose, it is observed that humins derived from glucose or fructose appear as spherical particles interconnected ≈ 3-5 mm in size. In contrast, xylose humins consist of spherical but isolated particles. On the other hand, a mixture of carbohydrates leads to agglomerated (amorphous) particles. Hence, the morphology of humins depends on their composition, which is based on the reactant. [72] Finally, the composition of the reacting system also has implications on humins' morphology. For instance, the addition of DMSO affects humins morphology. 5-HMF-derived humins morphology is invariant with spherical particles of 50-100 nm during the entire reaction time in presence of DMSO, in contrast to the particle size distribution observed for humins in pure water at different reaction time intervals. In conclusion, these results indicate that humins morphology is strongly affected by reaction conditions and the solvent environment. [69]
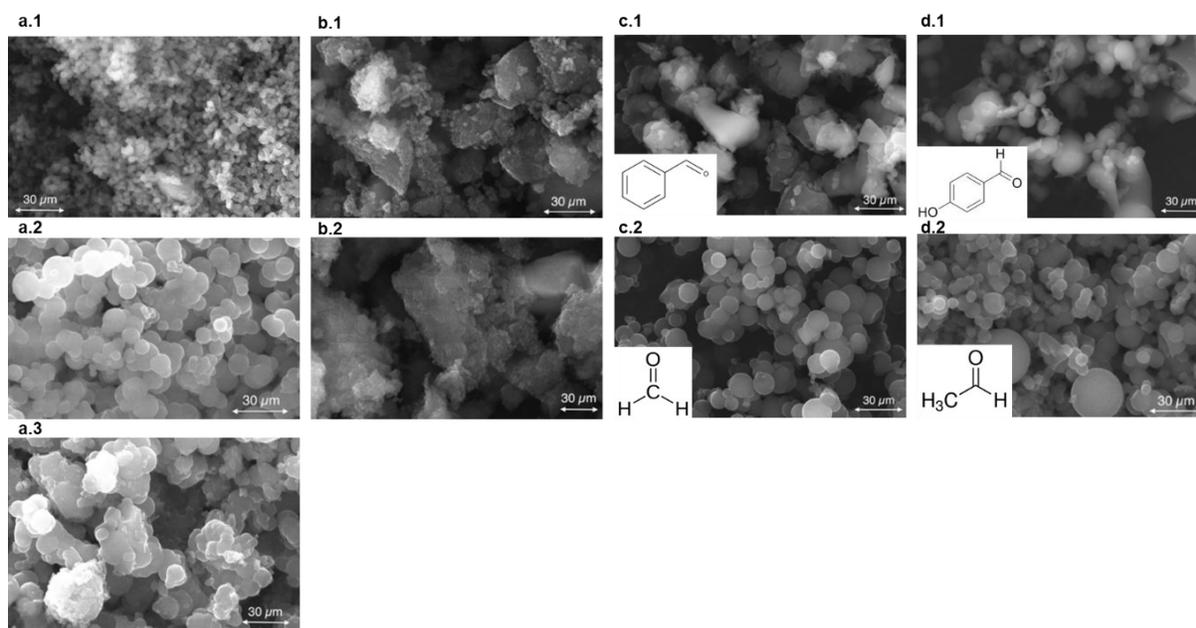
Figure 1.8 a) 0.1 M sulfuric acid 0.1 M Sulfuric acid 0.1 M 5-HMF at 135 °C a.1) 22 min (43% 5-HMF conversion) a.2) 2 hrs (90% 5-HMF conversion) a.3) 18 hrs. (100% 5-HMF conversion) b) no catalyzed reaction b.1) 49 h (45% HMF conversion) b.2) 120 h (85% HMF conversion) c & d) acid-catalyzed conversion of an equimolar mixture of 5-HMF c.1) benzaldehyde c.2) formaldehyde d.1) hydroxybenzaldehyde d.2) acetaldehyde. [72] Copyright © 2012, American Chemical Society.

1.3 Effect of solvent composition in condensed phase biomass reactions

It has been found that both the deconstruction of lignocellulose to structurally simpler molecules and the conversion of intermediate platforms to bio-based chemicals are highly sensitive to changes in the

composition of the reacting environment. Difficulties have been encountered in arriving at a solvent and reaction environment to enhance selectivity, facilitate separations, and can mediate the chemistry at high concentrations under continuous operation. Aprotic polar solvents are usually mixed in the aqueous reaction medium to enhance yield for several biomass conversion processes. [74–76] Aprotic solvent added to the aqueous reaction system has been widely studied for acid-catalyzed conversion methods, such as lignocellulose saccharification, lignin depolymerization or upgrading cellulose and hemicellulose C6/C5 molecules to building blocks chemicals, leading to enhanced yields. [77–80] Specifically, for the dehydration of biomass moieties to 5-HMF, adding an aprotic solvent enhanced yields and deterred the formation of humins. [81–84] Specific studies (experimental and simulations) are detailed in chapters 3 and 4. Here, an overview is presented about how solvents can influence kinetics, conversion, and selectivity when converting lignocellulose-derived molecules to chemicals. Apart from experiments, computational modelling insights are discussed to provide a fundamental molecular understanding of the role of the solvent in biomass reactions, rationalizing what is observed experimentally and suggesting modifications of the reaction conditions (solvent systems composition) and catalytic technologies to improve these processes.

1.4 Implications of solvent environment on reactivity and selectivity

Changes in reactivity in the presence of aprotic solvents are due to alterations in the interaction between the solvent medium and the reacting species during the reaction (reactant, transition state, and product). In addition, the solvent can modify the interaction between the catalyst and the reacting species, affecting the reactivity. On the other hand, as mentioned before, the highly functionalized bio-feed stocks lead to unselective reactions leading to humins, which can be seen analogous to coke formation in petroleum processing. Adding an aprotic solvent influences humins formation too. These factors are exemplified and discussed in the following subsections in the context of the acid-catalyzed bio-based molecules reactions, where most of them are accompanied by humins formation.

1.4.1 Competitive interaction between solvent and the reacting species with the catalyst

Competition between the solvent molecules and the reactants interacting with the catalytic surfaces may alter the overall reactivity. [85] For example, during the zeolite-catalyzed aldose to ketose isomerization of glucose to fructose (a critical step to produce 5-HMF from sugars) in water, higher turnover rates are observed for hydrophobic Ti-Beta zeolite compared to the hydrophilic Ti-Beta-OH zeolite. [86,87] Lower reactivity in the hydrophilic catalytic environment is associated with the stronger binding between water molecules with the silanol groups (Si-O-H), inhibiting the interaction of glucose with the active sites in contrast to the weak or non-interaction of water molecules with the hydrophobic catalytic surface, as observed in Figures 1.5 a & b. Besides, higher reactivity for the hydrophobic Ti-Beta zeolite is observed using methanol instead of water, as shown in Figure 1.5 c. The DFT-based calculation for the isomerization of glucose using hydrophobic Sn-Beta in water and methanol found that the change in reactivity for the hydrophobic catalyst is associated with the capacity of methanol to solvate the pores despite water molecules and, this provides better stabilization to the reacting specie throughout the reaction. [88]
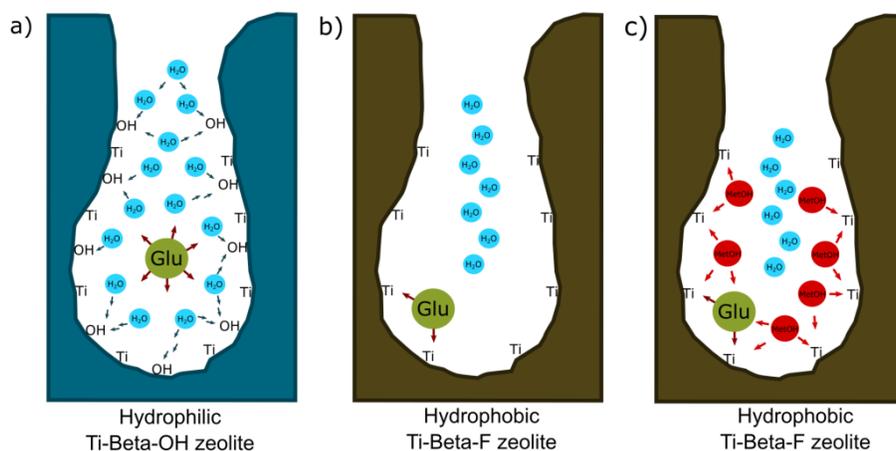
Figure 1.9 Illustration of a) the solvation of the hydrophilic zeolite pore, inhibiting contact of glucose with the Lewis acidic Ti active site due to competitive inhibition by strongly bound water molecules which are stabilized by hydrogen bonding with silanol groups b) the poor solvated hydrophobic zeolite pore, glucose in contact with the Lewis acidic Ti active site. c) the solvated hydrophobic zeolite pore which facilitates contact of glucose with the Lewis acidic Ti active site and enhanced by the stabilization from methanol molecules. Reproduced/Adapted from ref. 89 with permission from The Royal Society of Chemistry.

Alcohols undergo dissociative adsorption to the Pt/TiO2 catalyst forming alkoxides that lead to free protons for the hydrogenation of alkyl benzene compounds. Even using alcohols of different molecular structures, the substrate-catalyst interaction changes. Higher reaction rates are determined for secondary over primary alcohols because the latter form a densely well-packed catalyst inhibiting the reactant's adsorption to the catalytic surface, unlike the secondary ones with a spaced attachment to the catalyst.[90] In the case of other solvents such as aromatics (i.e. benzene and toluene), they have a strong interaction with the catalytic surface generating a low H^+ coverage, unlike cyclohexane, with a poor interaction with Pt-catalyst causing a high H^+ coverage.[91]

1.4.1.1 Relative stabilization of the reactant, transition state and product

The increase in reactivity considering could be due to the decrease in the energy barrier associated with the chemical reaction. The decrease in free energy can be achieved by destabilizing the reactant state or better stabilizing the transition state, as shown in Figures 1.6 a & b. As previously noted, acid-catalyzed dehydration reactions of biomass-derived carbohydrates show substantially higher reaction rates when a polar aprotic solvent is added to the aqueous reacting phase. Molecular dynamics simulations (MD) were performed on different carbohydrates such as fructose, xylitol and cellobiose in the mixture of water: 1,4dioxane at different proportions.[92] Local water-enriched domains near the -OH groups were identified by increasing the aprotic polar solvent proportion.[93] The same effects were encountered for other aprotic solvents such as DMSO or GVL. In the context of homogeneous acid-catalyzed dehydrations, the following two scenarios are possible (i) increasing local water domains promotes the protonation and subsequent dehydration of the reactants (destabilizing the reactant state)[94–96] and (ii) the nearest neighbor water

13

molecules configure themselves stabilizing the charged transition state (stabilizing the TS).[97–99] Finally, as previously mentioned, the synthesis gas is closely related to biomass's transformation as an intermediate product, for which its upgrading is essential. Hence, a plausible upgrading method is the hydrogenation of CO to produce valuable chemicals, which in turn illustrates the effect of the solvent in increasing reactivity and selectivity by stabilizing the product (thermodynamic stabilization), as shown in Figure 6c. The hydrogenation of CO in the gas phase to form formaldehyde is highly endergonic (energetically unfavorable). However, using alumina-supported Ni-Ru bimetallic catalysts in the aqueous phase, water molecules stabilize formaldehyde (the product) through H-bonds making the reaction energetically possible. [100] This enhancement in reactivity is amplified in the presence of water: methanol mixtures. In an aqueous system, formaldehyde (1st product) is converted to methylene glycol, while methanol: water mixtures alternatively lead to hemiacetal formation (2nd product in presence of methanol). Ab initio molecular dynamics simulations using the CPMD scheme revealed that methanol molecules stabilize hemiacetal via H-bonds forming an oligomeric-type structure (product stabilization). [101]
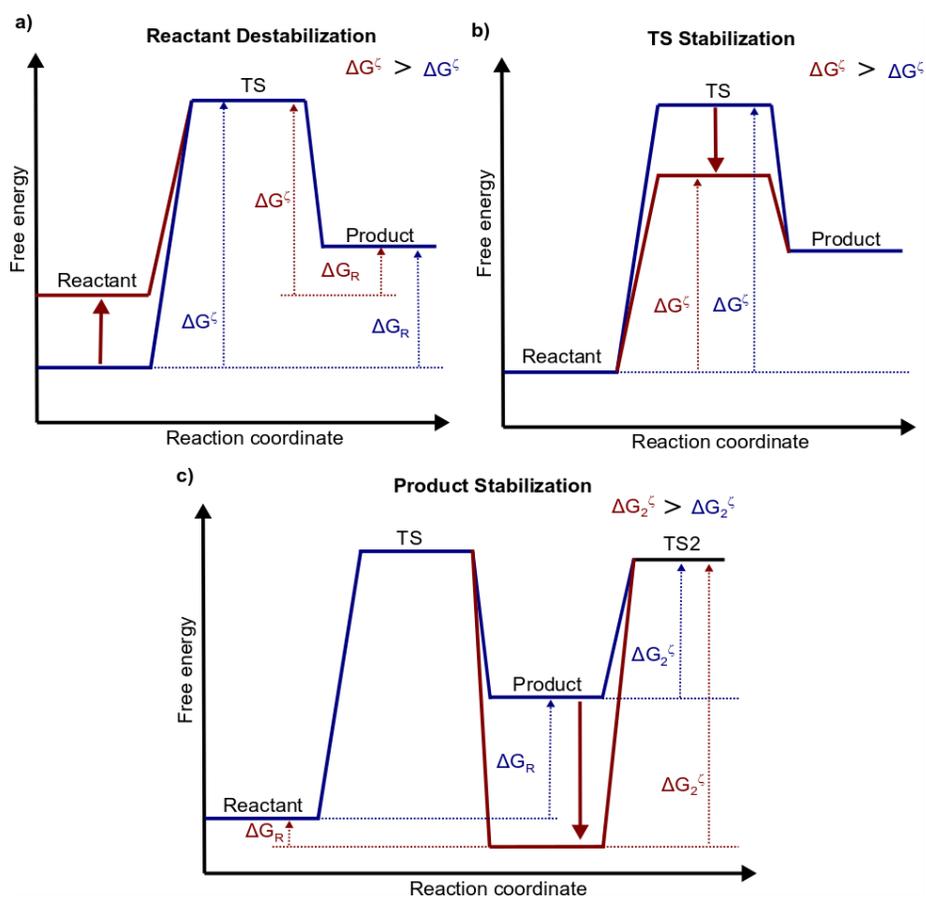
Figure 1.10 Effects on the energy profile along the reaction for a) destabilization of reactant, b) stabilization of the TS c) stabilization of the product. Figure 6 Effects on the energy profile along the reaction for a) destabilization of reactant, b) stabilization of the TS c) stabilization of the product.

1.4.1.2 Change in reaction mechanism

Solvent molecules can directly participate in the chemistry by acting as a catalyst. For the dehydration of fructose to 5-HMF in neat DMSO (no acid catalyst), DMSO can participate as a proton carrier facilitating the two dehydrations leading to 5-HMF, suggesting that in water: DMSO mixtures, the dehydration of fructose can be acid or DMSO-catalyzed, as shown in Figure 1.7.[102] Additionally, the changes in solvent composition sometimes influence the reaction's thermodynamics, altering the conversion and selectivity. For example, during the acid-catalyzed hydrolysis of cellulose to 5-HMF in an acid-aqueous environment, 5-HMF is produced from the triple dehydration of glucose. However, it is determined that using THF, the glucose-levoglucosan (LGA) chemical equilibrium changes towards LGA, which in turn is dehydrated to form 5-HMF. [103] The effect that the addition of the aprotic solvent has on the direct synthesis of 5-HMF from cellulose are (i) 20 times higher reaction yields, (ii) no need of Lewis acid catalyst (used for the isomerization of glucose to fructose) and (iii) higher 5-HMF yields by suppressing humins formation.
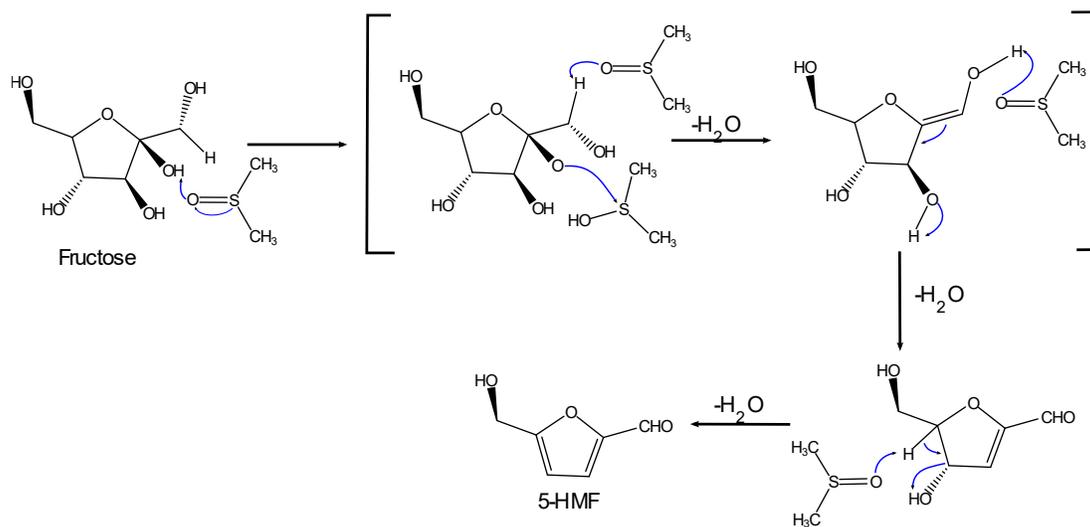
Figure 1.11 Proposed mechanism for the DMSO catalyzed dehydration of D-fructose furanose forms to 5-HMF. (Adapted from [102]) (Reproduced with permission from Elsevier, © 2008 Elsevier Ltd)

1.4.1.3 Inhibition of undesired reactions by physical and reactive solvation

Increased selectivity in biomass reactions in the condensed phase by changing the composition of the solvent environment can be achieved by adding a solvent insoluble in the reacting phase, capable of extracting the desired product and thus avoiding its degradation. For instance, high selectivity of 5-HMF from glucose are obtained in a biphasic system of water and 2-sec-butyl phenol using both Lewis and Bronsted acid catalysts for the isomerization of glucose to fructose and dehydration of fructose to 5-HMF, respectively. These two reactions occur sequentially in the reacting aqueous phase, while 97 % of 5-HMF

produced was extracted to the organic phase avoiding 5-HMF degradation into humins.[104,105] Also, the addition of aprotic polar solvent soluble in water can enhance conversion and suppress biomass-derived chemical degradation to humins. Force field based molecular dynamics of fructose in water: DMSO mixtures and glucose in aqueous environments mixed with DMSO, tetrahydrofuran (THF) or N, N-dimethylformamide (DMF) reveal that adding the aprotic solvents, promotes water molecules to be oriented towards the oxygen atoms of the –OH groups while the co-solvent molecules interact with the alpha-hydrogen atoms. Figure 1.8 shows how fructose in furanose form is solvated by (a) water and (b) water: DMSO mixture. It is observed how DMSO (blue colored) competes with water (red colored) to interact with the alpha-hydrogens located beside the –OH groups while water molecules interact stronger with the oxygen atoms. The spatial rearrangement near the reactant's functional groups leads to enhanced yield because of the proton donor capacity of water molecules, while aprotic solvent inhibits side reactions.[98,99]. Finally, there is a scenario where the solvent inhibits unwanted reactions from occurring: by deactivating the product's active sites through a solvent-product reaction. That is the case when formaldehyde is added during the hydrogenolysis to depolymerize lignin. NMR spectra show that formaldehyde is incorporated into the monomer structure in acidic conditions in the form of a 6-membered 1, 3 dioxane reacting with 1,3 diols on the lignin structure or via ortho/para-addition with the methoxy group of aromatic rings of lignin. Reaction yields for monomer formation increase from 26 to 78 %, avoiding uncontrolled polymerization reactions.[106]
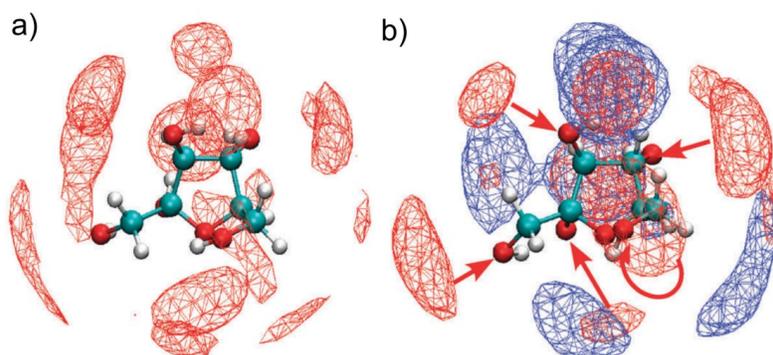
Figure 1.12 Volumetric map of the time averaged distribution of oxygen atoms in a) pure water and b) water and DMSO mixture around a fructose molecule in the simulation. Red colored surfaces indicate water oxygen atoms and blue-colored surfaces indicate DMSO oxygen atoms. Reproduced/Adapted from ref. 99 with permission from the PCCP Owner Societies.

1.5 Motivation of the thesis

It is imperative to diversify biomass conversion efforts, focusing on high-value-added specialty chemicals and materials. Humins formation results in the net loss of carbon (and hence the desired product) in the biomass to be converted to fuels and chemicals in liquid-phase catalytic processing. Their formation lowers the selectivity towards desired platform chemicals, deactivates catalysts and may cause severe processing and separation problems. Due to the high-functionalized molecules derived from biomass to produce bio-based chemicals, there are experimental difficulties in identifying the reactions that lead to humins

formation. Scientific investigations to study and analyze the effect of operating conditions, catalyst functionalities, and the choice of solvents on humins formation face barriers because of very little knowledge about the structure of humins and their precursors and almost no understanding of the catalytic chemistry of humins formation. Hence, in this work, we attempt to determine the reaction mechanisms that lead to the formation of humins, identifying critical intermediate reactions and calculating the associated energy barriers based on experimental evidence suggesting dominant reaction mechanisms according to the chemical composition of the oligomeric humins and kinetics studies in section 1.2 and further discussed in Chapter 2. The molecular modelling methodology allows us to study the kinetics and energetics of humins formation, which may not be feasible using a purely experimental approach. Additionally, to recover the economic loss in biomass conversion processes because of humins formation, the best possible alternative is to explore ways to suppress its formation. Biomass transformation to specialty chemicals in the condensed phase is highly susceptible to solvent composition, as discussed in section 1.3. Biomass acid-catalyzed reactions to biochemicals are highly enhanced in yield and avoid forming unwanted humins when adding aprotic solvents in the reacting mixture. The effects of solvent-altering humins formation chemistry are discussed in detail in Chapters 3 and 4. The solvent effects can be decoupled as physical and chemical interactions directly participating in the reaction by altering kinetics. Physical interaction encompasses preferential solvation of the reacting species. It can provide a shielding effect to avoid unwanted reactions, while chemical effects involve alterations in activations barrier, chemical equilibrium, or even changing reaction mechanisms. Hence, understanding the effects of solvent at the molecular level, functionalities in catalytic and solvent systems can be proposed to suppress humins formation by either physical and/or chemical participation. Different approaches can be formulated without incurring intense separation costs, minimizing humins formation in homogeneous and heterogeneously catalyzed biomass conversion reactions. The approach proposed using multiscale molecular modelling tools will provide a rigorous methodology to test solvent mixtures for carrying out biomass reactions instead of screening and selecting these solvents by a trial and error. The results obtained will not only benefit the biomass conversion process by minimizing humins formation. However, they will also significantly impact many chemical, pharmaceutical and food-related processes, where solvents are believed to play a critical role.

1.6 Objectives of the thesis

The overall objectives of this thesis are i) to reveal the mechanism of humins formation using homogeneous acid catalysts and identify the pathways that lead to the formation of humins and their precursors. Moreover, computing the reaction energetics associated with the pathways and recognizing the rate-limiting step in each pathway and ii) to identify the effect of solvent medium on humins formation, elucidating preferential physical solvation of biomass moieties and specialty chemicals derivatives by providing shielding effect for their degradation to humins. In addition, we attempt to investigate the role of solvent composition directly in humins formation chemistry.

The first objective is fully addressed in Chapter 2. Even though information about the chemistry and composition of humins is minimal, there is sufficient experimental evidence discussed in Chapter 2 to conclude that 5-HMF is a precursor for humins formation in the sugars dehydration reaction and to suggest that rehydration and addition-condensation reactions between 5-HMF and its derivatives initiate humins formation. Hence, in this section, for the first time, density functional theory (DFT)-based calculations are performed to compute reaction pathways, activation barriers, and reaction-free energies associated with all elementary reaction steps in the 5-HMF initiated, acid-catalyzed, humins formation.

The effects of the solvent environment in biomass reactions leading to humins formation are addressed in Chapters 3 and 4. Chapter 3 deals with the physical solvation (shielding effects and competitive interaction solvent-catalyst with the reactive species). To the best of our knowledge, no previous work has provided information about the effect of aprotic solvents on the interaction between reactants and products with the homogeneous acid catalyst, i.e., the hydronium ion, even though there is sufficient experimental evidence that indicates an increase in the activity of acid catalysts in the presence of aprotic solvents. Analogous to heterogeneous catalysis, where the interaction between the reactant and the catalyst and the competition between reactant and solvent to interact with the catalyst surface is believed to govern the reactivity, we aim to demonstrate here the effect of solvent environment, i.e., the concentration of an aprotic solvent like DMSO, on the interaction of hydronium ions (the catalyst) with the reactant fructose and 5-HMF molecules in a homogeneous environment. Its understanding is essential since the transformation of biomass to chemicals is typically facilitated by liquid-phase homogeneous acid-catalyzed reactions.

Finally, Chapter 4 addresses solvent's effects directly in humins formation chemistry. Here, we aim to compare and analyze three popular solvation approaches (implicit solvent, explicit equilibrated solvent molecules without dynamics and polarizability, and explicit solvent molecules with dynamics with polarizability and quantum mechanically treatment) to model the acid-catalyzed protonation 5-HMF, which is considered the first step for 5-HMF dehydration leading to humins. These three solvent models consider different aspects of solvent participation in chemistry and have different computational requirements. Molecular level understanding of the role of solvents on the reaction kinetics and thermodynamics is crucial to select/screen/design solvents and to optimize catalyst performance and the overall chemical transformation process.

The specific objectives addressed in each chapter are detailed in the following subsections.

1.6.1 Key objectives of chapter 2

- To compute the detailed reaction mechanisms, activation barriers, and reaction-free energies associated with all elementary reaction steps in the 5-HMF initiated, acid-catalyzed humins formation.

- To compare the kinetics and energetics for the early stages of humins formations along the different pathways to identify rate-determining steps and key intermediate reactions.

- To explain the subsequent chemistry for further polymerization after dimers formation leading to humins oligomeric compounds in agreement with the experimental spectroscopic data.

1.6.2 Key objectives of chapter 3

- To investigate the effect of aprotic solvent on the interaction between reactants and products with the catalyst (the hydronium ion) using fructose dehydration as a reaction model, 5-HMF (the model product) and fructose (the model reactant) in acidic water and water-DMSO mixtures.
- To elucidate the role of aprotic solvent on inhibiting humins formation by preferential interaction (physical solvation) with the reactive groups of 5-HMF over water and hydronium ions.
- To reveal the effect of DMSO addition on the interaction between hydronium ions and the polar groups of fructose to get insights about the enhanced reactivity of fructose dehydration in the presence of an aprotic solvent.
- To quantify the hydronium ion's relative stability in free energy in the first solvation shell of fructose and 5-HMF compared to the bulk of the solvent as a function of solvent composition.

1.6.3 Key objectives of chapter 4

- To systematically compare three molecular solvent modellings techniques such as implicit solvation, explicit equilibrated solvation without solvent dynamics and polarizability, and explicit solvation with solvent dynamics and polarizability.
- To reveal the effect of adding aprotic solvent on the homogeneous acid-catalyzed protonation in biomass chemistry since protonation is a critical step governing the kinetics of many acid-catalyzed reactions.
- To understand how the chemistry (kinetics and thermodynamics) of the 5-HMF degradation into humins is altered as a function of solvent proportion by adding DMSO in an aqueous reacting system.
- To quantify the changes of solvent reorganization energy and non-equilibrium solvation as a function of solvent composition (water-DMSO mixtures) in acid media and study the implications on the reaction energetics.

1.7 Organization of the thesis

This thesis includes three content chapters organized as follows. Every chapter has an introduction about the relevant literature on the state of the art and specifics of each research performed. Furthermore, the computational methodology, as well as the computational systems implemented, are described. The results and discussion are presented afterwards, and finally, each chapter concludes with a summary of the key findings.

1.7.1 Chapter 2: Mechanistic investigation into the formation of humins in acid catalyzed liquid phase biomass reactions.

In this chapter, density functional theory (DFT)-based calculations are performed to reveal the humins formation chemistry in detail and to explore the kinetic and thermodynamic feasibility of the reaction pathways. Studying the reaction pathways, activation barriers, and reaction-free energies associated with all elementary reaction steps in the 5-HMF-initiated, acid-catalyzed reactions leading to humins formation. The humins formation is initiated with the rehydration of 5-HMF to form DHH (key promoter of humins formation), followed by its keto−enol tautomerization, aldol addition, and condensation with 5-HMF. Within the implicit solvation approximation, the formation of the 5-HMF-DHH dimer is slightly endergonic, whereas the 5-HMF rehydration leading to DHH is thermodynamically downhill. This mechanistic understanding of initiation reactions for humins could help to design solvent and catalyst systems to deter their formation.

1.7.2 Chapter 3: Understanding the effect of solvent environment on the interaction of hydronium ion with biomass-derived species: A Molecular Dynamics and Metadynamics investigation.

This section is designed to analyze the role of solvents during a chemical reaction from the physical solvation perspective with the reacting species and the catalyst. To get a molecular understanding of the effects of the addition of certain co-solvents inhibiting the formation of humins during the dehydration of fructose to 5-HMF via preferential solvation of key functional groups, molecular dynamics (MD) simulations of 5-HMF (the model product) and fructose (the model reactant) in acidic water and water-DMSO mixtures are performed to analyze their interaction with the hydronium ions as a function of solvent composition. An essential factor in this context is the relative stability of the hydronium ion (the catalyst) near the biomass moieties compared to that in bulk, as it could determine its efficacy in the protonation step. Additionally, well-tempered metadynamics (WT-MTD) simulations are performed to determine the relative stability of the hydronium ion near fructose and 5-HMF, as compared to that in the bulk solvent phase, as a function of solvent composition. DMSO as a co-solvent improves the stabilization of the hydronium ions in the first solvation shell of fructose compared to that in the bulk solvent. On the other hand, hydronium ions become less stable near 5-HMF, as the concentration of DMSO increases in the reacting system.

1.7.3 Chapter 4: Energetics of acid catalyzed biomass reactions: How and why does the solvent model matter?

Condensed phase selective transformation of biomass to platform chemicals is susceptible to the solvent environment. Selectivities and conversions in several key acid-catalyzed biomass reactions change significantly by adding a co-solvent to the reacting system. Hence, when molecular modelling tools are implemented to investigate these reactions, selecting a solvent model that can compute reaction energetics accurately and capture and explain the "role" of solvent in altering them is necessary. Hence, this chapter compares popularly used implicit and explicit solvation approaches to model the protonation reaction of a

biomass-derived moiety, with and without reaction and solvent dynamics, using water and DMSO as co-solvents. This analysis reveals that unlike routinely performed DFT calculations in implicit solvents, where the free energy profiles do not change significantly due to changes in solvent composition, the reaction free energy using explicit equilibrated solvation without solvent dynamics was found to be ~ 67 kJ/mol higher at 75 % wt. DMSO compared to lower DMSO proportions. However, the effect of the solvent environment on free energy barriers was only captured when solvent and reaction dynamics were considered. The increase in activation-free energy with the addition of the co-solvent is due to changes in the solvent reorganization energy and non-equilibrium solvation.

1.8 Conclusions and future perspectives

The conclusions drawn from the key results of the four chapters are highlighted and perspectives on these are presented and summarized in Figure 1.9
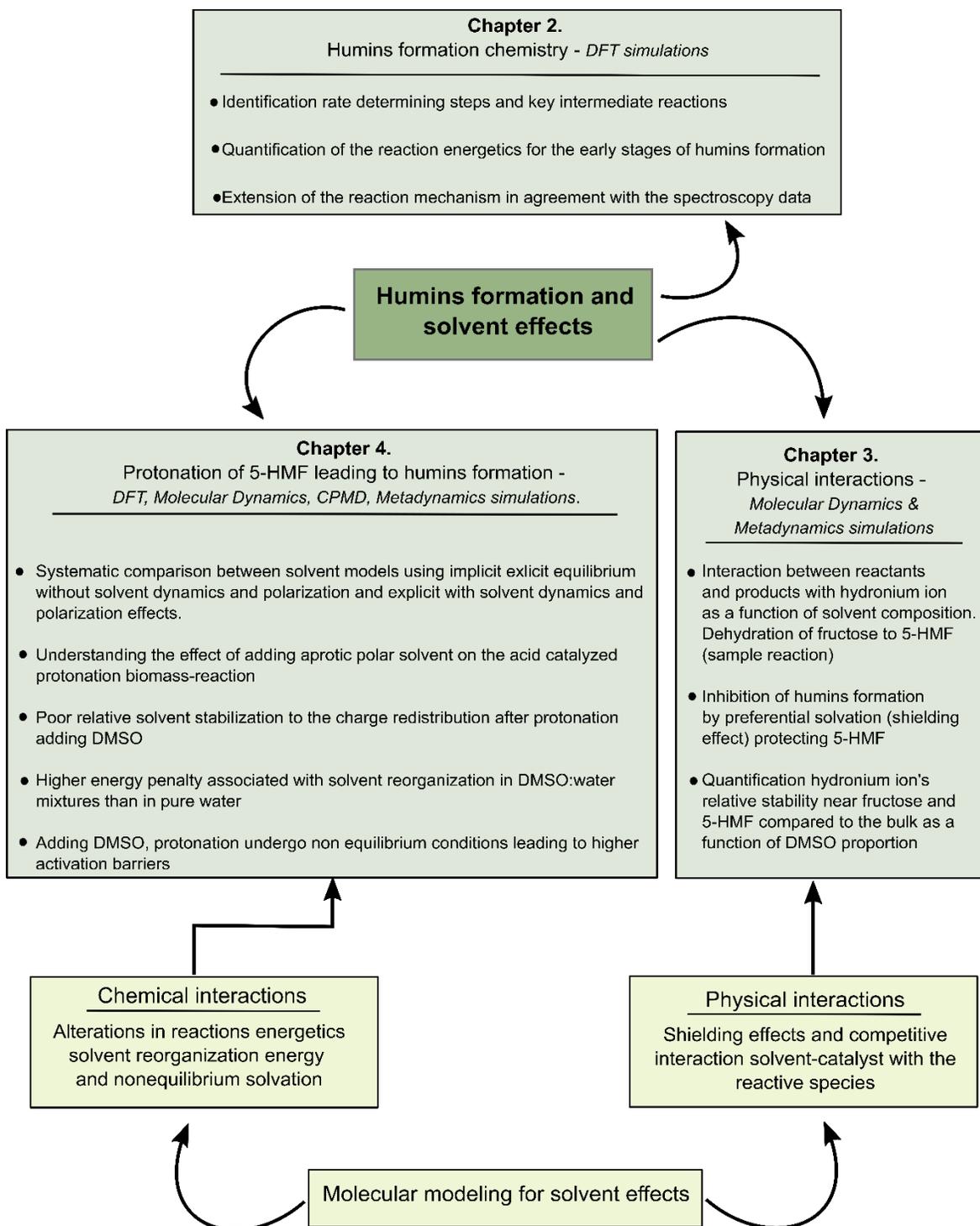
Figure 1.13 Summary of the thesis and organization.

Chapter 2 Mechanistic investigation into the formation of humins in acid catalyzed liquid phase biomass reactions.

2.1 Introduction

Investigating humins formation is challenging due to their complex and unknown molecular structure. Likewise, during the condensed phase catalytic conversion of biomass, several reactions occur simultaneously, making it difficult to isolate humins formation reactions. Since humins are polymeric and insoluble oligomers, it is challenging to determine their chemical structure and to identify the intermediates that can potentially lead to their formation. It has been proposed that humins formed during the catalytic dehydration reactions of cellulose derived biomass molecules to produce 5-HMF consists mainly of furan rings, aliphatic linkers and rich in functional groups such as carboxyl, carbonyl, and hydroxyl.[72] Moreover, Cheng et al. [107] reported that a certain fraction of humins are conglomerate of oligomers interacting weakly rather than being large macromolecules. Van Zandvoort et al. [62] characterized humins derived from glucose, fructose, and xylose in aqueous sulfuric acid. Based on FTIR spectroscopy, it was suggested that humins formation during the transformation of hexoses occurs via 5-HMF which further undergoes poly-condensation processes by electrophilic substitution with the formation of ether or acetal bonds between the furan rings. Whereas Dee and Bell[108] performed kinetic studies during the acid-catalyzed hydrolysis of cellulose, where glucose is the main intermediate species undergoing dehydration to 5-HMF. A corollary of that work was that the reduction of humins formation was attributed to the decrease of 5-HMF formation, suggesting that humins are formed due to the polymerization and condensation of 5-HMF and glucose. Patil and Lund[71] based on IR spectra of acid catalyzed humins formed from 5-HMF, fructose and glucose showed that the furan ring and the hydroxy methyl group of 5-HMF were present in humins, but the carbonyl group was not. Thus, the absence of the carbonyl group in humins suggested that 5-HMF is polymerized through its carbonyl group. Additionally, it was observed that humins were derived from 5-HMF and not from levulinic acid (LA) or formic acid (FA), which are also rehydration products of 5-HMF. Their findings are consistent with a mechanism proposed by Horvat[109] identifying 2,5-dioxo-6-hydroxy-hexanal (DHH) as the promoter of humins formation, which, in turn, is formed from the hydration of 5-HMF as shown in Figure 2.1. According to this, aldol addition and condensation are proposed as the main reaction pathways between DHH with 5-HMF or any carbohydrate molecule present in the reacting system to initiate humins formation. An overall pathway for the conversion of cellulose derived glucose to 5-HMF and subsequent degradation of 5-HMF to humins are other products is shown in Figure 2.2.

To summarize, there is sufficient experimental evidence to conclude that 5-HMF is a precursor for humins formation in the sugars dehydration reaction and to suggest that rehydration and addition-condensation reactions between 5-HMF and its derivatives initiate humins formation. Hence, to investigate the humins formation chemistry in detail and to explore the kinetic and thermodynamic feasibility of the reaction

pathways, density functional theory (DFT)-based calculations are performed in this paper. Individual elementary reaction steps are simulated, and associated activation free energy barriers and reaction free energies are computed. We aim to investigate aldol addition/condensation as a primary humins growth reaction, preceded by hydration of 5-HMF and its keto-enol tautomerization to DHH, considering this molecule as a key intermediate leading to humins formation. We also endeavor to get an insight into the effect of the solvation environment on the energetics of humins formation, using implicit water and DMSO solvent environments.
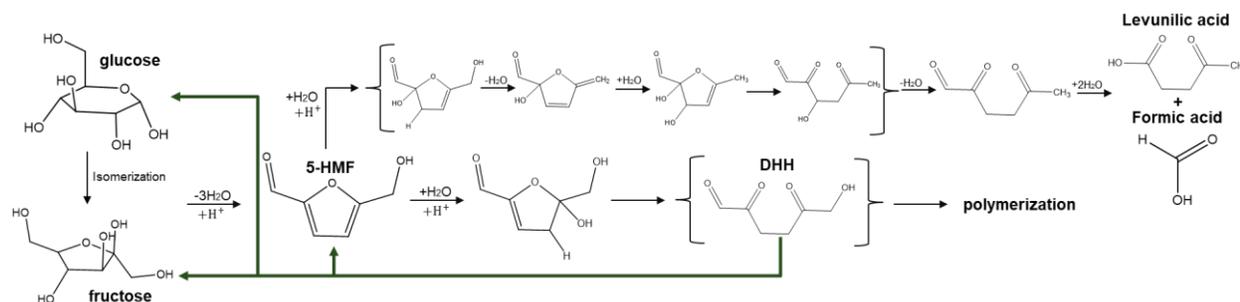
Figure 2.1 Glucose conversion pathway to 5-HMF and further rehydration of 5-HMF to LA and DHH, further leading to the formation of humins.

2.2 Computational methodology and simulation systems

DFT calculations are performed using Gaussian 09 code.[110] The hybrid functional RM06-2X is used in combination with 6-31+G (2d, p) basis set. This theory level is deemed sufficient to study the chemistry of cellulose derived molecules. [111,112] For the geometry optimization and transition state (TS) calculations, no constraints are applied on the atoms. The TS search is performed using Berny algorithm, as implemented in the Gaussian code. Geometry optimization is accompanied by frequency calculations at the same level to distinguish the optimized structures as the TS and the ground state, indicated by the presence and absence of an imaginary frequency (which corresponds to the reaction coordinate), respectively. Intrinsic reaction coordinate (IRC) calculations are performed to confirm that the obtained TS is linked with the anticipated reactant and product on the potential energy surface. Polarizable Continuum Model (PCM) using the integral equation formalism variant (IEFPCM) is considered during all geometry optimization calculations and TS search. Implicit solvation in the continuum model is implemented using the dielectric constant (ε) of DMSO and water. Reaction free energies in gas phase and in the implicit solvent phase are reported at 433 K. Self-consistent field (SCF) selected for finding electronic structure configuration was tight equivalent to 1.00D-06 in convergence in energy change.

The Zundel cation[113], which is an important and simplified structure for the solvated proton $H_5O_2^+$, was chosen as the acid catalyst.[114] A conformational analysis was carried out to determine the ground state of the initial system formed by a DHH molecule interacting with a Zundel structure. All energies are reported in kJ/mol.
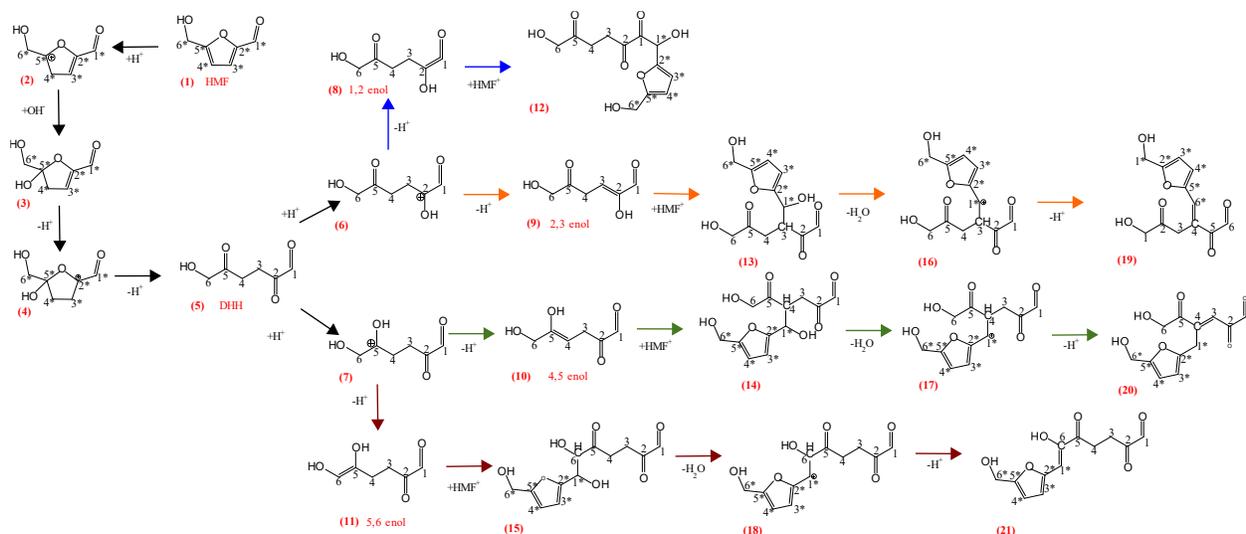
Figure 2.2 Schematic illustrating reaction pathways for humins formation through 5-HMF rehydration to DHH, DHH keto-enol tautomerization and further dimerization reaction with 5-HMF. Carbon atoms are numbered throughout the scheme and each reaction intermediate, and the elementary reaction step are numbered. Post DHH formation, reaction pathways corresponding to 1,2 enol, 2,3 enol, 4,5 enol and 5,6 enols are color coded using blue, orange, green and red arrows, respectively.

2.3 Results and discussion

Reaction steps involved are: (i) The rehydration of 5-HMF to DHH. (ii) the keto-enol tautomerization of DHH, resulting in the formation of four enols 1,2 enol; 2,3 enol; 4,5 enol and 5,6 enols. (iii) Subsequently, each enol reacts with the protonated 5-HMF (5-HMF⁺) via aldol addition. (iv) Except for the dimer derived from the aldol addition reaction between 1,2 enol and 5-HMF⁺, the rest of the dimers further undergo a two-step condensation mechanism. All the reaction steps and intermediate species are numbered in Figure 2.2 and referenced throughout the manuscript.

2.3.1 Rehydration of 5-HMF to form DHH.

The hydration of 5-HMF to DHH follows the mechanism described by Markonikov's electrophilic addition.[115, 116] Figure 2.3 shows the detailed reaction mechanism of DHH formation from 5-HMF, corresponding to step 1 to step 1.3 in Figure 2.2. Figure 2.4 shows the corresponding free energy profile. Step 1 corresponds to the protonation at C_{4*} of 5-HMF generating an electrophilic site at C_{5*}. Further, protonation at C_{4*} leads to the formation of intermediate 2. This proton addition at C_{4*} is the rate determining step for the rehydration of 5-HMF to DHH with a free energy barrier of ~93 kJ/mol, in both, water and DMSO implicit solvation. Further, hydroxylation at C_{5*} leads to the formation of intermediate 3. Subsequently, C_{3*} is protonated to form the intermediate 4, promoting furan ring opening. Finally, the -OH group at C_{5*} is deprotonated leading to a carbonyl group, regenerating the acid catalyst, and forming DHH, as species 5. The free energy profile reveals that it is an endergonic reaction, with ΔG_R slightly more than 30 kJ/mol, in both DMSO and water implicit solvation.
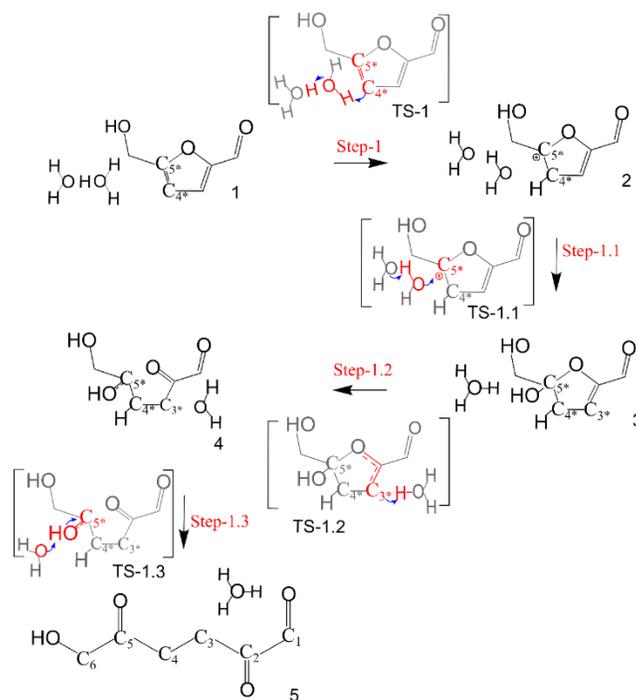
Figure 2.3 Hydration reaction mechanism of 5-HMF leading to DHH formation. The blue arrows in the molecular structures indicate electron flow. Transition states of the reactions are depicted using square brackets and atoms not taking part in the reaction are faded. Indices of the species and carbon atoms numbering correspond to those in Figure 2.2
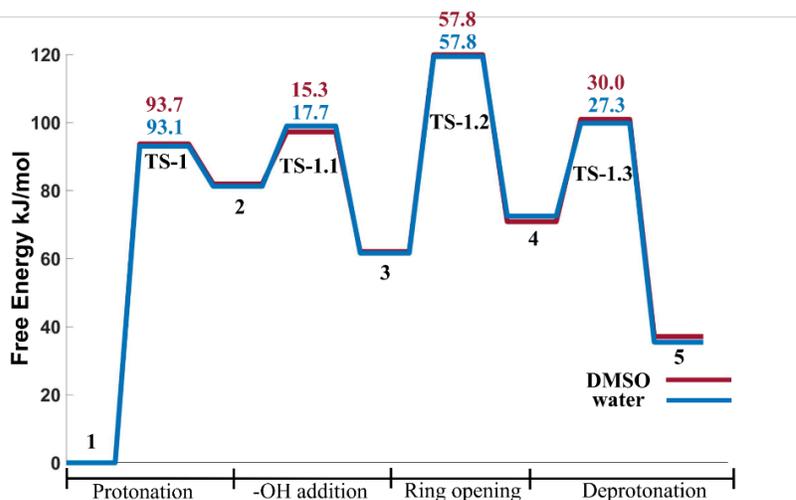
Figure 2.4 Free energy profile for the rehydration of 5-HMF to form DHH. Free energies barriers for each intermediate reaction for water and DMSO are shown in blue and red, respectively. Free energy values are reported in kJ/mol

26

2.3.2 Keto−Enol Tautomerization of DHH

Keto−enol tautomerization is often involved in the transformation of biomass molecules, such as carboxylic acids and olefins, or the transformation of levoglucosan (LGA) to produce levoglucosenone (LGO), to name a few. DFT calculations have been implemented to investigate these reactions in the past. The reaction mechanisms for the keto−enol tautomerization of DHH are shown in Figure 2.5. The keto-enol tautomerization of DHH takes place via a two-step mechanism. Firstly, DHH is protonated at carbonyl groups at C_2 or C_5 carbon, leading to intermediates 6 and 7, respectively. When the carbonyl group at C_2 is protonated, intermediate 6 is produced forming an electrophilic site at C_2. Subsequently, the deprotonation at C_1 or C_3 takes place leading to the formation of 1,2 or 2,3-enol respectively. On the other hand, when the carbonyl group at C_5 is protonated the intermediate 7 is produced, followed by the deprotonation at C_4 or C_6, leading to the formation of 4,5 or 5,6-enol, respectively. This step restores the Zundel structure which initially yielded proton to initiate the keto-enol tautomerization of DHH. The free energy profile for all four possible keto-enol tautomerizations of DHH is shown in Figure 2.6. The free energy of an isolated 5-HMF is added to allow us to map the complete free energy profile with the subsequent bimolecular reactions. It can be observed that the protonation leading to 4,5-enol and 5,6-enol is kinetically more favorable leading to intermediate 7 compared to the protonation leading to 1,2 and 2,3-enols through the formation of intermediate 8. Further, it is observed that the TS-3a corresponding to deprotonation at C_1 in step 3a leading to 1,2-enol is the least favorable compared to other alpha-hydrogen transfer of enols studied. This is because the double bond between C_1 and C_2, just beside the carbonyl group, makes the molecule rigid and thus unlikely to occur compared to another possible deprotonations.
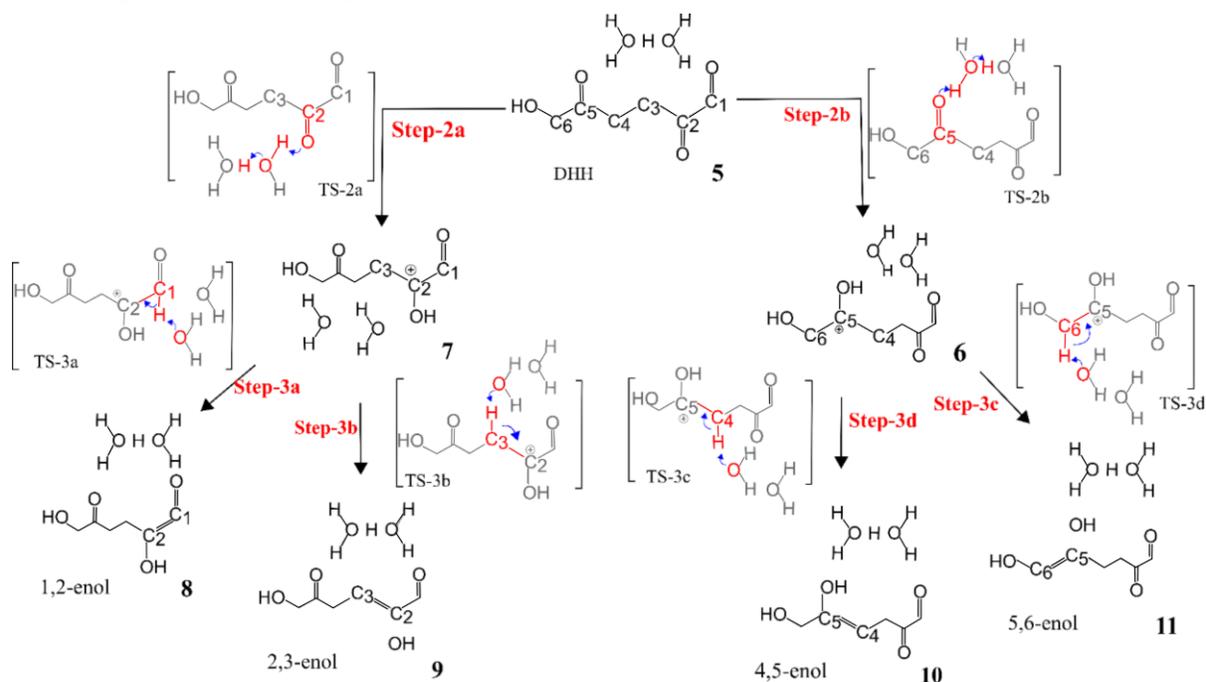
Figure 2.5 Reaction mechanism for the formation of the four possible enols from DHH. Blue arrows in the molecular structures indicate electron flow. Transition states of the reactions are depicted using square brackets and atoms not taking part in the reaction are faded. Indices of the species and carbon atoms numbering correspond to those in Figure 2.2
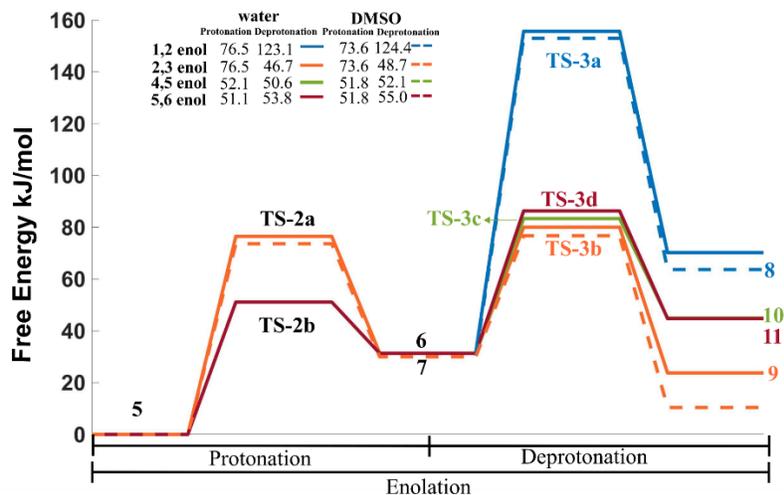
Figure 2.6 Free energy profile for the keto-enol tautomerization reactions. Activation free energy barriers in kJ/mol are also shown in the legend. The labeling of intermediates and transition states is the same as in Figure 2.2

Figure 2.7 shows the transition state configurations for the two possible protonation reactions of DHH leading to intermediates 6 and 7. In Figure 2.7 a, it can be seen for TS-2a that the hydroxyl group close to the protonated carbonyl oxygen attached to C_2 stabilizes the water cluster. While, for TS-2b in Figure 2.7 b, the two water molecules are well stabilized with the carbonyl group at C_6. It is suggested that this difference in the intermolecular interaction between the DHH and the Zundel structure causes the difference in activation free energies between the two possible protonation of DHH.
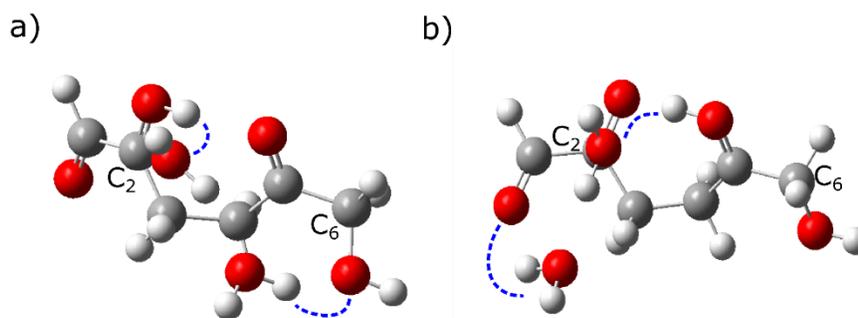
Figure 2.7 Transition states correspond to the protonation reaction of the keto-enol tautomerization of DHH a) protonation of the C2 carbonyl group of DHH leading to intermediate 6 b) protonation of the C5 carbonyl group of DHH leading to intermediate 7. The blue dotted arrows indicate the

intermolecular interactions between the Zundel structure and DHH polar groups. The arrows are oriented according to the direction of electron flow.

2.3.3 Aldol-Addition Reaction between DHH and 5- HMF

5-HMF becomes electro deficient when the carbonyl group is protonated leading to a nucleophilic addition with any enol available. The aldol addition reaction between 5-HMF$^+$ and the DHH derived enols has been found to take place in a concerted manner. Figure 2.8 shows 5-HMF protonation and four-aldol addition reactions for each enol derived from DHH. In the case of 1,2-enol, the -OH bonded to C_2 is deprotonated to restore the Zundel structure while C_2 is bonded to C_{1*} to form the dimer. Similarly for 2,3-enol, the same –OH group bonded to C_2 is deprotonated but in this case C_3 becomes electro deficient reaction with C_{1*} forming the corresponding dimer shown in Figure 2.8 c. In the case of 4,5 and 5,6-enols the -OH bonded to C_5 is deprotonated, while C_4 and C_6 now become nucleophilic carbons reacting with C_{1*} of 5-HMF respectively. Finally, the dimers are formed with a single carbon-carbon bond and a hydroxyl group present at the β-carbon atom, as shown for all the reactions.

The free energy barriers for the aldol addition reaction between 5-HMF$^+$ and the enols are reported in Figure 2.9. The barriers for aldol addition reaction of all enols with the HMF are between 90 and 100 kJ/mol. These values resemble the apparent activation energies of 94 ± 8.8 kJ/mol calculated by Patil et al.[69] from kinetic data analysis for the rehydration of 5-HMF at 118 - 135 °C starting with 0.1 M in acidic medium. Aldol addition reaction is identified to be the rate-determining step with the highest energy barriers along the entire reaction pathway. Moreover, no significant differences in the barriers were found between DMSO and water as implicit solvent.
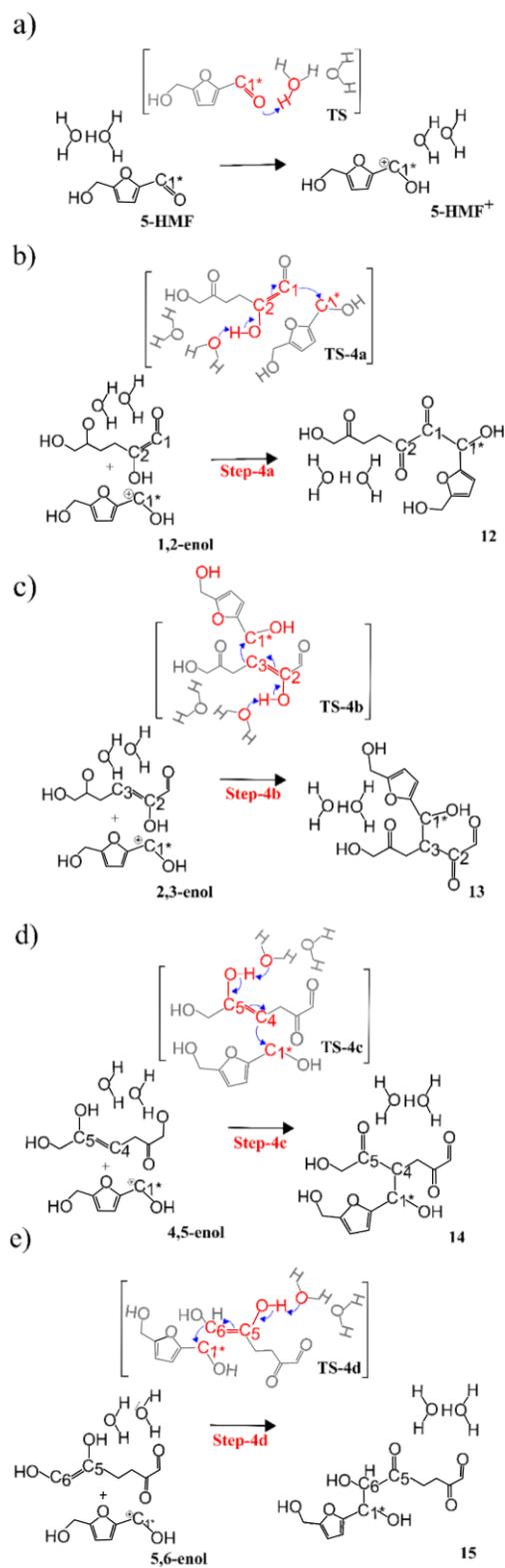
Figure 2.8 Reaction mechanism of the aldol addition reactions of a) 1,2 enol b) 2,3 enol c) 4,5 enol d) 5,6 enol with 5-HMF+. The blue arrows indicate electron flow during the reactions. Transition

states of the reactions are depicted using square brackets and atoms not taking part in the reaction are faded. Indices of the species and carbon atoms numbering correspond to those in Figure 2.2.
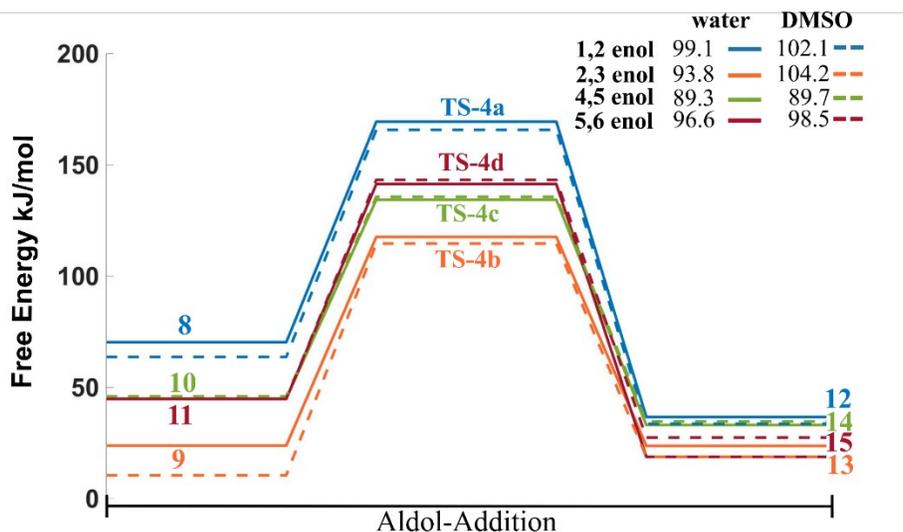
Figure 2.9 Free energy profile for Aldol-Addition reactions. The numbers in the legend indicate the activation of free energies. All energy values are reported in kJ.mol−1. The labeling of intermediates and transition states is the same as in Figure 2.2

2.3.4 Condensation of Dimers Derived from Enols and 5-HMF Aldol Addition

This reaction proceeds in two steps: loss of hydroxyl group and the subsequent deprotonation of the dimer leading to a C-C double bond, thus producing a water molecule. This reaction only occurs for the dimers 13; 14 and 15 due to the presence of alpha hydrogens (C atom which is alpha to the newly formed C=O group in the dimer) In the case of the dimer derived from 1,2-enol, the reaction does not take place due to the absence of an alpha hydrogen.

2.3.5 Dehydration Intermediate Reaction

Dehydration is the first step for the dimer. Figure 2.10 describes the detailed reaction mechanism for the dehydration of the dimers leading to the formation of humins precursors. The dehydration takes place in a single step involving the protonation of the -OH group attached to C_{1*} and the loss of the hydroxyl group from the dimer.
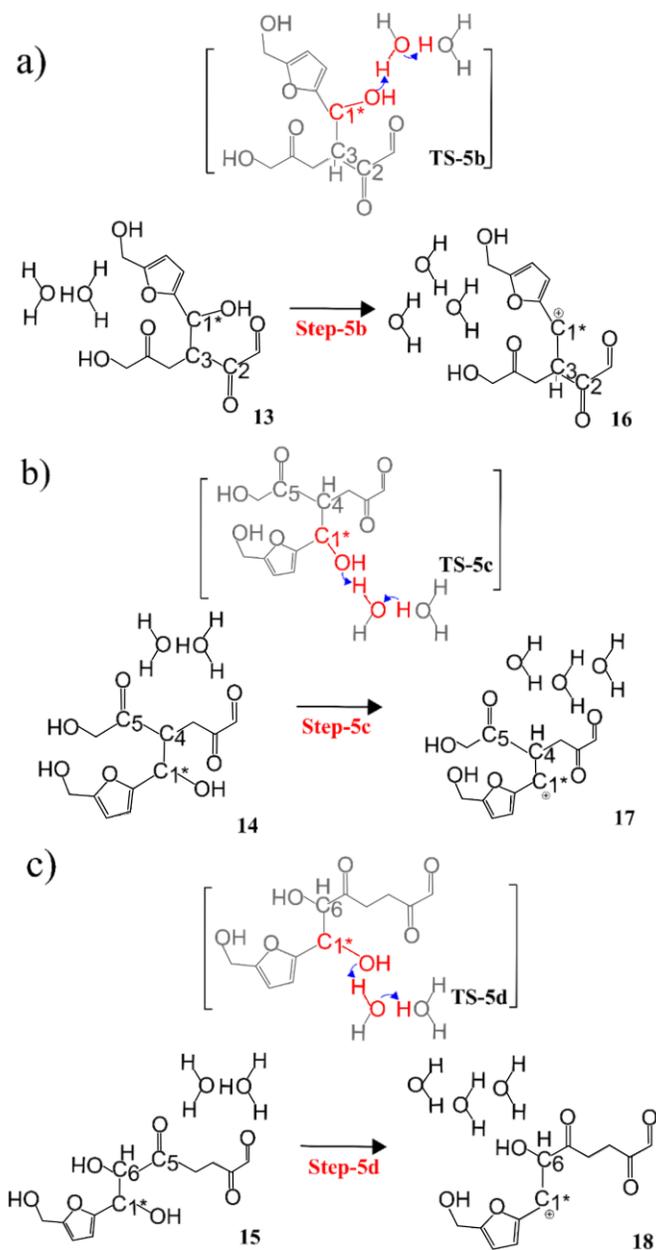
Figure 2.10 Reaction mechanism of the dehydration reactions for a) 2,3-enol derived dimer b) 4,5-enol derived dimer c) 5,6-enol derived dimer. Transition states of the reactions are depicted using square brackets and atoms not taking part in the reaction are faded. Indices of the species and carbon atoms numbering correspond to those in Figure 2.2.

Energetics of the dehydration reaction are shown in Figure 2.11. The activation barrier is highest for the dehydration of the dimer derived from 4,5-enol. On the other hand, the kinetically most favorable dimer dehydration is that of derived from 5,6-enol. Additionally, the difference in free energy barriers between water and DMSO implicit solvation, for 4,5 and 5,6-enols, is approximately 20 kJ/mol.
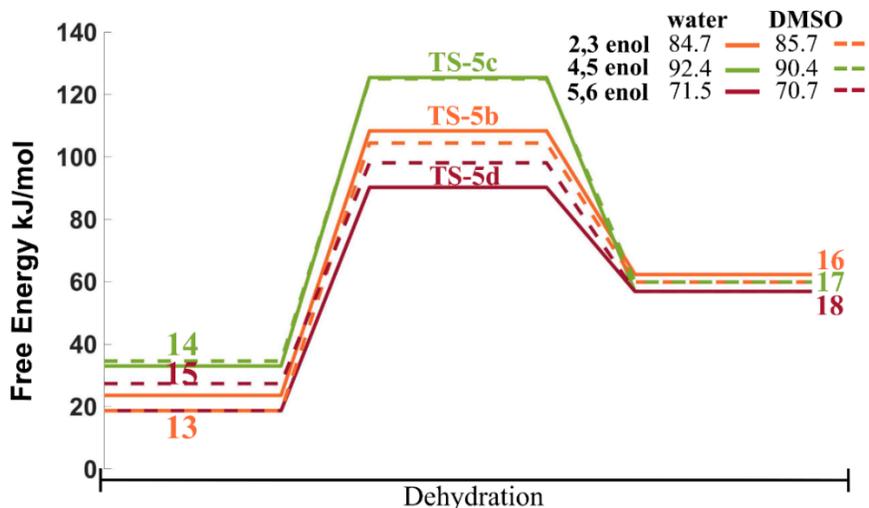
Figure 2.11 Free energy profile for the dehydration reactions of the dimers derived from 2,3 4,5 and 5,6 enol aldol addition reaction with 5-HMF. The numbers in the legend indicate the activation of free energies. All energy values are reported in kJ.mol−1. The labeling of intermediates and transition states is the same as in Figure 2.2.

As previously mentioned, the intermolecular interaction between the dimer and the Zundel structure can play a major role in the energetics of the reaction. Figure 2.12 shows the transition states for the three different dehydration reactions. The dehydration of the dimer derived from 5,6-enol has the lowest activation barrier because the -OH groups bonded to C_6 and C_1 stabilize the Zundel structure in the transition state. This does not occur in the case of the other dimers, because of the lack of -OH group in the immediate vicinity in the transition state. However, for 2,3-enol the carbonyl group at C_2 also stabilizes the water cluster, as shown in Fig 2.12 a.
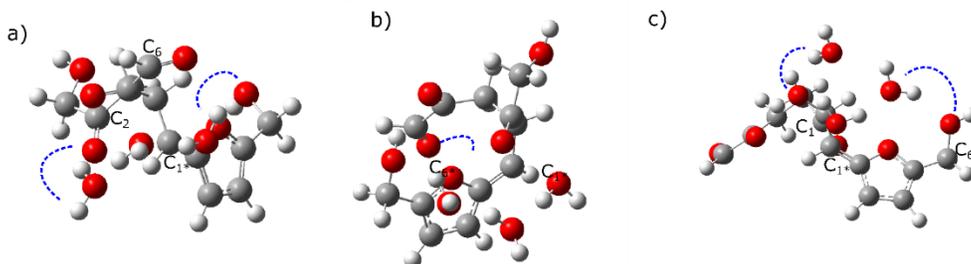
Figure 2.12 Transition states corresponding to the dehydration reaction of a) 2,3-enol-derived dimers, b) 4,5-enol-derived dimers, c) 5,6-enol-derived dimers. The blue dotted arrows indicate the intermolecular interactions between the Zundel structure and the dimer's polar groups. The arrows are oriented according to the electron flow.

2.3.6 Deprotonation as the Final Reaction Step

The last step in the reaction mechanism is the deprotonation of the dimer to restore the Zundel structure. It is observed in Figure 2.13 for the case of the dimer 16 (derived from 2,3-enol) that C_3 loses the only hydrogen attached to it forming a double bond with C_{1*}. The same occurs for the rest of the dimers 17 and 18 (derived from 4,5 and 5,6-enol where C_4 and C_6 respectively are deprotonated.
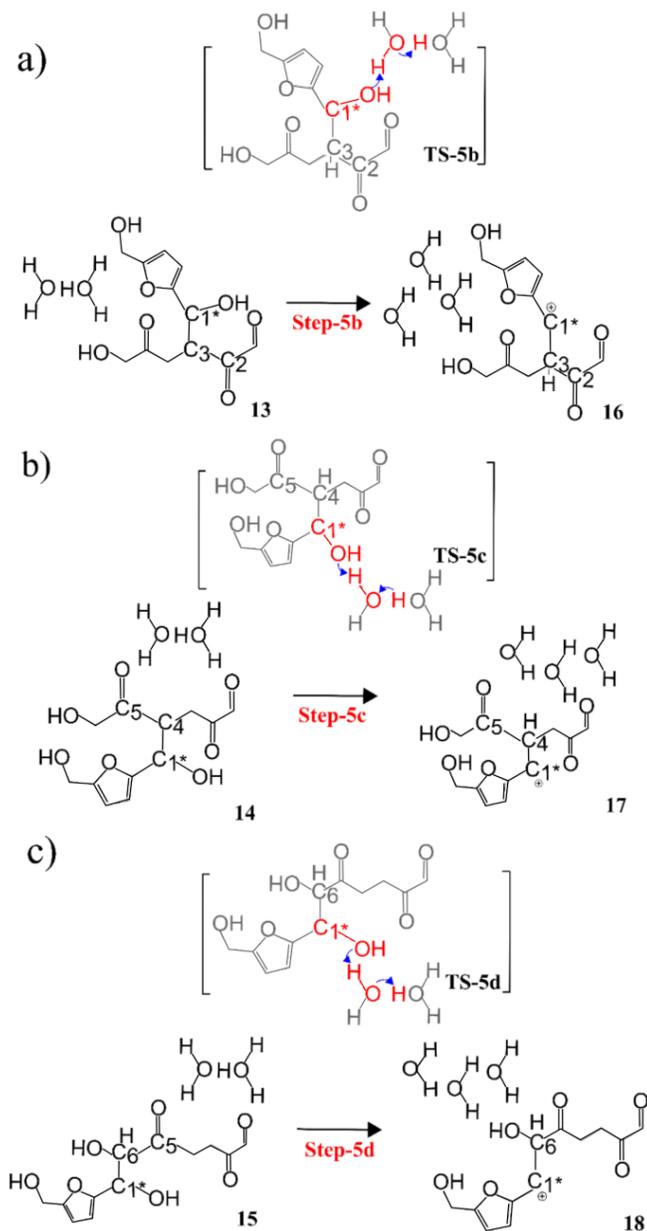
Figure 2.13 Reaction mechanism of the deprotonation reactions of a) 2,3 enol derived dimer b) 4,5 enol derived dimer c) 5,6 enol derived dimer. The blue arrows indicate electron flow during the reactions. Transition states of the reactions are depicted using square brackets and atoms not taking part in the reaction are faded. Indices of the species and carbon atoms numbering correspond to those in Figure 2.2.

The free energy barriers corresponding to the deprotonation reaction are shown in Figure 2.14. A relatively higher energy barrier is required to deprotonate the dimer derived from 4,5 enol. Furthermore, no significant differences are observed between the free energy values of the reaction in water and DMSO implicit solvation.
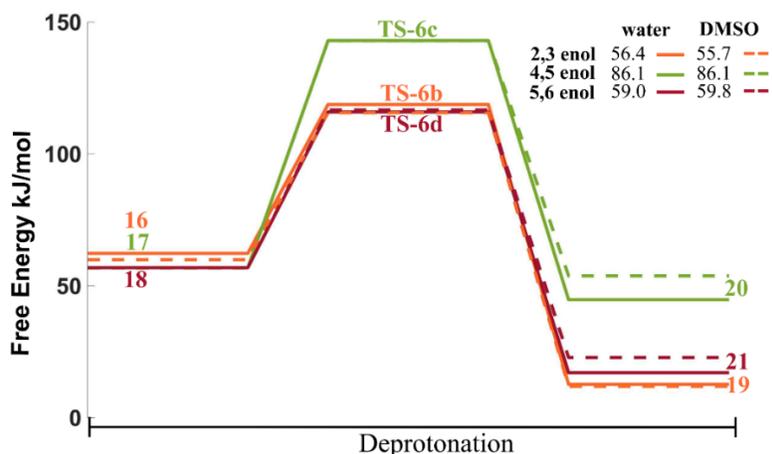
Figure 2.14 Free energy profile for the deprotonation reactions. The numbers in the legend indicate the activation of free energies. All energy values are reported in kJ.mol−1. The labeling of intermediates and transition states is the same as in Figure 2.2.

The transition state structures for the deprotonation of the three dimers are shown in Figure 2.15. It is noticeable that during the deprotonation of the dimer derived from 4,5-enol, water molecules of the Zundel structure only have one polar group to be stabilized, which is the -OH bonded to C_{6*}. The absence of more polar groups with which the water molecules can interact leads to higher energy barrier. On the other hand, for the cation 16 (dimer derived from 2,3-enol), the water molecules are not only stabilized by the -OH bonded to C_{6*} but also by the carbonyl group at C_1. In the same way, the cation 18 (dimer derived from 5,6-enol) has two -OH groups bonded to C_6 and C_{6*} with which the Zundel structure is well stabilized.
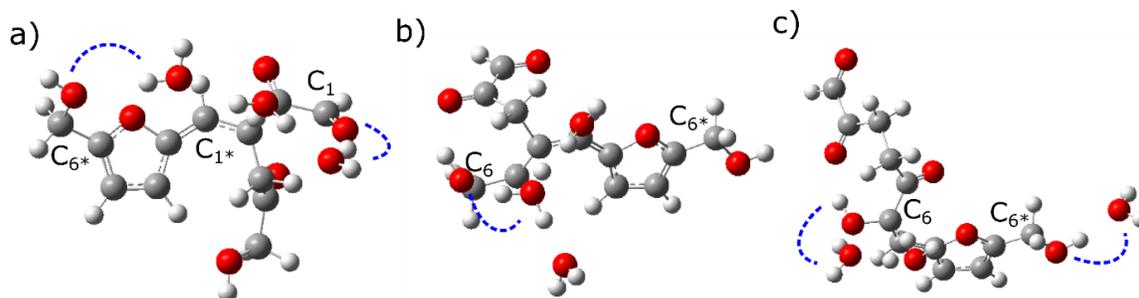
Figure 2.15 Transition states correspond to the deprotonation intermediate reactions of a) 2,3 enol-derived dimers, b) 4,5 enol-derived dimers, c) 5,6 enol-derived dimers. The blue dotted indicate the intermolecular interactions between the Zundel structure and the dimer's polar groups. The arrows are oriented according to the electron flow.

2.3.7 Complete Free Energy Profile

The free energy profile for the complete reaction pathway presented in Figure 2.2 is shown in Figure 2.16. Aldol-addition reactions have the highest activation free energy barriers in all the pathways except for the one in which 1,2-enol is involved. These free energy profiles also give us a thermodynamic perspective of the reactions. In general, reactions corresponding to steps 2, 3 and 5 (protonation, deprotonation & dehydration) show positive reaction free energies, while the aldol addition reaction as well as the final deprotonation of the dimers are thermodynamically downhill. However, the overall reaction is slightly endergonic. These results agree with the values reported by Patil et al [68]. This could be the reason that condensed-phase biomass- reactions that lead to humins formation are usually carried out at temperatures above 100 °C. Though there is strong experimental evidence that modifying the solvent medium (water vs. DMSO-water mixtures) alters the yield of humins, the implicit solvation model does not reveal significant differences in both, the kinetics and thermodynamics of the reactions leading to humins formation. Further investigation of the rate determining aldol addition reaction step, with explicit solvent molecules in condensed phase is needed. By identifying the elementary reaction steps and associated activation barriers, the present work laid the foundation for the same.
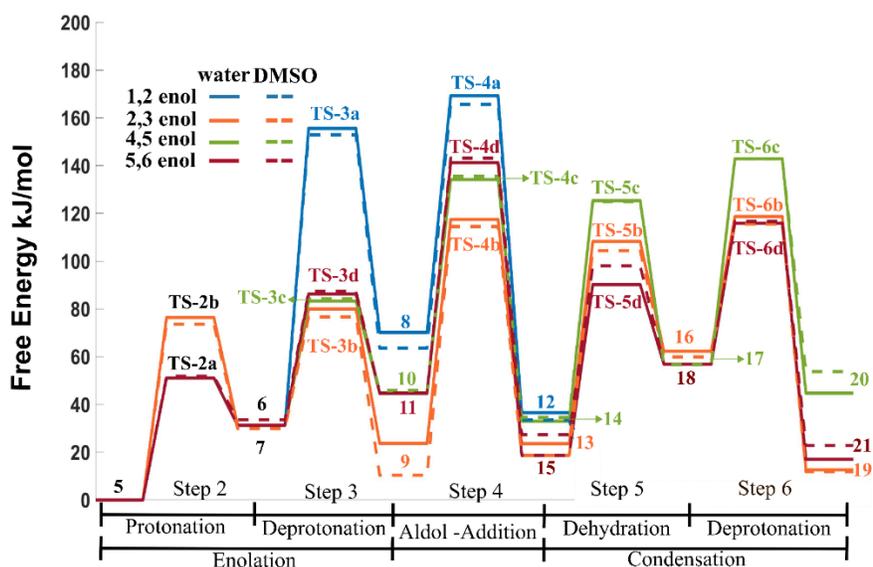
Figure 2.16 Free energy profile for all the reactions leading to the formation of the final dimers, as precursor for humins formation. The activation free energies are calculated with respect to DHH interacting with Zundel structure and an isolated 5-HMF as the reference. The labeling of intermediates and transition states is the same as in Figure 2.2.

2.3.8 Extended Reaction Mechanism for the Formation of Oligomers

The mechanistic understanding of the reaction mechanisms developed in this work allows us to predict the subsequent steps for the generation of larger oligomers, as previously pointed out by Patil et al[68]. Figure 2.17 shows the reaction mechanism extended from the dimer derived from 5,6 enol, as the pathway via this

enol has the lowest free energy barrier and is thermodynamically most facile. This dimer can further react with 5-HMF molecules until it becomes saturated, forming a pentamer. Once this macromolecule can no longer react with another 5-HMF molecule, it can be acid attacked to generate a nucleophilic carbon again (as initially occurred with a 5-HMF molecule) and further react with another enol available in the reaction medium. It can also subsequently undergo acid catalyzed keto-enol tautomerization.

This suggested mechanism supports aldol-addition and condensation as the main reaction for humins formation. The molecular structure corresponding to the oligomer shown in Figure 2.17 reveals that as more 5-HMF molecules react, their carbonyl functional groups are lost, forming repeated double bonds with carbons initially belonging to the enol molecule. Experimentally, it has been observed through IR spectra of humins derived from 5-HMF, that the peak corresponding to the carbonyl group of 5-HMF is absent in humins. On the other hand, the IR of humins derived from 5-HMF reveals that the structure of humins is rich in furan rings, which aligns with the structure of the oligomer shown in Figure 2.17.

5-HMF-derived macromolecules, as shown in Figure 2.17, would constitute humins observed in biomass reactions. However, it is important to highlight that humins derived from other biomass moieties, could be present too. Experimentally it has been shown that the aromatic ring of other organic molecules (i.e., benzaldehyde) added to the reaction system gets incorporated into the structure of humins, while the aldehyde group is lost. Additionally, benzaldehyde alone did not generate humins. These experimental observations suggest that regardless of the initial species, the aldol addition mechanism would be the main pathway leading to humins formation.
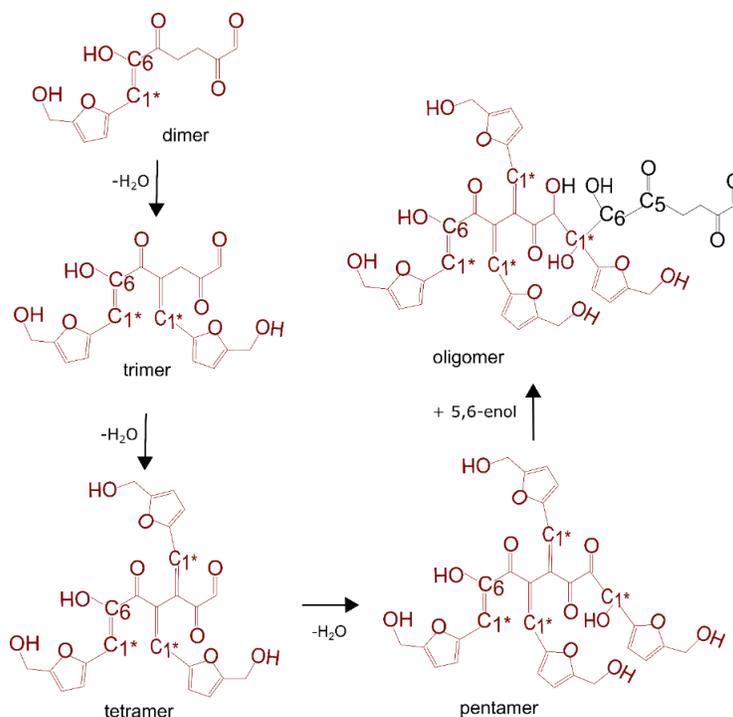
Figure 2.17 Extended reaction mechanism for oligomeric humins formation from dimer derived from 5,6 enol and 5-HMF.

2.4 Conclusions

Pathways for acid catalyzed reactions, initiating the formation of humins, from 5-HMF, are simulated using density functional theory (DFT) and reaction free energies and activation free energy barriers are computed. Elementary reaction steps involve hydration of 5-HMF to 2,5-dioxo-6-hydroxy-hexanal (DHH), followed by keto-enol tautomerization of DHH and aldol addition condensation reaction of the enols with 5-HMF. Depending on the location of the carbonyl group, 4 different enols can form from DHH. The rate determining step in this mechanism is the aldol addition reaction between the DHH-derived enols and 5-HMF, except for the pathway involving 1,2-enol, where the activation barrier is the highest for the keto-enol tautomerization step. Rehydration of 5-HMF to DHH (key promoter to humins formation) is endergonic. Moreover, the overall reaction from the keto-enol tautomerization of DHH to the formation of the dimers is found to be slightly endergonic. The initiation reaction pathway computed in this work explains the subsequent chemistry of further polymerization of the dimer by reacting with 5-HMF, to form humins. The computed mechanism and suggested polymerization scheme are in excellent agreement with the experimental spectroscopic evidence. Implicit solvation model in DFT computations did not reveal significant difference in reaction energetics for water and DMSO. Thus, to evaluate the effect of solvent on kinetics and thermodynamics of humins formation, key reaction steps identified in this work need to be simulated with explicit solvent molecules, in the condensed phase, and possibly with finite temperature dynamics.

Chapter 3 Understanding the effect of solvent environment on the interaction of hydronium ion with biomass derived species: A Molecular Dynamics and Metadynamics investigation.

3.1 Introduction

It is well known that the addition of an aprotic solvent like DMSO in the reacting system increases the reaction rate and improves the selectivity of 5-HMF from less than 50 % to slightly more than 90 % at high conversions. [103] In a biphasic reacting system, DMSO is also beneficial due to improved extraction of 5-HMF from the reacting phase to the extracting phase, further enhancing HMF selectivity.[117] However, the energy required to separate 5-HMF from a high boiling point solvent such as DMSO increases the production cost. Hence, a better understanding of the role of DMSO in the dehydration of biomass species to 5-HMF is needed, to identify other potential low boiling solvents. Besides, these findings are also relevant to improve the design of novel porous heterogeneous catalysts with functionalities that can mimic the solvent effect.[118] Tsilomelekis et al[119] found that solvation of 5-HMF with DMSO reduces the likelihood of a nucleophilic attack due to increase in the LUMO energy of 5-HMF, compared to that being solvated by water molecules. In addition to this, there are multiple explanations about the possible role of DMSO in the increased conversion and higher selectivity. Firstly, it has been suggested that fructose is predominantly present in the furanose form in aprotic solvents such as DMSO.[120] Hence, the rising proportion of furanose in DMSO at high temperature benefits HMF formation.[121] In addition, Quantum Mechanics/Molecular Dynamics simulations for the dehydration of fructose to HMF correlate the reorganization of solvent molecules to the high activation energies.[122] Amarasekara et al.[123] have also shown that DMSO can catalyze the dehydration of fructose to 5-HMF. Melmer et al.[124] suggested that the presence of aprotic solvents in the reacting system stabilize the transition state, speeding up the reaction rates.[125] Mushrif et al.,[126] based on molecular dynamics calculations, showed that DMSO improves the interaction between the fructose hydroxyl groups and water molecules, promoting the dehydration of fructose to HMF. Additionally, via, preferential solvation, DMSO protects the polar groups of 5-HMF and inhibits the formation of rehydration and condensation products. Similar solvation effects are also reported for other aprotic solvents such as tetrahydrofuran (THF) and N, N-dimethylformamide (DMF) in the dehydration of glucose to 5-HMF.[98] Acid-catalyzed reactions in aqueous medium with polar aprotic co-solvents have also been studied by a combination of molecular simulations and machine learning. Spatial configurations of the solvent around the reactant were used to predict the solvent effect on the reactions.[127] Kim et al [67] developed a multivariate model using different descriptors based on aprotic solvent's properties to predict yields for the dehydration of glucose to HMF. Even though the reaction mechanisms are not explicitly modeled, these techniques show significant advantages in computational efficiency and allow the optimization of reaction conditions to maximize the production of 5-HMF. These studies provide outstanding insights into the solvation effect of aprotic solvents such as DMSO on the reactivity of carbohydrates and furans. However, to the best of our knowledge, no previous work has investigated the effect of aprotic solvents on the interaction of reactants and products with the acid catalyst, i.e., the hydronium ion. There is sufficient experimental evidence that

indicates an increase in the activity of acid catalysts in the presence of aprotic solvents.[117] In heterogeneous catalysis, the interaction of the reactant with the catalyst and the competition between the reactant and the solvent to interact with the catalyst surface is believed to affect the reactivity. Similarly, here we are looking at the interaction between the catalyst (hydronium ion) and the reactant to determine if the presence of DMSO in the reacting system influences it. This has never been studied in homogeneous catalysis and we believe that its understanding is important since the transformation of biomass to chemicals is typically facilitated by liquid-phase homogeneous acid-catalyzed reactions. In the present paper, we perform molecular dynamics (MD) simulations of fructose-5-HMF systems in water and water–DMSO mixtures, with hydronium ions. We aim to investigate the structure and local ordering of hydronium ion, water and DMSO around fructose (reactant) and 5-HMF (desired product) to gain insights into the effect of solvation. Furthermore, to quantify the relative stability of the hydronium ion near fructose/HMF compared to being in the bulk solvent, free energy surfaces, as a function of selected coordinates, are computed using Well-tempered metadynamics (WT-MTD) simulations. Simulation details are discussed in Section 2 and the output of the MD simulations and WT-MTD are analyzed and discussed in Section 3. We conclude our findings in Section 4.

3.2 Computational methodology and simulation systems

Force field-based MD and WT-MTD are performed to study the effects of DMSO on the interaction of fructose and HMF with the hydronium ion. MD simulations are performed in a canonical NVT ensemble at 298 K using GROMACS Version 4.6.7.[128] Force-field parameters for all the species in the present work are given in Table S1 in the supporting information (SI). TIP3P model is used for water and OPLS-AA force-field is used to model DMSO, fructose and HMF molecules.[129] OPLS-AA has been extensively tested for lliquid phase simulations and successfully employed in predicting the strong hydrogen bonding between the hydroxyl groups of carbohydrates and DMSO sulfonyl group.[126,98] The OPLS–AA force field has also been employed by Pomata et al.[130] to study the dynamics of fructose solutions over a wide range of sugar concentrations and has been able to capture the variation in the mobility of water molecules as a function of fructose concentration. The temperature of the simulation system is maintained constant by Noose Hoover thermostat, with coupling time constants of 1 ps.[131] MD trajectories were analyzed using the Visual Molecular Dynamics (VMD), version 1.9.3.[132]

All simulations consisted of 0.1 ns of steepest descent energy minimization to bring the energy gradient below 10 kJ/mol nm, 0.4 ns of equilibration runs, 0.4 ns of isothermal-isobaric runs followed by 12 ns of production runs with a single MD time-step of 1 fs. All the simulations are performed using periodic boundary conditions in all directions. The cut-off of 0.7 nm is used with Particle-Mesh Ewald (PME) electrostatics. The simulation systems consisted of water and water-DMSO mixtures in the presence of hydronium ions, at experimentally relevant concentrations. We studied systems with 1 and 10 % wt. of fructose and HMF, since experimental investigations have shown that fructose conversion and HMF selectivity vary with

fructose concentration, with DMSO concentrations ranging from 0 to 80 %wt. The number molecules and the simulation cell size for each simulation system are given in Tables A.2 and A.3.

WT-MTD simulations are performed to calculate the relative stability of the hydronium ions in the immediate vicinity of fructose and 5-HMF as compared to that in the bulk solvent, as shown in Figure 3.1 a.[133] WT-MTD involves the addition of bias potential as a sum of Gaussians deposited along the system trajectory as a function of chosen collective variables (CVs).[134] Also, the potentials deposited during the WT-MTD simulation can be traced back and the free energy surface (FES), as a function of collective variables, can be reconstructed.

The coordination number between fructose or 5-HMF molecules and hydronium ions is taken as the first collective variable (CV1) since this work attempts to investigate the relative stability of the hydronium ion in the bulk solvent as compared to that in the immediate vicinity of fructose/5-HMF. However, for the hydronium to enter to the first solvation shell of fructose/5-HMF, the number of water and DMSO molecules in the first solvation shell of fructose or HMF may change, and hence the coordination number between water and DMSO molecules and fructose/HMF is chosen as the second collective variable (CV2), as described in Figure 3.1. WT-MTD simulations were performed coupled with MD using NVT ensemble at 298 K using GROMACS-PLUMED/2016.3. To equilibrate our systems, NVT simulations are performed for 12ns before initiating the WT-MTD run. WT-MTD production runs are performed in NVT ensemble for 300 ns with time-step of 1 fs. Potentials are deposited after every 500 steps (500 picoseconds) with the initial height as 1.2 kJ/mol and the bias factor used for reducing the hill height is 25. Sigma values, i.e., the width of the potential deposited is taken as 0.05 for both collective variables. The advantage of using WT-MTD is that the Gaussian size changes adaptively. As the simulation time increases, the hill height decreases. Details of the convergence criteria for WT-MTD simulations are provided in the SI.
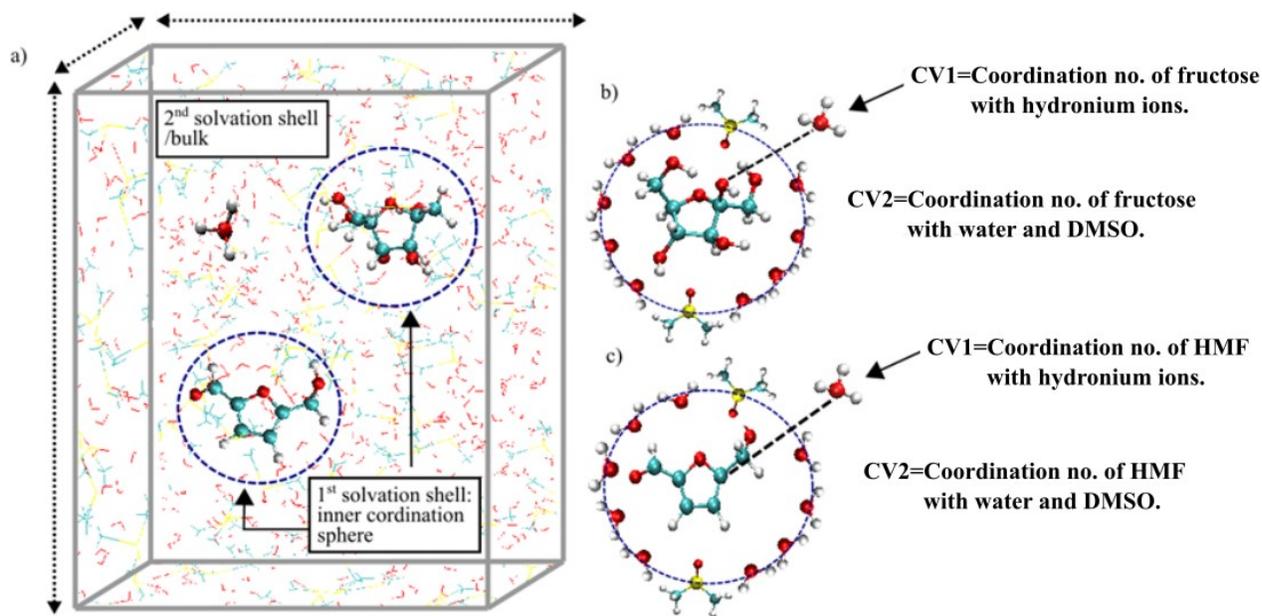
Figure 3.1 a) Schematic diagram of the simulation cell. b) Collective variables for fructose systems used in WT- metadynamics. c) Collective variables for 5-HMF systems used in WT-MTD. For atoms, yellow color is for sulphur, red is for oxygen, white is for hydrogen and cyan for carbons.

3.3 Results and discussion

Acid catalyzed conversion of fructose to 5-HMF and further rehydration of 5-HMF leading to the formation of levunilic acid, formic acid and humins is shown in Figure 3.2. Carbon atoms in Figure 3.2 are numbered to demonstrate their origin in the reactant. It is well known that the dehydration of fructose to HMF starts with the protonation of any of the hydroxyl groups bonded to carbon atoms 2, 3 and 4. However, the subsequent removal of the protonated OH of the C_2 carbon is much faster and thermodynamically more favorable than the removal of the other two during the first dehydration of fructose.[122] On the other hand, carbon atom C_5 is cleaved from HMF to form 2,5-dioxo-6-hydroxyhexanal (DHH). DHH is identified as a promoter for the formation of humins, whereas carbon atom C_1 is cleaved to form levunilic acid.[71] Since the composition of the solvent medium in which the acid catalyzed reactions shown in Figure 3.2 alters the conversion and the selectivity, we attempt to investigate the local organization of the hydronium ion in the immediate vicinity of fructose and 5-HMF by calculating radial distribution functions (RDF) between solute-hydronium atom pairs as the concentration of DMSO in the reacting system increases. Nevertheless, the local structure gets hidden in the computed radial averaging over the angular coordinates in RDFs. Therefore, three-dimensional arrangement of solvent and hydronium ions around the solutes are analyzed by constructing volumetric maps of the time-averaged densities. Later, we present and discuss the FESs reconstructed for fructose and 5-HMF at 50 % wt. of DMSO and the relative stability of the hydronium ion in the vicinity of fructose and 5-HMF for the entire range of DMSO concentrations studied.
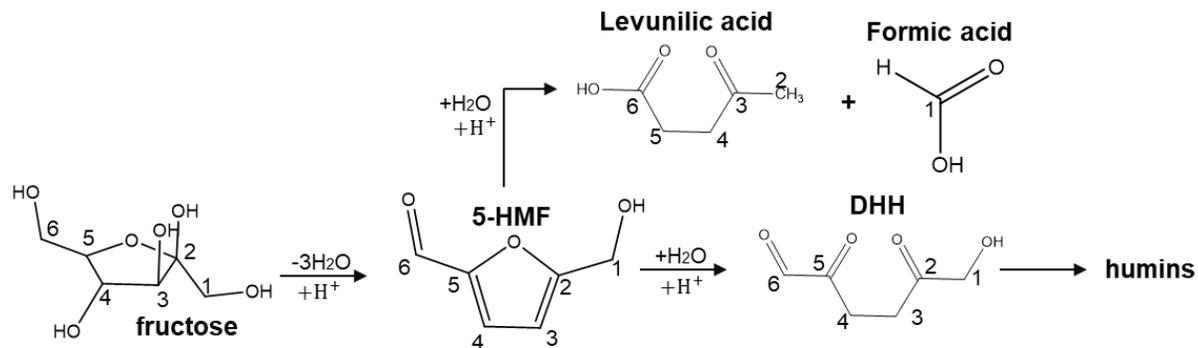
Figure 3.2 Dehydration of fructose to HMF and rehydration of HMF to DHH leading to the formation of humins. In all the species shown, the carbon atom numbering refers to the fructose molecule.

3.3.1 RDFs for Fructose and HMF systems

In a system of atoms, RDFs describe how the probability varies as a function of distance from a reference particle. In the context of acid-catalyzed reaction, here we attempt to understand how the hydronium ion interacts with the biomass species present in the reaction media, as a function of DMSO concentration. Figure 3.3 shows RDFs for fructose oxygen-hydronium oxygen atom pairs. The oxygen of fructose corresponds to the hydroxyl group attached to C_2. This oxygen has been chosen as the reference because the dehydration through this hydroxyl group is kinetically preferred over other hydroxyl groups.[135] For all DMSO concentrations, the first peak of the RDF appears at around 2.8 nm. However, the intensity of the peak increases remarkably as the concentration of DMSO increases. This shows that the interaction between hydronium ions with the polar group of fructose molecules becomes stronger at higher concentrations of DMSO, more so at the higher fructose concentration, as shown in Figure A.4. These results indicate that the presence of DMSO strengthens the interaction between the hydronium ion and the hydroxyl group of fructose. In the context of acid-catalyzed dehydration of fructose to 5-HMF, the enhancement of interaction between the hydronium ion and the hydroxyl group may facilitate the protonation of the hydroxyl group as the first step of the dehydration reaction. This finding correlates well with the experimental result reported by Chheda et al. [26], showing that the presence of DMSO leads to higher conversion and selectivity towards 5-HMF.
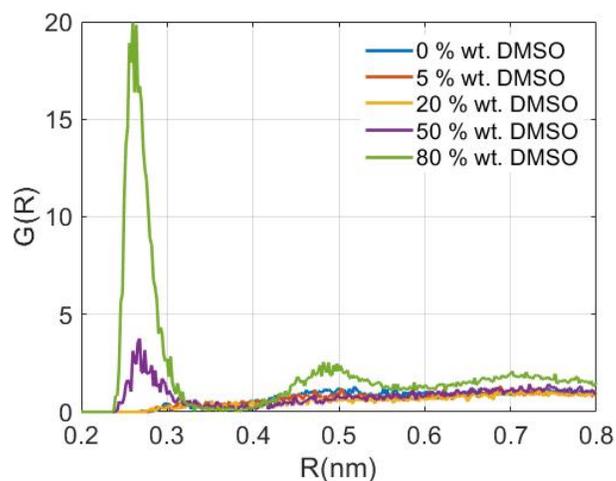
Figure 3.3 Radial pair distributions for the fructose oxygen (attached to C2 carbon)–hydronium ion oxygen atom pairs, at all DMSO concentration studied. Fructose concentration of 1% wt.

Additionally, we observe that the interaction of the fructose hydroxyl group with water is enhanced due to DMSO, and even more at a higher fructose concentration, as shown in Figures A.5 & A.6. Though the interaction between fructose hydroxyl groups and water is strengthened by adding DMSO, the number of hydrogen bonds between solvent molecules and fructose are slightly reduced due to fewer water molecules being available at higher DMSO concentrations, as shown in Table A.2.3. These results agree with previous computational simulations of fructose: DMSO: water system.[123] Water and hydronium interactions with fructose are expected to coincide since hydronium ions are typically solvated by water molecules. To have a better understanding of the solvation environment of hydronium ion in DMSO: water mixture, RDFs of the oxygen atoms of water and DMSO with hydronium oxygen atoms are calculated and shown in Figures 3.4 & 3.5 respectively. In both cases, as DMSO concentration increases, the interaction between the solvent molecules and the hydronium ions becomes greater. However, at each concentration of DMSO, based on peak intensities we conclude that hydronium ions show higher preference to interact with water than with DMSO.
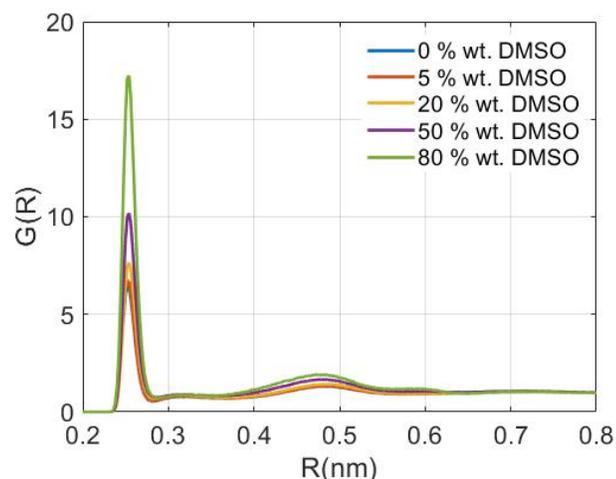
Figure 3.4 Radial pair distributions for the hydronium ion oxygen - water oxygen atom pairs, at all DMSO concentration studied. Fructose concentration of 1% wt.

RDFs for 5-HMF C_5-hydronium oxygen atom pairs are reported in Figure A.9. In contrast to fructose, no significant differences are observed between the interaction of C_5 and hydronium ions with increase in DMSO concentration, except for the RDF at a concentration of 80 wt. % DMSO, where the peak intensity is slightly lower compared to the others. These results do not vary at higher concentrations of HMF reported in Figure A.10. Additionally, the interaction between 5-HMF and DMSO was observed to be stronger than between 5-HMF and water, for all DMSO concentrations studied, as shown in Figs. A.11-A.14. These observations agree with the results reported by Mushrif et al.[126] for systems without hydronium ions. This could be due to the interaction of the hydrophobic methyl group or the sulfur atom in DMSO molecule with the atoms in the furanic ring of 5-HMF.
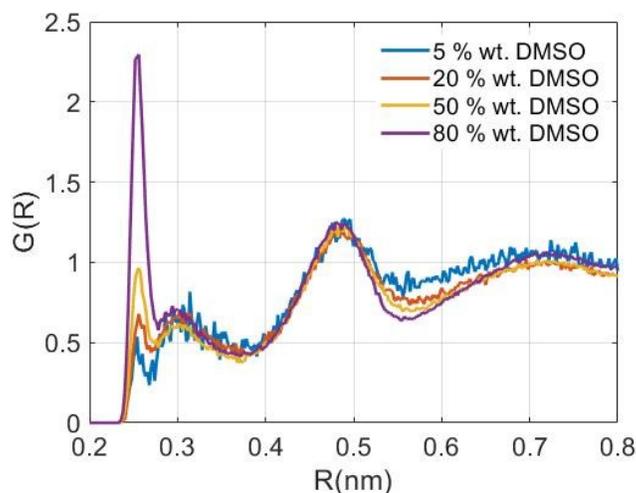
Figure 3.5 Radial pair distributions for the hydronium ion oxygen – DMSO oxygen atom pairs, at all DMSO concentration studied. Fructose concentration of 1% wt.

3.3.2 3-D arrangement of solvent molecules and hydronium ions around fructose and 5-HMF

To understand the local 3-dimensional structural arrangement of solvent around the solute, volumetric maps are created. In the generated maps, the solvents and the hydronium ions around the solutes can be visualized using isosurfaces representations. Figure 3.6 shows isosurfaces of hydronium ions (grey), water (red) and DMSO (blue). around the fructose molecule, for 5 to 50 % wt. DMSO concentrations. The isosurface of hydronium ions grows especially near the hydroxyl group of fructose bonded to C_2 carbon, with increasing DMSO concentration. In the context of the acid-catalyzed dehydration of fructose to 5-HMF, we suggest that DMSO possibly improves the selectivity due to the enhancement of the interaction of hydronium ion and the C_2 hydroxyl group, facilitating its protonation as the first step of fructose dehydration.[124] In addition, Figure 3.6 c reveals that DMSO molecules compete with water to be oriented towards the hydrogen atom of the hydroxyl group attached to C_2. However, as mentioned before, DMSO has preference over water to interact with the hydrogen atom due to the stronger hydrogen bonding with the sulfonyl oxygen than with water oxygen. These observations suggest that the presence of DMSO protects the hydroxyl hydrogen to be detached from the hydroxyl group, thus inhibiting the formation of glycosidic linkage with another sugar molecule.[136] The 3-dimensional arrangement of hydronium ions, water and DMSO molecules around an 5-HMF molecule is shown in Figure 3.7 for DMSO concentrations of 5, 20 and 50 % wt. Isosurfaces for other DMSO concentrations are provided in a movie format with the SI for better visualization. The preferential interaction between the carbon atoms of the furan ring of 5-HMF with DMSO over water molecules and hydronium ions can be clearly seen. Additionally, it is identified that hydronium ions and water molecules are in the immediate vicinity of the aldehyde group and the hydroxyl group of 5-HMF. It is also observed that with increasing concentration of DMSO, the presence of hydronium ions and water molecules is significantly reduced near the polar groups of 5-HMF. By the preferential solvation, DMSO protects the reactive sites of 5-HMF. Therefore, with an increase in DMSO concentration, the likelihood of 5-HMF protonation would go down. The reduction of humins or rehydration by-products by the addition of DMSO may be a consequence of the stabilization of 5-HMF by the preferential solvation of DMSO. This agrees with previous DFT calculations revealing that DMSO binds stronger than water to both polar groups of 5-HMF. [119]
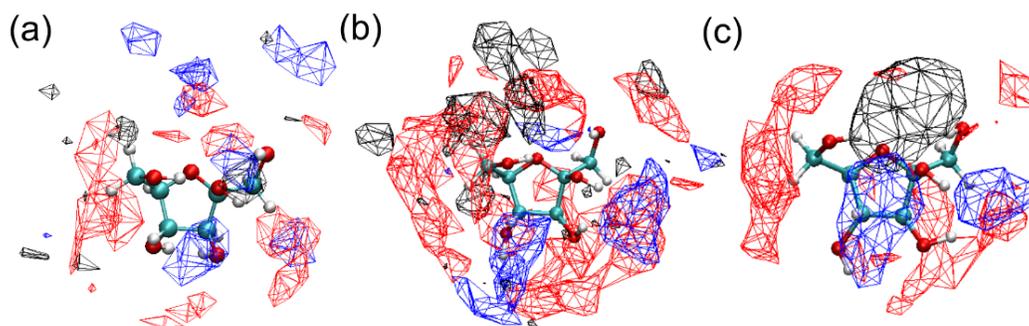
Figure 3.6 Volumetric map of the time-averaged distribution of water, DMSO and hydronium ions around a fructose molecule in the simulation, 1% wt. fructose is solvated in water-DMSO mixtures of different compositions, at 0.1 M concentration of hydronium ions. Black, red, and blue colored surfaces indicate hydronium ions, water, and DMSO oxygen atoms, respectively. (a) 5 % wt., (b) 20 % wt., (c) 50 % wt. of DMSO
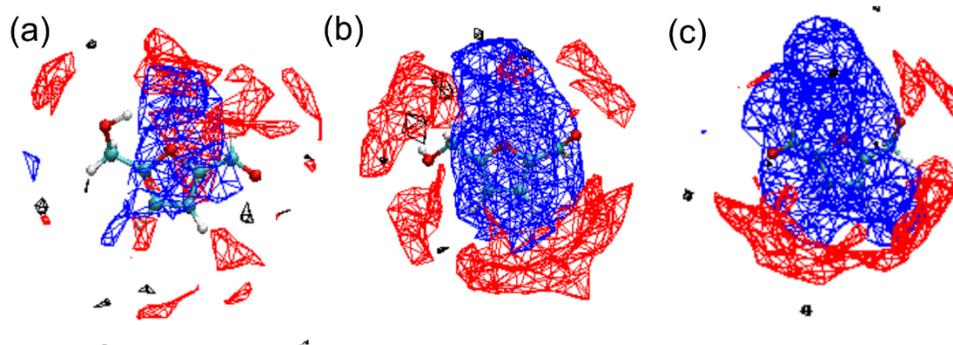
Figure 3.7 Volumetric map of the time-averaged distribution of water, DMSO and hydronium ions around a fructose molecule in the simulation, 1% wt. 5-HMF is solvated in water-DMSO mixtures of different compositions, at 0.1 M concentration of hydronium ions. Black, red, and blue colored surfaces indicate hydronium ions, water, and DMSO oxygen atoms, respectively. (a) 5 % wt., (b) 20 % wt., (c) 50 % wt. of DMSO

3.3.3 Well-Tempered-Metadynamics

FESs computed from WT-MTD simulations for 50 % wt. DMSO are shown in Figures 3.8 and 3.9. FESs for all other DMSO concentrations are provided in section A.2.5.

3.3.3.1 Fructose-hydronium ion interaction

Figure 3.8 shows the FES corresponding to the relative stability of hydronium ions near fructose as compared to that in the bulk solvent, as a function of CVs described in Figure 3.1 b and c. Like RDFs shown in Figure 3.3 for the systems with fructose, the oxygen of C_2 hydroxyl group is chosen for the CV definition. Multiple local minima in the free energy landscape can be observed. However, minima corresponding to CV_1 equal 0 and 0.7 are particularly important. The first minimum corresponds to the hydronium ion in the bulk solvent while the second corresponds to the hydronium ion present in the first solvation shell of fructose. The relative stability of the hydronium ion in the first solvation shell of fructose is quantified by the difference in free energies of the two minima. Lower the value of this ΔG (FES minimum corresponding to hydronium in the vicinity of Fructose – FES minimum corresponding to hydronium in bulk solvent), higher is the relative stability of the hydronium ion near fructose. In the context of acid catalyzed dehydration of fructose, the hydronium ion (or the proton) needs to shuttle from the bulk solvent to the first solvation shell of fructose, so that the protonation of the hydroxyl oxygen and the subsequent dehydration of fructose molecule, can happen. It is important to state that the barrier associated with the movement of proton to the immediate vicinity of fructose, due to the solvent reorganization, is insignificant.
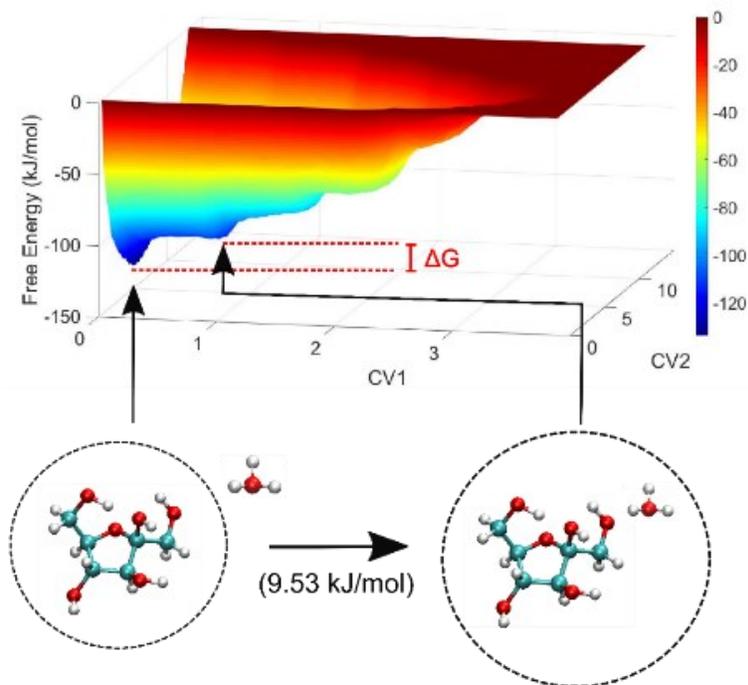
Figure 3.8 Free energy surfaces (FES) reconstructed from the WT-MTD simulations as function of coordination number of water & DMSO (CV2) and that of the hydronium ion (CV1) with 1 % wt. fructose. CV1 and CV2 are also defined in Fig. 3.1.

3.3.3.2 HMF-hydronium ion interaction

Figure 3.9 shows the FES corresponding to the relative stability of hydronium ions near 5-HMF as compared to that in bulk solvent. Carbon C_5 of 5-HMF is chosen for the CV definition. It is observed that the FES has a clear minimum at CV_1 equal to 0, corresponding to the hydronium ions in the bulk solvent. Same as the previous FES for fructose, this minimum corresponds to the most stable state of the system. At CV_1 to 0.8, another minimum (at higher free energy) is identified corresponding to one hydronium ion in the first solvation shell of 5-HMF. Higher the value of this ΔG (FES minimum corresponding to hydronium in the vicinity of 5-HMF – FES minimum corresponding to hydronium in bulk solvent), lower is the relative stability of hydronium ion in the first solvation shell of 5-HMF. As CV_1 further increases, no prominent local minima are found along the FES. Additionally, the free energy barrier associated with the solvent reorganization for the movement of hydronium ion form bulk to the first solvation shell of 5-HMF is ~ 23 kJ/mol.
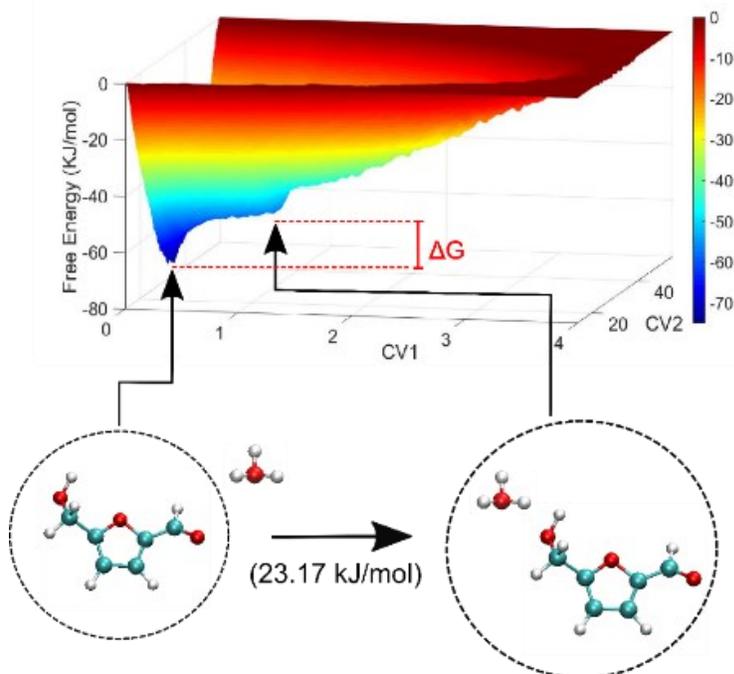
Figure 3.9 Free energy surfaces (FES) reconstructed from the WT-MTD simulations as function of coordination number of water & DMSO (CV2) and that of the hydronium ion (CV1) with 1 % wt. HMF. CV1 and CV2 are also defined in Fig. 3.1.

3.3.4 Relative stability of hydronium ions for different DMSO concentrations

As mentioned before, WT-MTD simulations were performed for all DMSO concentrations (FESs for all DMSO concentrations are reported in A.2.5). By comparing the ΔG (ΔG = FES minimum corresponding to hydronium in the vicinity of HMF/Fructose – FES minimum corresponding to hydronium in bulk solvent) for different DMSO concentrations, we analyze the effect of the addition of DMSO on the stability of the hydronium ion in the first solvation shell relative to that in the bulk solvent. Figure 3.10 shows the ΔG for fructose-hydronium ion at all DMSO concentrations. An increase in ΔG is observed when DMSO concentration increases from 0 to 5 % wt., indicating that the relative stability of the hydronium ion in the bulk solvent slightly increases compared to that in the vicinity of the fructose molecule. This may happen because of the instability of DMSO molecules in bulk water, generating a rich local domain of DMSO molecules near fructose hydroxyl groups, while water molecules in the bulk solvent stabilize hydronium ions. However, as DMSO concentration increases from 5 to 80 % wt., a clear descending trend is observed in ΔG values. This suggests that the relative stability of the hydronium ion in the first solvation shell of fructose increases. This may happen due to limited availability of water molecules in the bulk to stabilize the hydronium ion, forcing them to interact with the hydroxyl groups of the fructose molecule instead. These results agree well with the results obtained from the RDFs shown in Figure 3.3. It is observed that with increasing DMSO concentration, the interaction between hydronium ions and fructose oxygen increases significantly. Moreover, the volumetric maps in Figure 3.6 also show enhanced presence of hydronium ions near the C_2 hydroxyl group of fructose, as DMSO concentration increases. These findings are consistent

with the observation made by Mellmer et al.[137] Based on reaction kinetics and ab initio molecular dynamics studies of the acid catalyzed dehydration of fructose to 5-HMF, they concluded that increasing the proportion of DMSO would enhance the interaction between the protons accompanied by water molecules and the hydroxyl groups of fructose.

The change in ΔG for 5-HMF, as a function of DMSO concentration is also reported in Figure 3.10. An increasing trend in ΔG is clearly observed as DMSO concentration increases. As ΔG increases, the hydronium ion becomes less stable in the first solvation shell of 5-HMF. In other words, the addition of DMSO protects 5-HMF from protonation by inhibiting its interaction with the hydronium ion. It has been found that by-products, resulting from acid catalyzed condensation or rehydration reactions, are inhibited in the presence of DMSO.[117] The relative instability/lower likelihood of the hydronium ion in the immediate vicinity 5-HMF compared to that in the bulk can be correlated with the increase in the selectivity of 5-HMF. The opposite trends observed for the relative stabilization of the hydronium ion near fructose and 5-HMF are in consonance with the previous work done by Mellmer et al.[138] suggesting that the reactivity depends on the hydrophilicity of the biomass species. With five hydroxyl groups, fructose is lot more hydrophilic than 5-HMF, which has only one hydroxyl group and a carbonyl group.
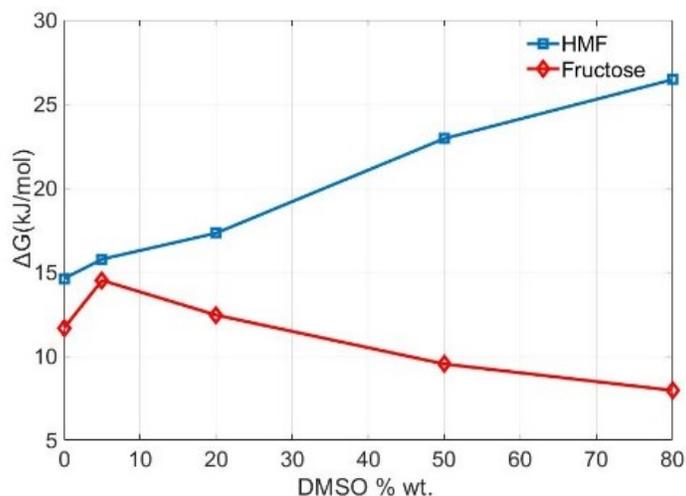
Figure 3.10 Comparison of the free energy difference between the two minima of the FES (ΔG = FES minimum corresponding to hydronium in the vicinity of Fructose/5-HMF – FES minimum corresponding to hydronium in bulk solvent) corresponding to the migration of the hydronium ion from the bulk of the solution to the fist solvation shell of Fructose/HMF at all DMSO concentration studied. Fructose/HMF concentration is 1 % wt.

3.4 Conclusion

The effect of the addition of DMSO on the interaction of hydronium ions with fructose and HMF in aqueous medium is investigated using molecular dynamics (MD) and Well-tempered metadynamics (WT-MTD) simulations. Analysis of MD trajectories using radial distribution functions revealed that with increase in DMSO concentration, hydronium ions' correlation with the C_2 fructose hydroxyl group is strengthened, more so at higher concentrations of fructose. Moreover, DMSO molecules are preferentially located in the

immediate vicinity of hydrogen atoms of the hydroxyl groups of fructose, allowing the direct interaction of hydronium ions with the oxygen atoms of the hydroxyl groups. This likely promotes the selective dehydration of fructose to 5-HMF by protonating the hydroxyl oxygen, to initiate the dehydration reaction and protects the fructose molecule from forming reversion and polymerization products. While for 5-HMF systems, DMSO clearly shows its preferential interaction with the molecule over hydronium ions and water molecules because of i) the hydrophobic methyl groups of DMSO interacting with the carbon atoms of 5-HMF and ii) the reduced number of hydrogen bonds between HMF and water than that between fructose and water. Moreover, the addition of DMSO reduces the presence of hydronium ions and water molecules in the immediate vicinity of the polar groups of HMF. This implies that DMSO protects the reactive sites of HMF, thereby suppressing the formation of humins or any rehydration by products via HMF. WT-MTD is performed to calculate the relative stability of the hydronium ions in the bulk solvent, as compared to that in the immediate vicinity of fructose and HMF. The free energy difference (ΔG = free energy in the immediate vicinity of fructose/HMF – free energy of the hydronium ion in bulk solvent) is used to quantify the relative stability. It was observed that with increase in DMSO concentration, the relative stability of the hydronium near fructose increases. In the case of HMF, WT-MTD elucidated that the relative stability of hydronium ions in the first solvation shell of C_5 carbon of 5-HMF decreases with increase in DMSO concentration. This proves that the addition of DMSO protects 5-HMF from interacting with the hydronium ions that act as catalysts for its degradation reactions. Conclusions from this work agree with the previously reported experimental work on the acid-catalyzed dehydration of fructose to 5-HMF, and the degradation of HMF to condensation and rehydration by-products. Analogous to the effect that solvent environment has on the interaction of the reactant molecule with a heterogeneous catalyst; we demonstrate here the effect of solvent environment, i.e., the concentration of an aprotic solvent like DMSO, on the interaction of hydronium ions (the catalyst) with the reactant fructose and 5-HMF molecules in a homogeneous environment.

Chapter 4 Energetics of acid catalyzed biomass reactions: How and why does the solvent model matter?

4.1 Introduction

Several investigations have been done to provide possible explanations of specialty chemicals and electrochemical and pharmaceutical transformations that are highly dependent on solvents.[139] In addition, solvents are key to the selective transformation of biomass-derived molecules to produce platform chemicals. [51] These processes for biomass deconstruction or upgrading reactions, such as hydrolysis, dehydrogenation, aldol addition, and dehydration, to name a few, are very commonly performed in a homogeneous acid environment. These acid-catalyzed biomass reactions are very susceptible to changes in solvent composition, leading to enhanced or worsened conversions and selectivity.[140] For lignocellulose saccharification, the addition of aprotic solvent in the aqueous media enhanced glucose recovery and minimized glucose derived oligomerization reactions under acid free conditions.[141] For example, during the acid catalyzed hydrolysis of cellulose, hydrolysis reaction rate as well as glucose yield are found to increase in presence of GVL, with a high glucose recovery from 70 to 91 % in 10 % GVL-water mixture, suppressing humins formation.[142] Also, during the dehydration of biomass moieties such as glucose, fructose, or xylose to furanic compounds using homogeneous acid catalyst, the addition of a polar aprotic solvent (such as DMSO, acetone or ɣ-valerolactone (GVL)) in the aqueous phase, manifests higher conversion and the selectivity increases from 30 up to 90 %.[81] In addition, during the acid-catalyzed hydrolysis of furfuryl alcohol to levulinic acid, conversion, and selectivity are enhanced in the presence of THF as a cosolvent. A THF-water mixture 19:1 using sulfuric acid leads to an increase in yield of ~ 50 % compared to pure water.

Investigations have been done in the past to provide possible explanations of the role of solvents in these biomass reactions. Experimental studies of reaction kinetics of the acid-catalyzed conversion of xylose into furfural show that the apparent activation energy decreases from 145 to 114 kJmol-1 when water was substituted by ɣ-valerolactone (GVL), whereas the activation barrier for furfural degradation is higher in GVL (from 85 to 114 kJmol-1). These changes in the kinetics of the formation of furfural using GVL favors the selectivity towards the desired product. Additionally, for strong Bronsted acid catalysts such as sulfuric acid, using GVL instead of water as solvent increases 18-fold its reactivity during the dehydration of 1,2 propanediol to propanal and 30-fold for the hydrolysis of cellobiose to glucose. [94] The presence of aprotic solvent in the reacting system alters reaction kinetics and changes the chemical equilibrium. During the acid-catalyzed hydrolysis of cellulose to 5-HMF in an acid-aqueous environment, 5-HMF is produced from the triple dehydration of glucose. However, Huber and coworkers demonstrated that in the presence of THF, the glucose-levoglucosan (LGA) chemical equilibrium changes towards LGA, which is further dehydrated to form 5-HMF. [103]

Molecular modeling can be an effective tool for studying condensed-phase catalytic biomass reactions. These computational tools can help to obtain a fundamental molecular understanding of the effect of solvent

on biomass reactions to biochemicals. This knowledge would allow better solvent selection and to identify which characteristics of the solvent molecules lead to enhanced reaction yields. In addition, customizing catalytic technologies and conditions can significantly improve these reactions.[118] Molecular modeling for investigating solvent effects decouples two levels: physical interactions (solvation effects) and chemical interactions directly participating in the reaction by altering kinetics. Physical interaction encompasses preferential solvation of the reacting species and can provide a shielding effect to avoid unwanted reactions, while chemical effects involve alterations in activations barrier, chemical equilibrium, or even changing reaction mechanisms.

Force-field-based molecular dynamics (MD) are often used to investigate the physical interactions of solvents with biomass-derived moieties at a moderate computational cost. For example, using MD, it has been shown that DMSO preferentially solvates 5-HMF in water-DMSO mixtures, explaining how this provides a shielding effect to the 5-HMF, thus inhibiting its degradation to humins.[99] The relative stability of the catalyst, hydronium ion, near fructose and 5-HMF was also computed as a function of solvent composition using MD-coupled with biasing methods like metadynamics.[143] It was observed that DMSO improved the stabilization of the hydronium ions in the first solvation shell of fructose compared to that in the bulk solvent, while hydronium ions become less stable near 5-HMF as the concentration of DMSO increases. In other words, adding DMSO protects 5-HMF from protonation by inhibiting its interaction with the hydronium ion. These observations agree with experiments where humins formation, along with acid-catalyzed dehydration of fructose, is inhibited in the presence of DMSO. Moreover, Dumesic and coworkers [92] extended molecular dynamics simulation outcomes to predict kinetics for acid-catalyzed reactions for biomass-derived oxygenates. Experimentally, the catalytic turnover rates increase in water-polar aprotic solvent mixtures such as ɣ-valerolactone, 1,4-dioxane, and tetrahydrofuran. Furthermore, they determined that adding organic solvents in the reacting system leads to water-enriched local solvent domains near hydrophilic reactants and increases hydrogen bonding intensity between reactants and local water molecules. Further, these effects were correlated into descriptors to develop a multi-descriptor model that predicts the rate constants as the function of solvent composition.

As mentioned before, the solvent composition can also alter reaction kinetics and thermodynamics during the acid-catalyzed reactions of biomass moieties. Therefore, to analyze the chemical effects of the solvent environment, implicit and explicit solvent environment is incorporated in quantum mechanical modeling methods. The implicit solvation approach in the first-principles density functional theory (DFT) calculations is a mean-field approximation where the solvent is represented as a homogeneous dielectric continuum, which may not always be adequate. [144] For example, propanal is the primary product of the acid dehydration of 1,2 propanediol in pure water; however, the selectivity shifts towards acetone in water: DMSO mixtures. Both reaction pathways are initiated with the protonation of two different hydroxyl groups. Van Lehn and coworkers [145] determined the stability of the reactant protonated in pure water and pure DMSO using an implicit solvation model. They found that models cannot capture the influence of the local solvent

environment of the reactant. A few explicit solvent molecules were placed near the hydroxyl groups considering multiple solvent configurations to overcome this. DMSO molecules highly stabilize the protonated hydroxyl group leading to acetone formation, which agrees with experiments. However, this leads to a considerable increase in the computational cost and a limited number of solvent configurations to explore. For explicit solvent modeling for chemical effects, equilibrium solvation and solvent dynamics and polarizability are the two main categories. Equilibrium solvation models assume that solvent molecules stabilizing the reactant, transition state (TS), and products are in an equilibrated configuration and that the relative stabilities of the reacting species are governed by the interactions with those explicit equilibrated solvent molecules in the neighborhood and that the dynamics/timescale of the changes in solvent configuration along the reaction coordinate are much faster than the dynamics/timescale of the reaction.[146] On the other hand, solvent dynamics could be of great importance when timescales/dynamics for solvent reorientation during the reaction are slower and when that may result in non-equilibrium solvation of the transition state or products.[147] Janik and coworkers[148] quantified the energetics for the reaction steps during the first dehydration of fructose to 5-HMF in GVL-water mixtures combining DFT with force-field-based MD, taking the equilibrium solvation approach. They calculated the activation-free energies in the gas phase (using DFT) and incorporated each species' solvation-free energies (calculated using molecular mechanics) along the reaction coordinates. The results obtained using this approach resemble the reaction-free energy profiles for the same dehydration reaction reported by Mellmer and coworkers [149] , who treated the explicit solvent molecules quantum mechanically and with finite temperature effects, using Ab Initio molecular dynamics (AIMD). The main difference between AIMD and the hybrid DFT method combined with molecular mechanics is that AIMD considers solvent dynamics, and the solvent environment is polarizable in response to charge redistribution in the reactant. However, Caratzoulas et al. [150] investigated the subsequent steps after the first dehydration of fructose, leading to 5-HMF. It was found that the intramolecular hydride transfer just after the first dehydration is the rate-limiting step of the entire mechanism; the reorganization of the polar solvent environment near the reactant affects the activation energy associated with it. This observation was confirmed when the same group demonstrated that experimental kinetic data for this reaction could only be predicted when the solvent reorganization energy (using Marcus' theory rate constants) was considered for the rate-limiting hydride transfer step. [151] Finally, Mushrif simulated the hydride transfer reaction in water and methanol using AIMD and metadynamics.[152] Hydride transfer had a higher free energy barrier of 50 kJmol-1 in methanol than in water. Because methanol is more polarizable than water, it had to undergo a higher level of molecular reorientation than that of water and that too with slower dynamics, during hydride transfer, resulting in non-equilibrium solvation of the transition state.

These examples demonstrate how some reaction steps within the same reaction mechanism can be more susceptible to the local density of solvent molecules to stabilize the reactive species (relative free energies of reactant, TS, and products). Whereas some other steps involving charge transfer leading to an intense charge redistribution in the reactant are more affected by solvent polarization and dynamics. Furthermore, a wide range of computational schemes, such as first-principles-DFT[153], AIMD with enhanced sampling

methods[154], quantum/molecular mechanics (QM/MM)[155], and hybrid DFT-MD[156], are implemented for studying the chemical phenomena of solvents in reactions. These methods broadly fall into three categories, implicit solvation, explicit equilibrated solvation without solvent dynamics and polarizability, and explicit solvation with solvent dynamics and polarizable solvent. A variety of literature analyzes and compares the scope and parameterization of these methods[157]; however, to the best of our knowledge, no single study systematically compares these three different solvent modeling approaches for the same reaction system.

Hence, in this work, we attempt to analyze three different solvent categories (implicit, explicit without solvent dynamics and polarizability, and solvent dynamics with polarizable quantum mechanically treated solvent) for a key reaction step within the acid-catalyzed biomass transformation. The sample reaction for this study is the protonation of 5-HMF, as shown in Figure 4.1. We have taken 5-HMF as the model compound since it is an important platform chemical and an intermediate leading to humins formation and structurally like many furanic, cellulose-derived moieties. Besides, the protonation reaction is the first step of acid-catalyzed biomass reactions, such as dehydration, rehydration, hydrolysis, and keto-enol tautomerization.[145] Additionally, in many cases, protonation step governs the kinetics. The DFT-based calculations perform the protonation reaction in an implicit solvent, where a forcefield is generated as a homogeneous dielectric medium (the dielectric constant value referred to the solvent to model).[158] For explicit equilibrium solvation without dynamics, a hybrid method of DFT-gas phase calculations with MD allows us to estimate the reaction-free energies in the gas phase and incorporate the solvation free for the reactant, TS, and product. Finally, solvent dynamics and polarization are performed using AIMD with metadynamics to accelerate the dynamics, accessing a larger timescale for the protonation reaction. Simulation details are reported in section 4.2. Then, in section 4.3, we discuss the results, elucidating the difference in computed reaction energetics due to the choice of solvent modeling scheme. We also provide a molecular-level explanation of the observed differences in reaction energetics. Finally, we conclude our findings in section 4.4.
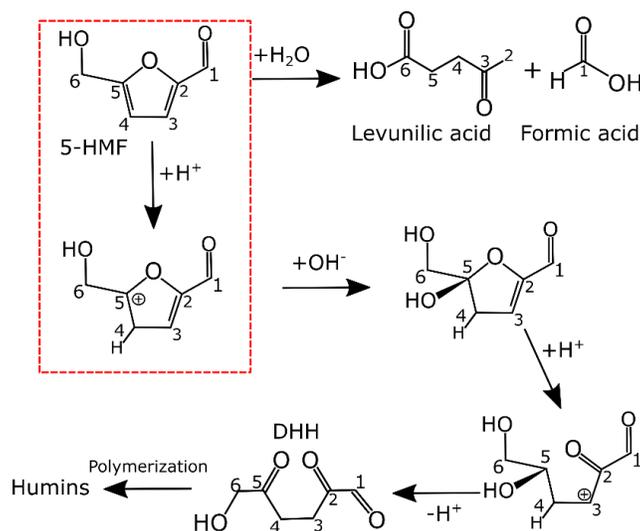
Figure 4.1 5-HMF possible rehydration reactions leading to levunilic and formic acids as well as DHH promoting humins formation. The protonation reaction highlighted in red is the reaction investigated in this study.

4.2 Methodology
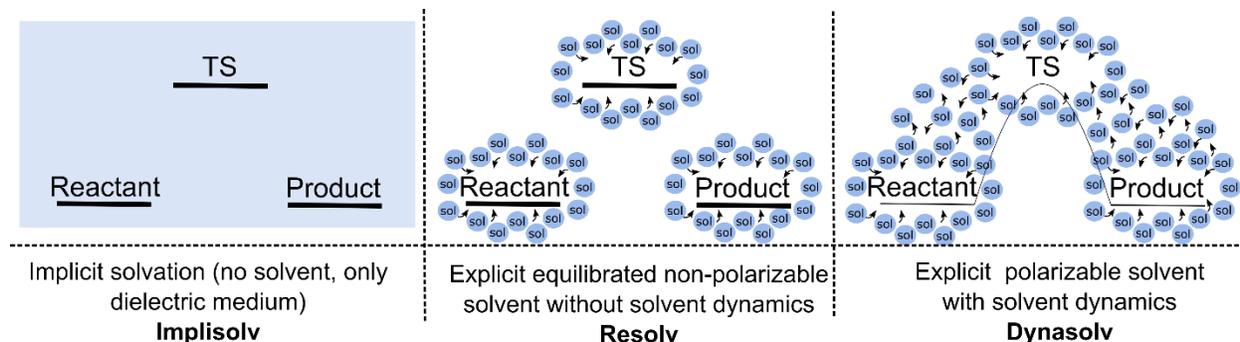
Figure 4.2 Schematic representation of the different approaches to model solvent systems.

4.2.1 First principles DFT calculations with implicit solvation model – Implisolv method.

DFT calculations are performed using Gaussian 09 code.[110] The hybrid functional M06-2X combined with 6-311++G (2d, p) basis set is used.[159] Gas phase and implicit solvent calculations are performed to compare the effect of implicit solvation on gas phase reaction energetics. Additionally, gas phase values are combined with the free energy values during the ReSolv method implementation, explained in section 2.2. For implicit solvent calculations, the polarizable continuum model (PCM) using the integral equation formalism variant (IEFPCM) is used during all geometry optimization calculations and TS search. This method consists of the replacement of the solvent with a homogeneous continuum polarizable medium to calculate the electrostatic interaction between the solvent medium and the solute. This approach reduces the solvation effects to a subject of a classical electrostatic involving solute charge distribution and permittivity. [160] Implicit solvation is implemented using the dielectric constant (ε) of DMSO and water. The transition state (TS) search is performed using the Berny algorithm. Intrinsic reaction coordinate (IRC) is performed to confirm that the TS is linked with the reactant and product on the reaction coordinates. The reaction-free energies in the gas phase and implicit solvent are reported at 433 K, which is in the temperature range in which biomass reactions are performed in the condensed phase that involves the protonation step.[71] The Zundel cation $H_5O_2{}^+$ is chosen as the acid catalyst.[113] All energies are reported in kJ/mol.

4.2.2 Explicit equilibrated solvation, non-polarizable without solvent dynamics - Resolv method.

The computational approach, previously named as ReSolv, [161] is used to calculate the free energy barrier and reaction free energy correction in the condensed phase. It involves the dissociation of the reaction into multiple associated steps in a thermodynamic cycle, as shown in Figure 4.3. [162] The reaction in the condensed phase is represented as a combination of associated reactions. This thermodynamic cycle

estimates the energetics in the condensed phase by calculating the activation and reaction-free energies in the gas phase and the relative interaction energies of the reacting species with the solvent molecules (reactant, TS, and product) in the condensed phase. Gas phase barrier calculations are explained in section 4.1.1. On the other hand, the interaction free energies of the reacting species with the solvent molecules (solvation-free energies) are calculated with the thermodynamic integration considering intermediate states connected by a coupling parameter λ. Solvation free energies are calculated using molecular mechanics (non-polarizable force-field).

The thermodynamic integration (TI) method is implemented using GROMACS 2020.4. [163] The structure for the reactant, TS, and product is incorporated in the explicit solvent system. OPLS-AA force-field parameters are used in combination with Langevin dynamics. [129] All simulations are first energy minimized using the steepest descent algorithm for 0.1 ns and equilibrated for 0.4 ns. Isobaric-isothermal runs were performed for 0.4 ns to bring the system to the appropriate density. Then, TI calculations are performed during the NVT ensemble for 2 ns for all the solvent compositions. [164] The intermediate states are optimized, and the number of 30 λ-windows is found to be adequate. After the entire calculations are done, the potential obtained from the MD simulations corresponding to each λ-windows is integrated to obtain the solvation-free energy (ΔG^{sol}). Finally, ΔG^{sol} is incorporated into the gas phase free energy values for reactant, TS, and product to calculate the energy profiles using explicit solvent molecules condensed phase. MD trajectories are analyzed using the Visual Molecular Dynamics (VMD), version 1.9.3. [132]
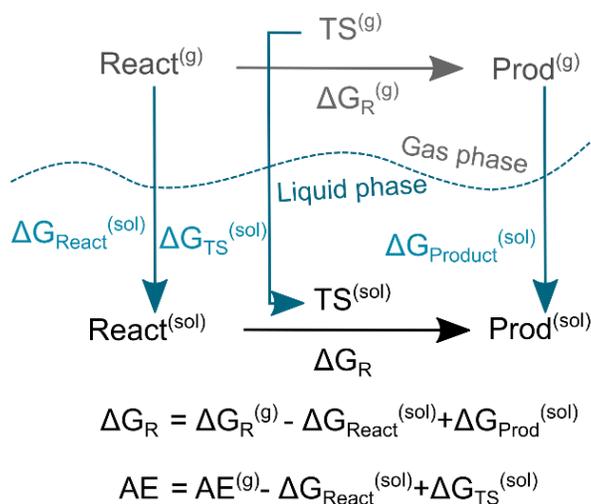
Figure 4.3 Thermodynamic cycle using Re-Solv method. The activation free energy change and reaction free energy in the gas ($\Delta G_R^{(g)}$, top) and solution (ΔG_R, bottom) phases.[148] Reprinted (adapted) with permission from ACS Catal. 2022, 12, 13193−13206. © 2022 American Chemical Society.

4.2.3 Ab initio molecular dynamics (CPMD scheme)-metadynamics implementation – Dynasolv method

A prior molecular mechanics-based MD in an isothermal-isobaric ensemble is performed to obtain correct density, followed by two ns of NVT ensemble. The same solvent compositions as for the TI method were analyzed using AIMD.   Ab initio MD simulations are performed using the CPMD package, version 4.3. [165] The first principles calculations were carried out using the plane-wave-pseudopotential implementation of the Kohn-Sham density functional theory.[166] Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional using generalized gradient approximation (GGA).[167] Also, the gamma point was implemented for integration over the Brillouin zone in the reciprocal space. An energy cut-off of 70 Ryd was optimal for energy convergence. The fictitious electron mass parameter in CPMD was set to 400 au. Geometry optimization was performed on the system before the MD run. Energies, including the fictitious electronic kinetic energy, were monitored to ensure that the system did not deviate from the Born– Oppenheimer surface during the MD simulation. A Nose–Hoover chain thermostat was used for controlling ionic and electronic temperatures. The frequency for the ionic thermostat was set to 3000 cm-1 as this is the approximate frequency for the C-H and O-H bond vibrations. The electronic frequency was set to 10000 cm-1 to avoid coupling between the ionic and electronic dynamics for all systems. The MD time step used in the simulation was 0.0967 fs.

Even though CPMD formulation reduced the computational cost of AIMD simulations, the accessible timescale is limited to a few picoseconds. Therefore, metadynamics has been implemented coupled with CPMD to accelerate the dynamics and to reconstruct the free energy surface in the coordinate space of interest. The metadynamics technique is based on the principle of filling up the energy surface with arbitrary potentials at small time intervals to "flatten" the energy surface (FES), as described by Laio and Gervasio. [168] In this way, the potentials added during the simulations can be tracked and the FES can be constructed as a function of the collective variables (CVs) specified.

The selected CVs to simulate the protonation of 5-HMF are defined in Figure 4.4.[134] Before protonation, the proton H_α is bonded to water-oxygen O_{water} and hence, CV1 ≈ 1 and CV2 ≈ 0. However, after protonation either H_α or any H_{water} is bonded to 5-HMF C_4 , therefore CV2 ≈ 1 and CV2 could be 0 or 1 depending on if H_{water} or H_α is bonded to C_4, respectively. The height hill of the potential added was kept fixed at 2.6 kJ/mol. Analogous to the original Car–Parrinello scheme, the dynamics of the collective variables were separated from the ionic and fictitious electronic motion by choosing an appropriate value for the fictitious mass of the collective variables. The temperature of the collective variables is set to 433 K and is controlled in a window of 150 K using velocity rescaling.

Before Protonation → After Protonation

Coordination no. of C_4 with H_{water} & $H\alpha$ (CV1)

Coordination no. of $H\alpha$ with O_{water} (CV2)
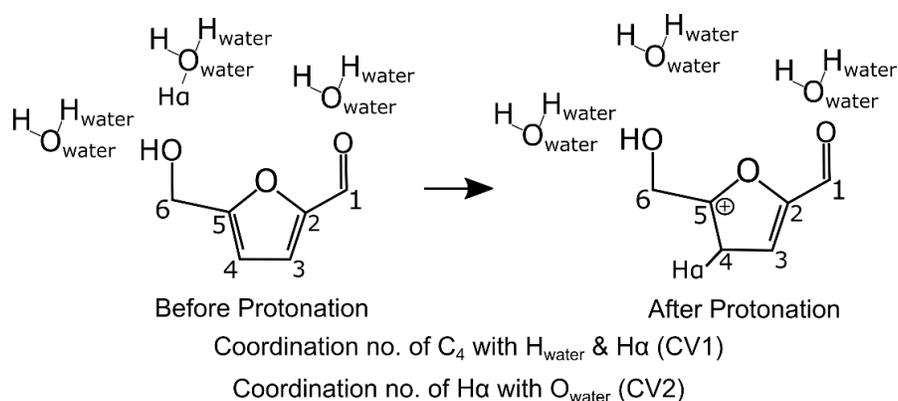
Figure 4.4 Collective variables CVs for the 5-HMF protonation at C_4 using CPMD-metadynamics.

4.3 Results and discussions

4.3.1 Gas phase and Implisolv method for implicit solvent

The free energy profiles of the quantum mechanical calculations using DFT in gas phase as well as in the implicit solvent environment are shown in Figure 4.5 a. The activation energy for all the cases is ~ 90 kJ/mol, with an insignificant difference between the gas phase and implicit solvents (water and DMSO). However, there is a reaction free energy (ΔG_R) difference of ~ 16 kJ/mol between the gas phase and the solvents (gas phase being more favorable), without any disparity between water and DMSO. The reaction is endergonic in all cases. The energy barrier for the protonation of 5-HMF C_4 is close to a similar reaction, such as the protonation of a ketone carbon from levunilic acid (LA) to form pentanoic acid of ~ 84 kJ/mol using an implicit water.[169] Also, various biomass reactions reported in the literature, such as acid-catalyzed hydrolysis of pyranose, acid catalyzed conversion of LA to valeric acid or pentane, and the cleavage of cellulose glycosidic bonds, have been studied using DFT-implicit solvent environments. For the reactions mentioned earlier, some of their steps reveal (i) differences in the energy profiles of the reaction in the gas phase and implicit solvent and (ii) the changes in energetics between the implicitly modeled solvents are negligible despite having different dielectric constants. [169] Figures 4.5 b and c show the most stable product conformers for the gas phase and implicit solvent environment respectively. Differences between DMSO and water are negligible, so only one conformation is chosen to represent the product in implicit solvation. It can be seen how oxygen O_ω forms hydrogen bonds (H-bond) with the hydroxyl group bonded to carbon C_6 and the hydrogen bonded to oxygen O_β, while in implicit solvation, O_ω interacts with carbonyl group of carbon C_1. Although the stabilization of the electrophilic C_5 seems to be similar in both cases, the difference in water molecules' orientation between gas phase and implicit solvation may be the cause of changing the relative stabilization of the product conformer. The almost unchanged energy profile using different implicit solvents has been observed for other acid catalyzed biomass reactions in the literature too, even though there is significant difference in their polarities.[152] However, these computational predictions fail to explain the experimental observation that the presence of DMSO reduces 5-HMF degradation into humins. [107] This suggest that the role of solvent in increasing the selectivity towards 5-HMF by the addition of aprotic solvent

cannot be explained using implicit solvation model in DFT calculations.
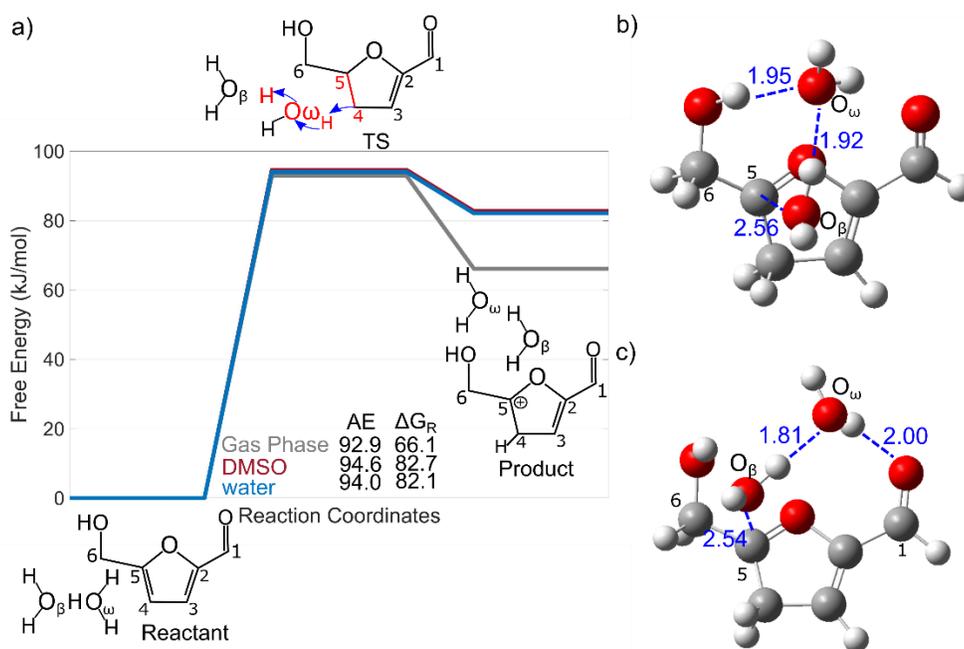
Figure 4.5 (a) Free energy profile for the protonation step of 5-HMF. Free energy values in gas phase, water and DMSO are shown in gray, blue and red, respectively. Activation energies (AE) and Reaction Free Energies (∆GR) are reported in kJ/mol. (b) product conformation in implicit solvent. The blue dotted lines indicate the intermolecular interactions between the water molecules and the product polar groups

4.3.2 Explicit, equilibrated, non-polarizable solvent without solvent dynamics using Resolv method

In contrast to the implicit solvation, the solvent molecules are explicitly modeled to interact with the reacting species (reactant, TS, and product) along the reaction coordinates in this method. The Resolv method allows explicit interactions of solvent molecules with the reacting species to investigate its effect on the relative stabilization of reacting species in the equilibrated solvent environment. Regarding solvent polarizability, there are methods to implement polarizability by parameterizing the force fields (fluctuating charges [170], drude oscillator better known as "charges on springs" [171] or inducible dipoles [172]) so that the charges of the solvent molecules can be distorted. Here, we study the interaction of the solvent molecules with the reactant, TS, and product as independent systems in equilibrium, where there is no charge redistribution due to the chemical reaction. Hence, in this work, a non-polarizable forcefield is used.

We computed the free energy profile along the reaction coordinates for the protonation of 5-HMF at C_4 in pure water and in water-DMSO mixtures of different compositions, as shown in Figures 4.6 a and b. It is observed that there is almost no effect on the activation free energies as the solvent composition changes from 0 to 75 wt.% DMSO. The activation barriers are like those computed using implicit solvent Implisolv method. However, in contrast to the results using an implicit solvent environment, there are changes in the

ΔG_R as a function of solvent composition. In the presence of explicit solvent molecules, the reaction is highly exergonic. The solvent environment with 75 % wt. DMSO increases ΔG_R by ~ 67 kJ/mol compared to 0, 25 and 50 wt.% DMSO.
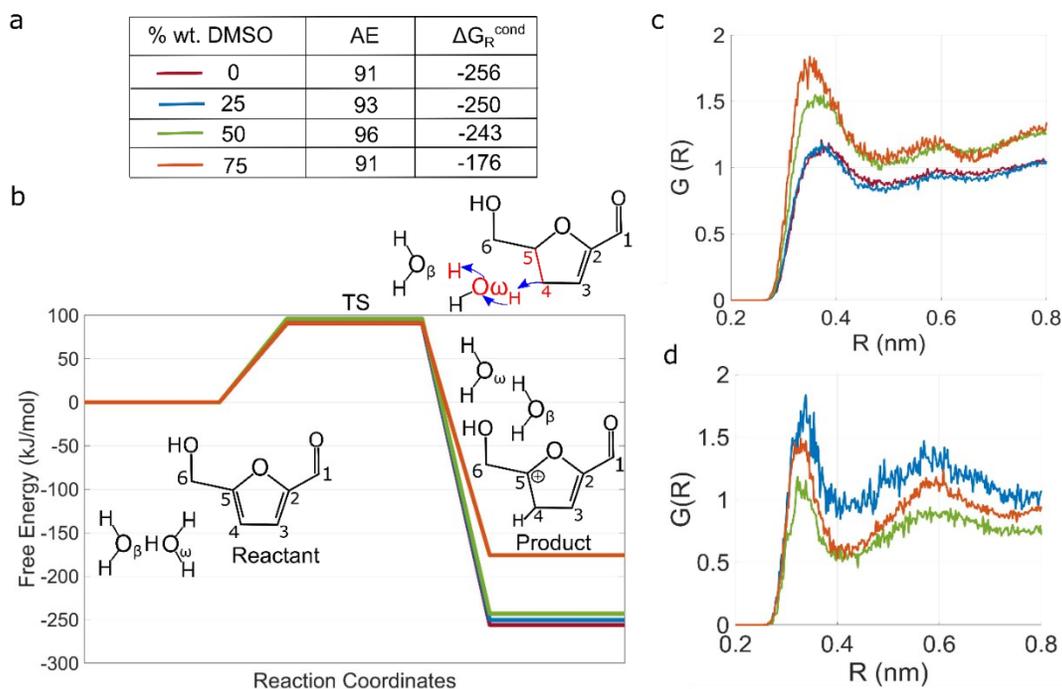
Figure 4.6 (a, b) Free energy profile for the protonation step of 5-HMF. Free energy values for corrections based on solvation in condensed phase at different concentrations of DMSO. Activation energies (AE) and Reaction Free Energies (ΔGR) are reported in kJ/mol. Radial pair distribution (c) C_5-water oxygen (d) C_5-DMSO oxygen atom pairs, at all DMSO concentrations studied.

This change in ΔG_R at higher DMSO proportions could be attributed to change in the relative stabilization of reacting species. After the protonation, there is a charge redistribution in the product conformer, in which the electrophilic C_5^+ would prefer to be stabilized by neighboring water molecules. To get insights about the changes in the relative stability of the product as solvent composition changes, radial distribution function (RDF) between 5-HMF C_5^+ carbon and oxygen of water (O_{water}) and 5-HMF C_5^+ carbon - oxygen of DMSO (O_{DMSO}) were calculated as shown in Figures 4.6 c and d, respectively. The first peak for C_5^+-O_{water} and C_5^+-O_{DMSO} atoms pairs appear at ~ 3 nm for all DMSO concentrations. However, the intensity of the peak increases monotonically for C_5^+- O_{water} atom pair as the DMSO proportion increases. However, for C_5^+-O_{DMSO} atom pair, the peak intensity changes non-monotonically. The peak intensity decreases when DMSO proportions increases from 25 to 50 wt.% followed by an increase at 75 % wt. DMSO. This non-monotonic trend is identified for the product, not in the reactant and the TS. The general trend for reactant's carbon atoms and O_{DMSO} atom pairs is that the intensity of the peak decreases by adding DMSO. In contrast, reactant's carbon atoms with O_{water} atom pairs show the opposite trend, as DMSO increases, the intensity of the peak increases too. The same trend is observed in the case of O_{DMSO} and O_{water} with the carbon

atoms' interactions in the TS. The RDFs for reactant and TS are reported in section 6.3.3. Since DMSO interaction with C_5^+ increases at 75 % wt. DMSO molecules compete with water molecules to be near C_5^+ in the product. Water molecules stabilize C_5^+ better than DMSO. Hence, stabilization of C_5^+ with water molecules is a crucial factor governing the relative stability of the product. However, the electrophilic C_5^+ becomes relatively unstable because there are not enough water molecules around to stabilize it due to the presence of DMSO.
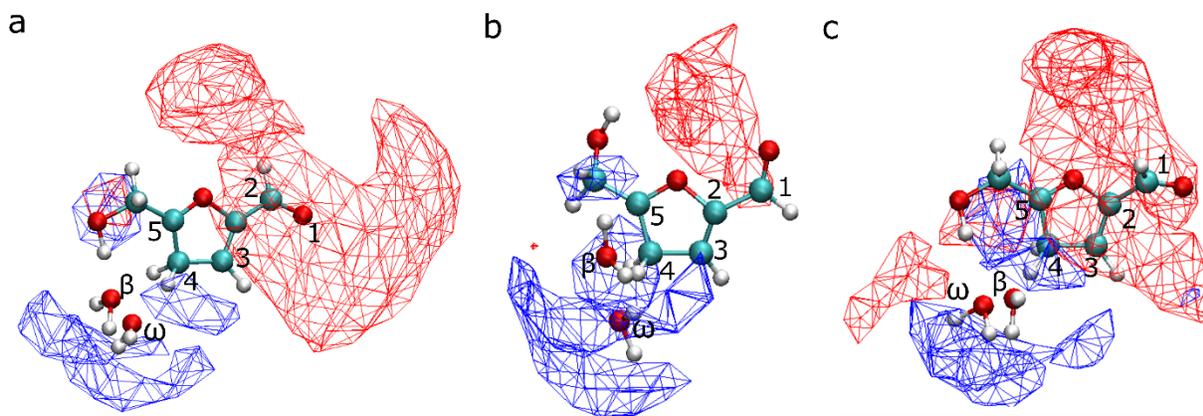
Figure 4.7 (a) Volumetric map of the time-averaged distribution of water, DMSO around the product conformer. Red, black, and blue colored surfaces indicate DMSO, hydronium ions and water atoms, respectively. (a) 25%wt., (b) 50%wt., (c) 75%wt. of DMSO.

To understand the 3-D local arrangement of the solvent around the product conformer, volumetric maps are created for all DMSO proportions.[98] These maps can be visualized by isosurface representations of the probability density of O_{water} and O_{DMSO} around the product. Volumetric maps at 25, 50 and 75 % wt. DMSO are shown in Figure 4.7. It can be seen how the isosurface of water molecules grows towards C_4, C_5 and C_6 of 5-HMF with increasing DMSO proportion from 25 to 50 % wt. However, Figure 4.7 c reveals that at 75 % wt. DMSO, DMSO competes with water molecules to be oriented towards C_4, C_5 and C_6 of 5-HMF. In the context of acid catalyzed protonation of 5-HMF, we suggest that the presence of DMSO near the electrophilic C_5^+, hinder the water molecules in stabilizing C_5^+, thus the protonation of 5-HMF becomes thermodynamically unfavorable in the presence of DMSO at 75 % wt. compared to lower DMSO proportions. These results align well with the experimental observations about the enhanced selectivity during the acid-catalyzed dehydration of sugars to 5-HMF and the inhibition of the side reaction of HMF.[8182173174175] The destabilization of the product by increasing the DMSO concentration is not seen using the implicit solvent model since the solvent competition near C_5^+ can only be observed with an explicit solvent system, where the interaction between the solvent molecules with the reactive species is captured.

4.3.3 Energetics of the protonation of 5-HMF using Dynasolv method.

Unlike Implisolv and Resolv, the Dynasolv method considers solvent dynamics and polarizability (since solvent is treated quantum mechanically). As a result, the Dynasolv method allows the solvent reorientation

and polarization to be modeled along the reaction trajectory/dynamics, in contrast to the previous approach, where the equilibrated explicit solvent molecules interact with only 3 configurations along the reaction, i.e., the reactant, TS and product, and the solvent could not be polarized due to being modeled using a non-polarizable forcefield. The CPMD-metadynamics computed the free energy surfaces (FESs) for the protonation step, as a function of collective variables defined in Figure 4.4, at different DMSO proportions. The FESs is shown in Figure 4.8 a, b, and c respectively. All FESs shows two defined wells representing the 5-HMF before and after being protonated. In addition, the free-energy profiles indicating activation barrier and ΔG_R at all DMSO concentrations are reported in Figure 4.8 d. Like that of Resolv, the reaction is exergonic, to a lesser extent though. The free energy barriers in pure water and 25 % wt. DMSO are only 40 kJmol-1 and 50 kJ/mol-1, respectively. However, higher DMSO proportion of 50 % wt. leads to a significant increase in the activation barrier to 137 kJmol-1. The reaction at 75% wt. DMSO is not reported here. This is because DMSO molecules create a shielding effect by preferentially solvating 5-HMF, thus preventing the access of the hydronium to the reactant. This is consistent with the previous literature that shows how the presence of DMSO reduces the stability of the hydronium ion near 5-HMF.[176] The reaction-free energy (ΔG_R) increases by 50 and 59 kJmol-1 with the addition of 25 and 50 % wt. DMSO, respectively, compared to pure water. Unlike the energy profiles for the protonation of 5-HMF obtained using Implisolv and Resolv, Dynasolv method, using CPMD-metadynamics, shows significant changes in reaction energetics as a function of solvent composition.
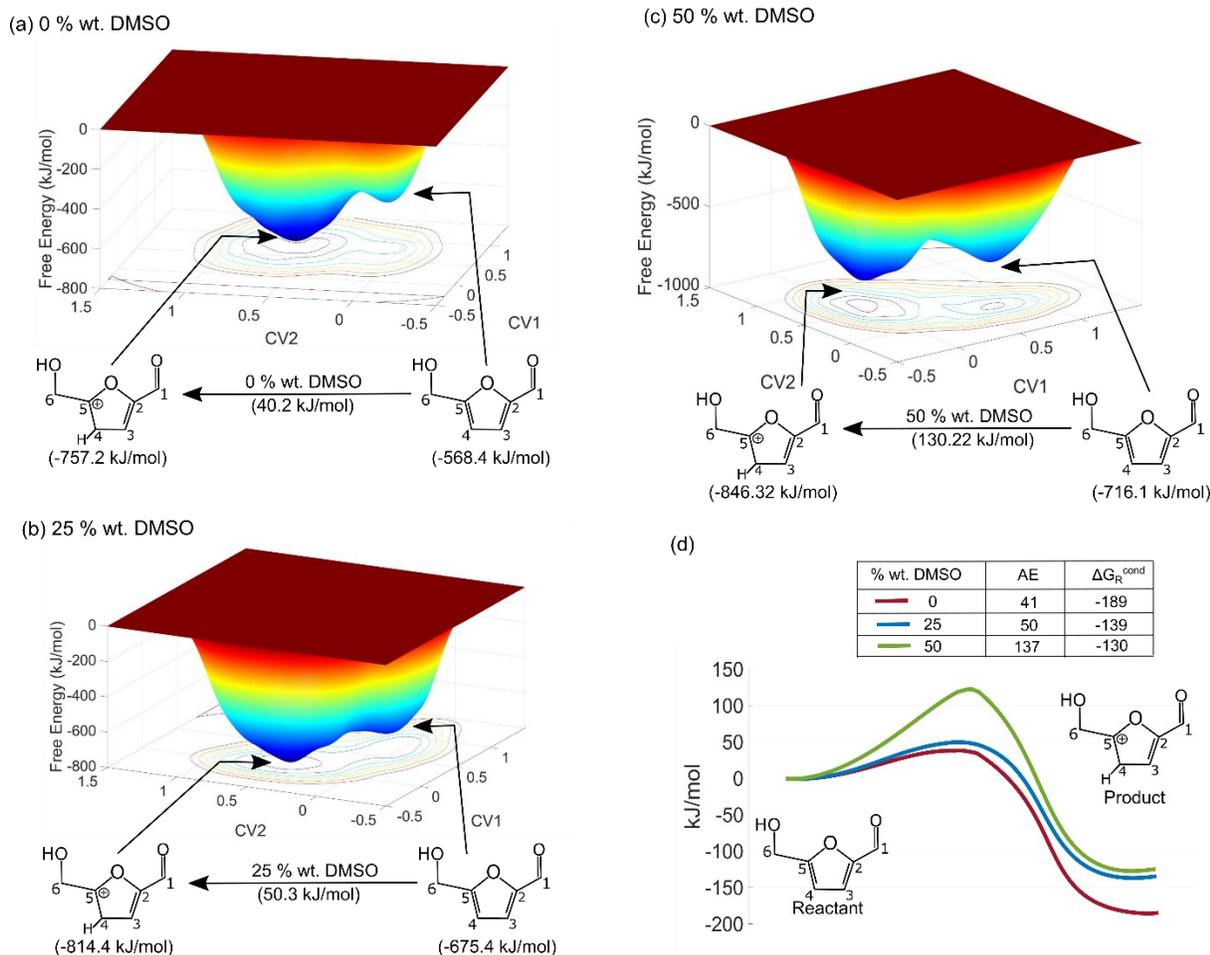
Figure 4.8 Metadynamics calculated free energy landscape for the 5-HMF C$_4$ protonation in the presence of explicit, quantum mechanically treated solvent molecules. Structures corresponding to two minima in the free energy surface are shown. The free energy barrier in the reaction step is shown below the arrow and free energies of the system before and after the protonation are shown in parentheses. Refer to Fig. 4.4 for the definition of collective variables CV1 and CV2. (a) 0% wt., (b) 25% wt., (c) 50% wt. of DMSO. (d) Free energy profile for the protonation step of 5-HMF. Free energy values for Ab Initio CPMD simulations at different concentrations of DMSO. Activation energies (AE) and Reaction Free Energies (ΔG_R) are reported in kJ/mol.

4.3.4 Solvent dynamics effects on reaction energetics

The charge redistribution in the reacting species during a chemical reaction significantly influences the immediate solvent environment. It can polarize the solvent, which in turn can induce solvent reorientation. According to Marcus' theory, there is an intrinsic barrier (λ) during a chemical reaction associated with solvent reorganization, also known as solvent reorganization energy. The activation free energy contains the free energy change of charge transfer and λ. The latter is the energy required to reorganize the solvent in equilibrium near the reactant to the product's structure.[177] For all DMSO proportions using Dynasolv, the kinetically most favorable pathway for the protonation of 5-HMF at all DMSO proportions, were projected

to obtain 2D free energy profiles. To determine λ, the reactant well (before protonation) is projected using numerical curve fitting to the product well (after protonation), as shown in Figure 4.9.
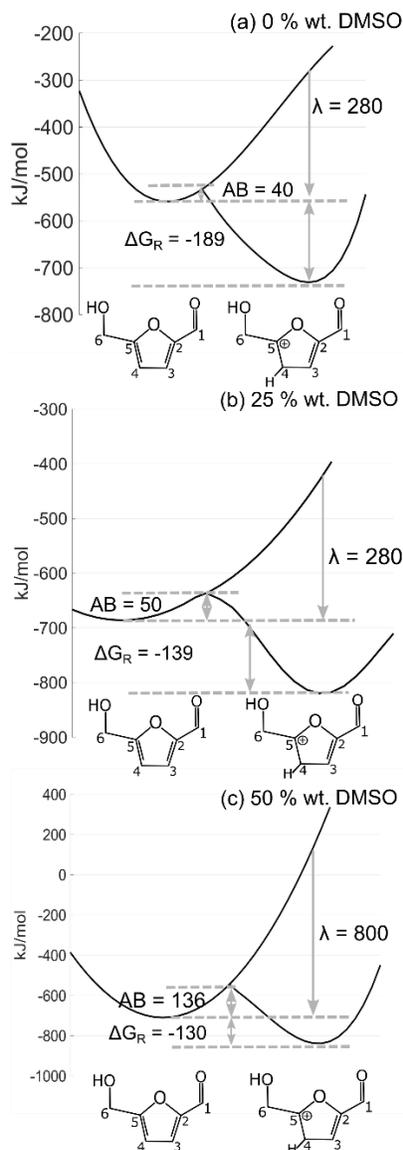
Figure 4.9 Free energy surface data projection in 2-D for the protonation of 5-HMF at all DMSO proportion. Marcus's reorganization energy λ changes significantly at 50 % wt. DMSO. All free energy values are in kJ/mol. (a) 0 % wt. DMSO, (b) 25 % wt. DMSO (c) 50 % wt. DMSO.

As observed in Figures 4.9 a and b, the solvent reorganization energy is ~280 kJ/mol in 0 and 25 % wt. DMSO. However, at 50 % wt. DMSO, λ increases to ~800 kJ/mol, and the activation barrier increases by ~ 86 kJmol-1. The λ values reported here are in the same range as those reported for the intramolecular hydride transfer reaction steps for the glucose and fructose dehydration mechanisms.[178] These results reveal the energy penalty associated with the solvent reorganization at 50 % wt. DMSO is significantly higher than lower DMSO proportions or pure water. Therefore, higher λ causes higher activation free energy

for the protonation reaction. In addition, higher λ values may mean more solvent reorganization during the reaction. It is thus possible that at higher DMSO proportions, solvent reorganization occurs over a longer timescale, possibly leading to non-equilibrium solvation.

4.3.5 non-equilibrium solvation

Nonequilibrium solvation occurs when reorientation dynamics of solvent molecules near the reactant is slower than that of the chemical reaction, resulting in a solvent configuration that is not fully relaxed/equilibrated along the reaction trajectory. To quantify the extent of solvent reorientation (also known as nuclear/molecular reorganization) at each DMSO proportion at a molecular scale, the orientational dipole vectors of the solvent molecules are computed by assigning partial charges (in the post-simulation analysis) to the solvent atoms in the simulation system along the CPMD-metadynamics trajectory. Figure 4.10 shows the time evolution of the $C_4 - H^+$ bond (black line) as the system undergoes protonation at different DMSO proportions. The angle of the solvent-orientational dipole vector is also reported over simulation time (grey lines) with respect to the x-axis of the simulation cell.
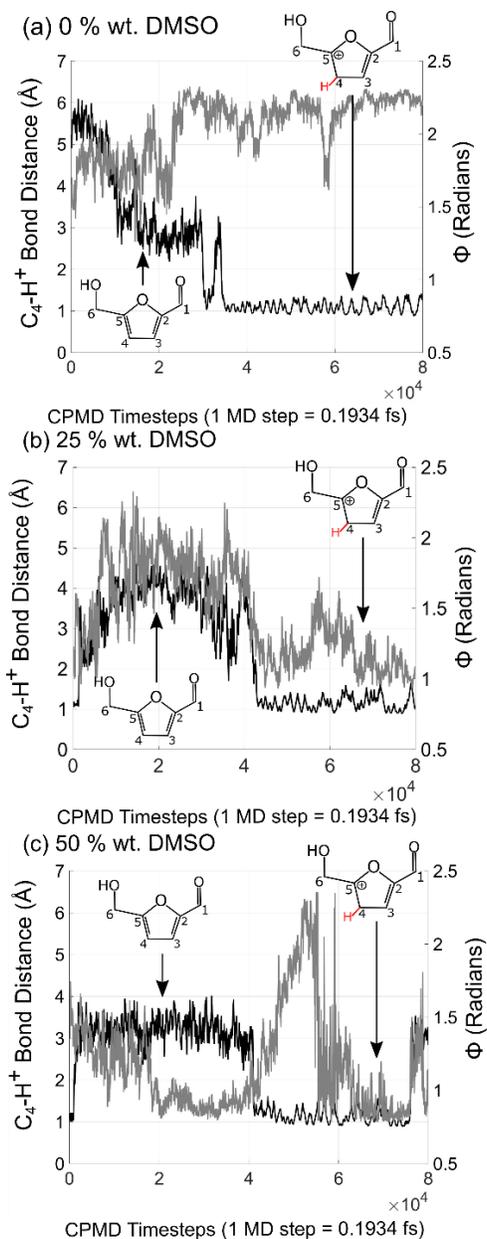
Figure 4.10 Variation in the orientation of the solvent dipole vector with respect to z-axis (or the angle that the solvent dipole vector makes with the z-axis of the simulation cell) during the protonation step in the CPMD-metadynamics trajectory. (a) 0 % wt. DMSO (b) 25 % wt. DMSO (c) 50 % wt. DMSO. The C4–H+ bond distance along the trajectory is also shown as a black line.

It can be seen in Figure 4.10 a for pure water that H^+ gradually moves towards C_4 until it binds after 30,000 MD steps (~6 ps). While the orientation of the dipole vector undergoes a sudden shift just before the protonation occurs, which appears to be a rapid response of the solvent molecules to reorient as the protonation reaction occurs (within the same time scale range), and a stable oscillating trend is observed afterward. Additionally, by increasing the DMSO proportion to 25% wt., it is observed in Figure 4.10 b how the reaction occurs after 40,000 MD steps (~8 ps). However, during the protonation, the orientation of the

dipole vector slightly shifts from around 1.75 to 1.25 Rad. This shift in the orientational dipole vector during the protonation of 5-HMF suggests molecular reorganization within the same timescale as that of the reaction. The reorientation in DMSO's presence occurs slightly slower than that in pure water. The reorganization of the solvent molecules is however significantly more for 50% wt. DMSO. as shown in Figure 4.10 c. The orientation of the dipole vector changes from 1 to 2.5 rad after the protonation reaction, over a time-period of ~2 ps. These results suggest that (i) the charge redistribution within the 5-HMF molecule due to protonation leads to reorientation of solvent molecules and that (ii) the molecular reorientation of the solvent occurs significantly slower at a higher DMSO proportion, which may lead to nonequilibrium solvation. Overall, not only a higher λ leads to the increase in activation-free energy for 50% wt. DMSO, but looking at the dynamics of solvent reorganization, it may also lead to non-equilibrium solvation conditions. In pure water as the solvent, water molecules reorient themselves relatively quickly, stabilizing the reacting species along the dynamic trajectory. However, at 50 % wt. DMSO and, to a lesser extent, at 25% wt. DMSO, the solvent system exhibits slower and more intense molecular reorganization, providing relatively lower stability to the reacting molecule.

The relatively smaller molecular reorganization and efficient stabilization associated with the pure water system is analogous to the improved stabilization enzymes provide to different charged groups via electrostatic interaction, as explained by Warshell and coworkers. They suggested that enzymes do not need significant molecular reorientation during the chemical reaction because of the enzyme's dipole preorientation towards the reactive molecule (enhanced charged stabilization compared to solvents). In that way, lower activation energies can be achieved compared to solvents, in which dipoles are less systematically oriented. For the protonation of 5-HMF, we observe that in pure water, the solvent undergoes lesser molecular reorganization compared to water-DMSO mixtures suggesting that either (i) DMSO molecules contribute to the more intense molecular reorganization, which in turn would lead to more significant orientational dipole vector alterations for the overall solvent, or (ii) the presence of DMSO molecules inhibit the water molecules from adapting quickly and efficiently (within the hydrogen bonded water network) to the changes in the reacting species, forcing them to undergo to a higher molecular reorganization. Hence, we analyze the contribution of each solvent to the overall molecular reorientation at different DMSO proportions. The orientational dipole vector is computed for water and DMSO separately. The orientational dipole vectors for each solvent are computed and plotted along the CPMD-metadynamics trajectory. Figure 4.11 show the orientational dipole vectors for water and DMSO, for all DMSO proportions.
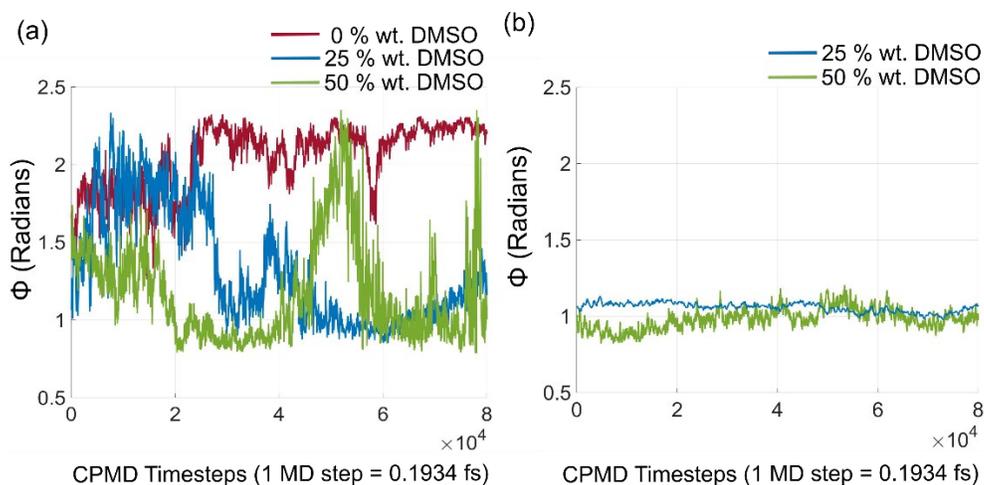
Figure 4.11 Variation in the orientation of the solvent dipole vector for the water and DMSO molecules separately with respect to z-axis (or the angle that the solvent dipole vector makes with the z-axis of the simulation cell) during the protonation step in the CPMD-metadynamics trajectory. (a) water molecules (b) DMSO molecules at different DMSO proportions.

The dynamics of the orientation of dipole vectors of water (Figure 4.11 a) are like the overall solvent orientational dipole vectors (Figure 4.10), while for DMSO, hardly any changes in the orientation are observed. These trends clearly show that the contribution to the overall solvent reorganization mainly comes from the water molecules, and not from DMSO. The change in the orientational dipole vectors for x and y axes are provided in section A 3.5. However, similar trends in the orientation of the solvent molecules are also observed. In principle, DMSO molecules are expected to show higher molecular reorganization due to higher polarizability (7.9 Å^3) than water molecules (1.4 Å^3) and their larger size. However, as shown in the volumetric maps in Figure 4.7, the probability of DMSO molecules in the nearest vicinity of the furanic ring of 5-HMF at 25 wt.% is low and even lower at 50% wt. DMSO (DMSO prefers to preferentially solvate the aldehyde group). Therefore, the changes in charge redistribution in the furanic ring of 5-HMF after protonation would mostly affect the nearby water molecules. In pure water, the relatively fast molecular reorientation of the water molecules near 5-HMF generates a systematic realignment of the water's dipoles restoring the H-bonds network almost instantly. However, it has been determined for water-DMF mixtures that DMF's oxygen atoms replace water's oxygen atoms in the H-bonds network. Still, the addition of DMF gradually disrupts the highly interconnected H-bonds network due to the lack of existing H-donating groups in the aprotic solvent molecules.[179] Due to the structural similarities between DMF and DMSO molecules, we suggest that the H-bonds water network also breaks down by adding DMSO in a similar fashion. These changes in the water interconnected H-bonds network by the addition of DMSO results in the loss of the ability of extensive water molecules' network to adapt almost instantaneously during and after the protonation reaction, and the fewer water molecules (now interacting with DMSO too) need to undergo a more extensive molecular rearrangement due to the hindrance caused by the DMSO molecules. The participation of DMSO oxygens to the H-bond network suggests that the water molecules stabilize the DMSO molecules so that despite being highly polarizable compared to water, DMSO do not significantly

reorient (also since they are not as close to the reactant as water) as shown in Figure 4.11 b. Thus, it can be inferred that the increase in activation free energy barrier of the protonation reaction, with the addition of DMSO in high proportions, is a combined effect of higher reorganization energy of water molecules and non-equilibrium solvation of the transition state of the reaction due to relatively sluggish dynamics of water molecules.

4.4 Conclusions

In this work, we systematically compare and analyse three solvent modeling approaches to investigate energetics of an acid catalyzed biomass reaction. The study focuses on the acid-catalyzed protonation of 5-HMF using water and DMSO as co-solvents. The computational models studied are: (i) implicit solvation using a dielectric continuum (Implisolv), (ii) explicit equilibrated solvation with a non-polarizable solvent and not considering reaction and solvent dynamics (Resolv), (iii) explicit solvation with reaction and solvent dynamics and polarizable solvent (Dynasolv). Using Implisolv, reaction free energies ΔG_R and activation free energies remained almost unchanged between water and DMSO. It cannot explain the experimentally observed differences in acid-catalyzed reactions of sugars and furans due to the presence of DMSO. In contrast, the Resolv method reveals significant differences in ΔG_R. The reaction was exergonic, but when DMSO proportion in the solvent medium was increased to 75 % wt., ΔG_R increased substantially. The activation free energy barriers computed using ReSolv and the Implicit solvation model were almost identical. Radial distribution function (RDF) based analysis suggests that DMSO competes with water to interact with the electrophilic site generated after the protonation reaction. 3-D local arrangements of the solvent molecules around the product conformer also reveal that increasing DMSO proportion from 50 to 75 % wt. resulted in DMSO displacing water molecules. In the context of acid-catalyzed protonation of 5-HMF, we suggest that the presence of DMSO near the electrophilic C_5^+ prevents the water molecules from stabilizing C_5^+, thus resulting in making the reaction thermodynamically less favorable at higher DMSO proportions. Finally, Dynasolv method that considers reaction and solvent dynamics, and solvent polarizability was able to reveal changes in activation energy and ΔG_R as a function of solvent composition. The free energy barrier in pure water and at 25 % wt. DMSO is 40 kJmol-1 and 50 kJ/mol-1, respectively. This is almost half of that computed using Implisolv and Resolv (~90 kj/mol). However, a higher DMSO proportion of 50 % wt. leads to a significant increase in the activation barrier to 137 kJmol-1. Also, ΔG_R increases by 50 and 59 kJmol-1 by adding 25 and 50 % wt. DMSO, respectively, compared to that in pure water. Thus, a solvent model that considers reaction and solvent dynamics and solvent polarizability predicted higher activation barriers and unfavorable reaction free energies with increasing DMSO proportion, an effect significantly pronounced beyond 50 wt.% DMSO. Our analysis revealed that Marcus' solvent reorganization energy λ increases significantly from 25 to 50 % wt. DMSO. Solvent orientation dipole vectors are computed to quantify the reorganization of the solvent with changing DMSO proportion. The results show that the dynamics of the molecular reorientation the solvent are slower at 50 wt.% DMSO in the solvent and that would also lead to nonequilibrium solvation of the transition state. Such

nonequilibrium solvation of the transition state would also contribute to the higher activation-free energy barriers at higher DMSO proportions in the solvent media. Analysis of individual water and DMSO reorientation dynamics suggests that water is the major contributor to solvent reorganization and that DMSO molecules contribute negligibly. It is suggested that the ability of interconnected water molecules to adapt to the change in charge distribution of the reacting species gets hindered due to the presence of DMSO molecules competing with water in the solvent environment and that results in higher reorganization energy and slower dynamics of neighboring water molecules.

Chapter 5 Conclusions and perspectives

This thesis employs a multiscale molecular modelling methodology to understand humins chemistry and solvent effects in the condensed phase acid-catalyzed biomass conversion process. Quantum mechanical calculations are performed to compute the reaction pathways, activation, and reaction-free energies associated with the elementary reaction steps in the 5-HMF initiated, acid-catalyzed humins formation. The detailed reaction mechanism is elucidated, involving first, the dehydration of 5-HMF to DHH, the keto-enol tautomerization of DHH, and the derivative to 4 possible enols. Each enol, in turn, can react with 5-HMF again via aldol addition, and finally, the dimers undergo a condensation reaction. For the hydration of 5-HMF to DHH, the free energy profile reveals that it is an endergonic reaction, with free energy slightly more than 30 kJ/mol, and electrophilic addition reaction is carried out in two steps, the protonation and subsequent hydroxylation of the 5-HMF furan ring, being the first one the rate-determining step of this reaction. From the keto-enol tautomerization of DHH to the formation of the dimers, the aldol addition reaction is identified to be the rate-determining step with the highest energy barrier along the entire reaction pathway. The barriers for the aldol addition reaction of all enols with the 5-HMF are between 90 and 100 kJ/mol. However, the overall reaction is slightly endergonic. This study provides kinetic data (i.e., reaction-free energies and barriers) for the subsequent kinetic modelling of the humins polymerization reaction, aiming to optimize the reaction conditions while minimizing the formation of humins. The computed and suggested polymerization mechanisms are in excellent agreement with the experimental spectroscopic data reported in the literature and discussed in this work. This mechanistic understanding of initiation reactions for humins paves the way to systematically screen and design solvent and catalyst systems that can hinder humins formation and, thus, improve the selectivity towards desired reactions.

This work also elucidates the two different roles of the solvent in a chemical reaction. The solvent can alter the interaction between the reacting species and the catalyst by preferential physical solvation/interaction, enhancing the selectivity towards desired products and protecting the products from further degrading into humins. The solvent can also directly participate in the chemistry by altering reaction-free and activation-free energy barriers via solvent dynamics and non-equilibrium solvation. A computational toolbox consisting of molecular mechanics-based molecular dynamics and well-tempered metadynamics methods is implemented to elucidate the effect of aprotic solvent on the interaction of reactants and products with the catalyst (hydronium ion). 5-HMF (the model product) and fructose (the model reactant) in acidic water and water-DMSO mixtures are taken as model simulation systems. Analogous to heterogeneous catalysis, where the interaction between the reactant and the catalyst and the competition between reactant and solvent to interact with the catalyst surface affect the chemistry, here, it is demonstrated that the presence of aprotic solvent molecules like DMSO alters the interaction of hydronium ions (the catalyst) with the reactant fructose and 5-HMF molecules in a homogeneous environment. Adding DMSO leads to enhanced interaction between the hydronium ions and the hydroxyl group of fructose. In contrast, for 5-HMF, the preferential interaction of the furan ring of 5-HMF with DMSO over water molecules and hydronium ions is

observed. Also, well-tempered metadynamics simulations determine the relative stability of the hydronium ion in the bulk solvent phase, as compared to near fructose and 5-HMF, as a function of solvent composition. It is found that the addition of DMSO improves the stabilization of the hydronium ions in the first solvation shell of fructose compared to that in the bulk solvent. On the other hand, hydronium ions become less stable near 5-HMF as the concentration of DMSO increases. In the context of acid-catalyzed degradation of 5-HMF to humins, the decreasing stability of the hydronium ions in the first solvation shell of 5-HMF reduces the likelihood of the side reaction leading to humins formation.

This thesis also provides an understanding of the role of solvent in altering the kinetics and thermodynamics of the protonation of 5-HMF, the first step leading to DHH and further humins formation. DFT calculations using an implicit solvation model do not differ in reaction kinetics and thermodynamics, irrespective of the solvent (water or DMSO). The reaction-free energy using explicit equilibrated solvation shows a significant increase in the endergonicity of the reaction at higher co-solvent concentrations. Molecular analysis revealed that the change in reaction-free energy is due to relatively poor product stabilization in the presence of co-solvent. On the other hand, no change in reaction kinetics, as a function of solvent composition, could be observed even with explicit equilibrated solvent molecules. The effect of the solvent environment on free energy barriers was only captured when solvent and reaction dynamics were considered and when the solvent molecules were polarizable. It is determined that the increase in activation- and reaction-free energies with the addition of the co-solvent is due to changes in the solvent reorganization energy and non-equilibrium solvation. It is observed that Marcus' solvent reorganization energy λ increases significantly at higher co-solvent concentrations. Moreover, water reorientation dynamics is the main contributor to the overall solvent reorganization, and co-solvent molecules contribute by altering the ability of water molecules to adapt to the change in charge distribution of the reacting species via their interaction with water. Since protonation and deprotonation are the most inevitable and usually rate-limiting steps in various reactions, and 5-HMF is also structurally like many furanic, cellulose-derived moieties, these findings are not limited to humins formation chemistry. However, they can be extended to other biomass transformation reactions in the condensed phase, particularly charge transfer reactions that can be highly susceptible to solvent composition.

In summary, this thesis provides an extensive understanding of the molecular level of the chemistry involved in the mechanisms leading to humins. This work provides information about energy barriers, reaction-free energies and the reaction mechanism for the early stages and extended step growth oligomerization of humins laying the foundations to develop kinetic Monte Carlo or microkinetic type models for the entire biomass moieties (i.e., glucose or fructose) conversion to 5-HMF, including humins formation chemistry. This future work also helps to suggest process conditions to optimize 5-HMF yield and minimize the formation of humins. Additionally, this thesis discloses the physical and chemical effects of adding aprotic solvent during the acid-catalyzed transformation of biomass molecules. Understanding the effects of aprotic solvent leading to yield enhancement during the transformation of biomass molecules to chemicals helps

the design and develop novel heterogeneous catalysts with functionalities that can selectively hinder humins chemistry. Finally, testing and revisiting key biomass reactions with different solvents and different solvent modelling approaches is necessary to have a better understanding at the molecular level of the effects of aprotic solvents on the chemistry involved in acid-catalyzed biomass conversion and identify the critical descriptors that can be implemented in computational tools (i.e. machine learning) leading to a less computational intensive solvent screening approach and selection helping to boost the deployment of biomass to chemicals routes.

References

1. McKendry, P. Energy production from biomass (part 1): Overview of biomass. *Bioresour Technol* **83**, 37–46 (2002).

2. Srivastava, R. K. *et al.* Biomass utilization and production of biofuels from carbon neutral materials. *Environmental Pollution* **276**, (2021).

3. Blanco, G. *et al. Bioenergy for the energy transition: Ensuring sustainability and overcoming barriers. Bharadwaj Kummamuru (World Bioenergy Association* (Jack Saddler, 2022).

4. Lyons, M., Durrant, P. & Kochhar, K. *Reaching Zero with Renewables: Capturing carbon*. (International Renewable Energy Agency, 2021).

5. Technology Roadmap Delivering Sustainable Bioernergy. (2017).

6. Oecd, Iea & Fao. *Technology Roadmap : How2Guide for Bioenergy Roadmap - Development and Implementation*. (2017).

7. Wyman, C. E. *Handbook on Bioethanol: Production and Utilization*. (CRC Press, 1996).

8. Halder, P., Azad, K., Shah, S. & Sarker, E. Prospects and technological advancement of cellulosic bioethanol ecofuel production. in *Advances in Eco-Fuels for a Sustainable Environment* 211–236 (Elsevier, 2018).

9. Zhou, C. H., Xia, X., Lin, C. X., Tong, D. S. & Beltramini, J. Catalytic conversion of lignocellulosic biomass to fine chemicals and fuels. *Chem Soc Rev* **40**, 5588–5617 (2011).

10. International Renewable Energy Agency (IRENA). *Innovation Outlook Advanced Liquid Biofuels*. (2016).

11. Technologies Office, B. *Bioenergy Technologies Office 2019 R&D State of Technology*. (2019).

12. RFA. *2023 Ethanol Industry Outlook*. (2023).

13. International Energy Agency, I. *World Energy Outlook 2022*. (2022).

14. Yousuf, A. Biodiesel from lignocellulosic biomass - Prospects and challenges. *Waste Management* **32**, 2061–2067 (2012).

15. W.B. Betts. *Springer Series in Applied Biology Biodegradation: Natural and Synthetic Materials*. (Springer London, 2012).

16. Kikas, T. *et al.* Basis of energy crop selection for biofuel production: Cellulose vs. lignin. *Int J Green Energy* **13**, 49–54 (2016).

17. Yousuf, A., Pirozzi, D. & Sannino, F. Fundamentals of lignocellulosic biomass. in *Lignocellulosic Biomass to Liquid Biofuels* 1–15 (Elsevier, 2019).

18. Gibson, L. J. The hierarchical structure and mechanics of plant materials. *J R Soc Interface* **9**, 2749–2766 (2012).

19. Zhang, Y. H. P. *et al.* Fractionating recalcitrant lignocellulose at modest reaction conditions. *Biotechnol Bioeng* **97**, 214–223 (2007).

20. Tursi, A. A review on biomass: Importance, chemistry, classification, and conversion. *Biofuel Research Journal* **6**, 962–979 (2019).

21. Sun, Y. & Cheng, J. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresour Technol* **83**, 1–11 (2002).

22. Imai, M., Ikari, K. & Suzuki, I. High-performance hydrolysis of cellulose using mixed cellulase species and ultrasonication pretreatment. *Biochem Eng J* **17**, 79–83 (2004).

23. Harmsen, P. F. H. (Paulien). *Literature review of physical and chemical pretreatment processes for lignocellulosic biomass*. (Wageningen UR, Food & Biobased Research, 2010).

24. Francesco Dalena, A. B. and C. R. *Bioenergy Systems for the Future Prospects for Biofuels and Biohydrogen*. (ELSEVIER, 2017).

25. Gupta, V. K. & Tuohy, M. G. *Biofuel technologies: Recent developments*. *Biofuel Technologies: Recent Developments* (Springer-Verlag Berlin Heidelberg, 2014).

26. Siva Reddy, V., Kaushik, S. C., Ranjan, K. R. & Tyagi, S. K. State-of-the-art of solar thermal power plants - A review. *Renewable and Sustainable Energy Reviews* **27**, 258–273 (2013).

27. Kumar, M., Olajire Oyedun, A. & Kumar, A. A review on the current status of various hydrothermal technologies on biomass feedstock. *Renewable and Sustainable Energy Reviews* **81**, 1742–1770 (2018).

28. López Barreiro, D., Prins, W., Ronsse, F. & Brilman, W. Hydrothermal liquefaction (HTL) of microalgae for biofuel production: State of the art review and future prospects. *Biomass Bioenergy* **53**, 113–127 (2013).

29. Dutta, A. *et al. Ex Situ Catalytic Fast Pyrolysis of Lignocellulosic Biomass to Hydrocarbon Fuels: 2020 State of Technology*. (2022).

30. Kookana, R. S., Sarmah, A. K., Zwieten, L. Van & Krull, E. *Biochar Application to Soil: Agronomic and Environmental Benefits and Unintended Consequences*. *Advances in Agronomy* vol. 112 (Elsevier Inc., 2011).

31. Rahimpour, A. *et al.* Novel functionalized carbon nanotubes for improving the surface properties and performance of polyethersulfone (PES) membrane. *Desalination* **286**, 99–107 (2012).

32. Molino, A., Chianese, S. & Musmarra, D. Biomass gasification technology: The state-of-the-art overview. *Journal of Energy Chemistry* **25**, 10–25 (2016).

33. Dutta, A. *et al. Process Design and Economics for Conversion of Lignocellulosic Biomass to Ethanol: Thermochemical Pathway by Indirect Gasification and Mixed Alcohol Synthesis*. (2007).

34. Dutta, A. *et al. Process Design and Economics for Conversion of Lignocellulosic Biomass to Ethanol: Thermochemical Pathway by Indirect Gasification and Mixed Alcohol Synthesis*. (2007).

35. Tan, E. C. D. *et al. Process Design and Economics for the Conversion of Lignocellulosic Biomass to Hydrocarbons via Indirect Liquefaction: Thermochemical Research Pathway to High-Octane Gasoline Blendstock Through Methanol/Dimethyl Ether Intermediates*. (2012).

36. Brethauer, S. & Studer, M. H. Biochemical conversion processes of lignocellulosic biomass to fuels and chemicals - A review. *Chimia (Aarau)* **69**, 572–581 (2015).

37. Ward, A. J., Hobbs, P. J., Holliman, P. J. & Jones, D. L. Optimisation of the anaerobic digestion of agricultural resources. *Bioresour Technol* **99**, 7928–7940 (2008).

38. Olsson, L. & Hahn-Higerdal, B. *Fermentation of lignocellulosic hydrolysates for ethanol production*. (1996).

39.   Humbird, D. *et al. Process Design and Economics for Biochemical Conversion of Lignocellulosic Biomass to Ethanol: Dilute-Acid Pretreatment and Enzymatic Hydrolysis of Corn Stover*. (2002).

40.   Nagamatsu, S. & Ohara-Imaizumi, M. Cell biology: IP7 debut in insulin release. *Science* vol. 318 1249–1250 (2007).

41.   International Energy Agency IEA. *The Future of Petrochemicals Towards more sustainable plastics and fertilisers Together Secure Sustainable* (2018).

42.   International Energy Agency IEA. *Renewables 2021 - Analysis and forecast to 2026*. 2021).

43.   Pires, J. C. M., Martins, F. G. & Simoes, M. Chemical Engineering Research and Design Recent developments on carbon capture and storage : An overview. *Chemical Engineering Research and Design* **89**, 1446–1460 (2011).

44.   Environment Agency, G. *Power-to-Liquids Potentials and Perspectives for the Future Supply of Renewable Aviation Fuel*. (2020).

45.   ICF International. *Power-to-Liquids Roadmap: Fuelling the Aviation Energy Transition in the United Arab Emirates*. (2022).

46.   Geoff Bell, A. *et al. IEA BIOENERGY Task42 BIOREFINING Sustainable and synergetic processing of biomass into marketable food & feed ingredients, products (chemicals, materials) and energy (fuels, power, heat)* (2014).

47.   Coker, A. K. *Ludwig's Applied Process Design for Chemical and Petrochemical Plants*. (Elsevier Inc., 2010).

48.   Han, X., Wang, H., Zeng, Y. & Liu, J. Advancing the application of bio-oils by co-processing with petroleum intermediates: A review. *Energy Conversion and Management: X* **10**, (2021).

49.   Wang, C., Venderbosch, R. & Fang, Y. Co-processing of crude and hydrotreated pyrolysis liquids and VGO in a pilot scale FCC riser setup. *Fuel Processing Technology* **181**, 157–165 (2018).

50.   Erin Chan. Converting a petroleum diesel refinery for renewable diesel. *Hydrocarbon Processing* (2021).

51.   Chheda, J. N., Huber, G. W. & Dumesic, J. A. Liquid-phase catalytic processing of biomass-derived oxygenated hydrocarbons to fuels and chemicals. *Angewandte Chemie - International Edition* **46**, 7164–7183 (2007).

52.   Werpy, T. & Petersen, G. *Top Value Added Chemicals from Biomass Volume I. US NREL/* (2004).

53.   U.S. Department of Energy Efficiency and Renewable Energy. *Top Value-Added Chemicals from Biomass Volume I—Results of Screening for Potential Candidates from Sugars and Synthesis Gas*. (2004).

54.   Onel, O., Niziolek, A. M. & Floudas, C. A. Integrated biomass and fossil fuel systems towards the production of fuels and chemicals: state of the art approaches and future challenges. *Curr Opin Chem Eng* **9**, 66–74 (2015).

55.   Hayes, G. *et al.* Polymers without Petrochemicals: Sustainable Routes to Conventional Monomers. *Chem Rev* **123**, 2609–2734 (2023).

56.   Gallezot, P. Catalytic Conversion of Biomass: Challenges and Issues. *ChemSusChem* **1**, 734–737 (2008).

57. Chheda, J. N., Roma, Y. & Dumesic, J. A. Production of 5-hydroxymethylfurfural and furfural by dehydration of biomass-derived mono- and polysaccharides. *Green Chemistry* 342–350 (2007).

58. Román-Leshkov, Y. & Dumesic, J. A. Solvent effects on fructose dehydration to 5-hydroxymethylfurfural in biphasic systems saturated with inorganic salts. *Top Catal* **52**, 297–303 (2009).

59. Asghari, F. S. & Yoshida, H. Kinetics of the decomposition of fructose catalyzed by hydrochloric acid in subcritical water: Formation of 5-hydroxymethylfurfural, levulinic, and formic acids. *Ind Eng Chem Res* **46**, 7703–7710 (2007).

60. Chheda, J. N. & Dumesic, J. A. An overview of dehydration, aldol-condensation, and hydrogenation processes for production of liquid alkanes from biomass-derived carbohydrates. *Catal Today* **123**, 59–70 (2007).

61. Huber, G. W., Cortright, R. D. & Dumesic, J. A. Renewable Alkanes by Aqueous-Phase Reforming of Biomass-Derived Oxygenates. **2**, 1575–1577 (2004).

62. Sumerskii, I. V., Krutov, S. M. & Zarubin, M. Y. Humin-Like substances formed under the conditions of industrial hydrolysis of wood. *Russian Journal of Applied Chemistry* **83**, 320–327 (2010).

63. Carniti, P., Gervasini, A. & Marzo, M. Absence of expected side-reactions in the dehydration reaction of fructose to HMF in water over niobic acid catalyst. *Catal Commun* **12**, 1122–1126 (2011).

64. Van Zandvoort, I. *et al.* Structural characterization of 13C-enriched humins and alkali-treated 13C humins by 2D solid-state NMR. *Green Chemistry* **17**, 4383–4392 (2015).

65. Agarwal, S., van Es, D. & Heeres, H. J. Catalytic pyrolysis of recalcitrant, insoluble humin byproducts from C6 sugar biorefineries. *J Anal Appl Pyrolysis* **123**, 134–143 (2017).

66. Mija, A., van der Waal, J. C., Pin, J. M., Guigo, N. & de Jong, E. Humins as promising material for producing sustainable carbohydrate-derived building materials. *Constr Build Mater* **139**, 594–601 (2017).

67. Kim, Y. *et al.* Prediction of hydroxymethylfurfural yield in glucose conversion through investigation of lewis acid and organic solvent effects. *ACS Catal* **10**, 14707–14721 (2020).

68. Patil, S. K. R. & Lund, C. R. F. Formation, and growth of humins via aldol addition and condensation during acid-catalyzed conversion of 5-hydroxymethylfurfural. *Energy and Fuels* **25**, 4745–4755 (2011).

69. Tsilomelekis, G. *et al.* Molecular structure, morphology and growth mechanisms and rates of 5-hydroxymethyl furfural (HMF) derived humins. *Green Chemistry* **18**, 1983–1993 (2016).

70. Patil, S. K. R., Heltzel, J. & Lund, C. R. F. Comparison of structural features of humins formed catalytically from glucose, fructose, and 5-hydroxymethylfurfuraldehyde. *Energy and Fuels* **26**, 5281–5293 (2012).

71. Patil, S. K. R. & Lund, C. R. F. Formation and growth of humins via aldol addition and condensation during acid-catalyzed conversion of 5-hydroxymethylfurfural. *Energy and Fuels* **25**, 4745–4755 (2011).

72. Van Zandvoort, I. *et al.* Formation, molecular structure, and morphology of humins in biomass conversion: Influence of feedstock and processing conditions. *ChemSusChem* **6**, 1745–1758 (2013).

73.   Cheng, Z. *et al.* Structural analysis of humins formed in the Brønsted acid catalyzed dehydration of fructose. *Green Chemistry* **20**, 997–1006 (2018).

74.   Cao, F. *et al.* Environmental Science Dehydration of cellulose to levoglucosenone using polar aprotic solvents. *Energy Environ. Sci.* **8**, 1808–1815 (2015).

75.   Jessop, P. G. Green Chemistry. *Green Chem.* **13**, 1391–1398 (2011).

76.   Komarova, A. O., Dick, G. R. & Luterbacher, J. S. produced from renewable biomass †. *Green Chem* **23**, 4790–4799 (2021).

77.   Li, Y., Lu, X., Yuan, L. & Liu, X. Fructose decomposition kinetics in organic acids-enriched high temperature liquid water. *Biomass Bioenergy* **33**, 1182–1187 (2009).

78.   Wang, W. *et al.* An efficient process for dehydration of xylose to furfural catalyzed by inorganic salts in water / dimethyl sulfoxide system. *Chinese Journal of Catalysis* **35**, 741–747 (2014).

79.   Dias, A. S., Pillinger, M. & Valente, A. A. Dehydration of xylose into furfural over micro-mesoporous sulfonic acid catalysts. **229**, 414–423 (2005).

80.   Molina, M. J. C., Mariscal, R., Ojeda, M. & Granados, M. L. Bioresource Technology Cyclopentyl methyl ether: A green co-solvent for the selective dehydration of lignocellulosic pentoses to furfural. *Bioresour Technol* **126**, 321–327 (2012).

81.   Román-Leshkov, Y., Chheda, J. N. & Dumesic, J. A. Phase modifiers promote efficient production of hydroxymethylfurfural from fructose. *Science (1979)* **312**, 1933–1937 (2006).

82.   Chheda, J. N., Román-Leshkov, Y. & Dumesic, J. A. Production of 5-hydroxymethylfurfural and furfural by dehydration of biomass-derived mono- and polysaccharides. *Green Chemistry* **9**, 342–35 (2007).

83.   Brown, D. W., Floyd, A. J., Kinsman, R. G. & Roshan-ali, Y. Dehydration Reactions of Fructose in Non-aqueous Media. *J. Chem. Tech. Biotechno* **32**, 920–924 (1982).

84.   Szmant, H. H. & Chundury, D. D. The Preparation of 5-Hydroxymethylfurfuraldehyde from High Fructose Corn Syrup and Other Carbohydrates. *J. Chem. Tech. Biotechnol.* **31**, 135–145 (1981).

85.   D'Agostino, C. *et al.* Solvent effect and reactivity trend in the aerobic oxidation of 1,3-propanediols over gold supported on titania: Nmr diffusion and relaxation studies. *Chem. Eur. J.* **19**, 11725–11732 (2013).

86.   Gounder, R. & Davis, M. E. Monosaccharide and disaccharide isomerization over Lewis acid sites in hydrophobic and hydrophilic molecular sieves. *J Catal* **308**, 176–188 (2013).

87.   Gounder, R. & Davis, M. E. Titanium-Beta Zeolites Catalyze the Stereospecific Isomerization of D-Glucose to L-Sorbose via Intramolecular C5−C1 Hydride Shift. *ACS Catal.* **3**, 1469–1476 (2013).

88.   Christianson, J. R., Caratzoulas, S. & Vlachos, D. G. Computational Insight into the Effect of Sn-Beta Na Exchange and Solvent on Glucose Isomerization and Epimerization. *ACS Catal* **5**, 5256–5263 (2015).

89.   Varghese, J. J. & Mushrif, S. H. Origins of complex solvent effects on chemical reactivity and computational tools to investigate them: A review. *React Chem Eng* **4**, 165–206 (2019).

90.   McManus, I. *et al.* Effect of solvent on the hydrogenation of 4-phenyl-2-butanone over Pt based catalysts. *J Catal* **330**, 344–353 (2015).

91. Bertero, N. M., Apesteguía, C. R. & Marchi, A. J. Catalytic and kinetic study of the liquid-phase hydrogenation of acetophenone over Cu/SiO2 catalyst. *Appl Catal A Gen* **349**, 100–109 (2008).

92. Walker, T. W. *et al.* Universal kinetic solvent effects in acid-catalyzed reactions of biomass-derived oxygenates. *Energy Environ Sci* **11**, 617–628 (2018).

93. Mellmer, M. A. *et al.* Solvent-enabled control of reactivity for liquid-phase reactions of biomass-derived compounds. *Nat Catal* **1**, 199–207 (2018).

94. Mellmer, M. A. *et al.* Solvent effects in acid-catalyzed biomass conversion reactions. *Angewandte Chemie - International Edition* **53**, 11872–11875 (2014).

95. Kelly, C. P., Cramer, C. J. & Truhlar, D. G. Single-ion solvation free energies and the normal hydrogen electrode potential in methanol, acetonitrile, and dimethyl sulfoxide. *Journal of Physical Chemistry B* **111**, 408–422 (2007).

96. Marcus, Y. Gibbs energies of transfer of anions from water to mixed aqueous organic solvents. *Chem Rev* **107**, 3880–3897 (2007).

97. Warshel, A. Energetics of enzyme catalysis. *Proc Natl Acad Sci U S A* **75**, 5250–5254 (1978).

98. Vasudevan, V. & Mushrif, S. H. Insights into the solvation of glucose in water, dimethyl sulfoxide (DMSO), tetrahydrofuran (THF) and N,N-dimethylformamide (DMF) and its possible implications on the conversion of glucose to platform chemicals. *RSC Adv* **5**, 20756–20763 (2015).

99. Mushrif, S. H., Caratzoulas, S. & Vlachos, D. G. Understanding solvent effects in the selective conversion of fructose to 5-hydroxymethyl-furfural: A molecular dynamics investigation. *Physical Chemistry Chemical Physics* **14**, 2637–2644 (2012).

100. Bahmanpour, A. M., Hoadley, A. & Tanksale, A. Formaldehyde production via hydrogenation of carbon monoxide in the aqueous phase. *Green Chemistry* **17**, 3500–3507 (2015).

101. Bahmanpour, A. M., Hoadley, A., Mushrif, S. H. & Tanksale, A. Hydrogenation of carbon monoxide into formaldehyde in liquid media. *ACS Sustain Chem Eng* **4**, 3970–3977 (2016).

102. Amarasekara, A. S., Williams, L. T. D. & Ebede, C. C. Mechanism of the dehydration of d-fructose to 5-hydroxymethylfurfural in dimethyl sulfoxide at 150 °C: an NMR study. *Carbohydr Res* **343**, 3021–3024 (2008).

103. Weingarten, R. *et al.* Selective conversion of cellulose to hydroxymethylfurfural in polar aprotic solvents. *ChemCatChem* **6**, 2229–2234 (2014).

104. Pagán-Torres, Y. J., Wang, T., Gallo, J. M. R., Shanks, B. H. & Dumesic, J. A. Production of 5-hydroxymethylfurfural from glucose using a combination of lewis and brønsted acid catalysts in water in a biphasic reactor with an alkylphenol solvent. *ACS Catal* **2**, 930–934 (2012).

105. Schwartz, T. J. & Bond, J. Q. A thermodynamic and kinetic analysis of solvent-enhanced selectivity in monophasic and biphasic reactor systems. *Chemical Communications* **53**, 8148–8151 (2017).

106. Hunter-Cevera, K. R. *et al.* Formaldehyde stabilization facilitates lignin monomer production during biomass depolymerization. *Science (1979)* **354**, 326–329 (2016).

107. Cheng, Z. *et al.* Structural analysis of humins formed in the Brønsted acid catalyzed dehydration of fructose. *Green Chemistry* **20**, 997–1006 (2018).

108. Dee, S. J. & Bell, A. T. A study of the acid-catalyzed hydrolysis of cellulose dissolved in ionic liquids and the factors influencing the dehydration of glucose and the formation of humins. *ChemSusChem* **4**, 1166–1173 (2011).

109. Horvat, J., Klaic, B., Metelko, B. & Sunjic, V. Mechanism of levulinic acid formation in acid catalyzed hydrolysis of 2-(hydroxymethyl) furan and 5-(hydroxymethyl)-2-furancarboxaldehyde. *Croatica Chemica Acta* **59**, 429–438 (1986).

110. Frisch, Æ., Plata, R. E. & Singleton, D. A. Gaussian 09W Reference. *J. Am. Chem. Soc.* **137**, 3811–3826 (2009).

111. Arora, J. S., Chew, J. W. & Mushrif, S. H. Influence of Alkali and Alkaline-Earth Metals on the Cleavage of Glycosidic Bond in Biomass Pyrolysis: A DFT Study Using Cellobiose as a Model Compound. *Journal of Physical Chemistry A* **122**, 7646–7658 (2018).

112. Mayes, H. B. & Broadbelt, L. J. Unraveling the reactions that unravel cellulose. *Journal of Physical Chemistry A* **116**, 7098–7106 (2012).

113. Park, M., Shin, I., Singh, N. J. & Kim, K. S. Eigen and Zundel forms of small, protonated water clusters: Structures and infrared spectra. *Journal of Physical Chemistry A* **111**, 10692–10702 (2007).

114. Decka, D., Schwaab, G. & Havenith, M. A THz/FTIR fingerprint of the solvated proton: Evidence for Eigen structure and Zundel dynamics. *Physical Chemistry Chemical Physics* **17**, 11898–11907 (2015).

115. Yang, Z. Z., Ding, Y. L. & Zhao, D. X. Insight into Markovnikov reactions of alkenes in terms of ab initio and molecular face theory. *ChemPhysChem* **9**, 2379–2389 (2008).

116. Hinchliffe, A. Ab initio SCF MO study of the hydrogen-bonded dimers of cyanogen with HF and HCl. *Chem Phys Lett* **87**, 417–419 (1982).

117. Chheda, J. N., Román-Leshkov, Y. & Dumesic, J. A. Production of 5-hydroxymethylfurfural and furfural by dehydration of biomass-derived mono- and polysaccharides. *Green Chemistry* **9**, 342–35 (2007).

118. Sun, Q. *et al.* Creating solvation environments in heterogeneous catalysts for efficient biomass conversion. *Nat Commun* **9**, (2018).

119. Tsilomelekis, G., Josephson, T. R., Nikolakis, V. & Caratzoulas, S. Origin of 5-hydroxymethylfurfural stability in water/dimethyl sulfoxide mixtures. *ChemSusChem* **7**, 117–126 (2014).

120. Bicker, M., Kaiser, D., Ott, L. & Vogel, H. Dehydration of D-fructose to hydroxymethylfurfural in sub- and supercritical fluids. *Journal of Supercritical Fluids* **36**, 118–126 (2005).

121. Mok, W. S. L., Antal, M. J. & Varhegyi, G. Productive and Parasitic Pathways in Dilute Acid-Catalyzed Hydrolysis of Cellulose. *Ind Eng Chem Res* **31**, 94–100 (1992).

122. Caratzoulas, S. & Vlachos, D. G. Converting fructose to 5-hydroxymethylfurfural: A quantum mechanics/molecular mechanics study of the mechanism and energetics. *Carbohyd Res* **346**, 664–672 (2011).

123. Amarasekara, A. S. & Razzaq, A. Mechanism of 1-(1-propylsulfonic)-3-methylimidazolium chloride catalyzed transformation of d-glucose to 5-hydroxymethylfurfural in DMSO: An NMR study. *Carbohyd Res* **386**, 86–91 (2014).

124. Mellmer, M. A. *et al.* Solvent effects in acid-catalyzed biomass conversion reactions. *Angewandte Chemie - International Edition* **53**, 11872–11875 (2014).

125. Gallo, J. M. R., Alonso, D. M., Mellmer, M. A. & Dumesic, J. A. Production and upgrading of 5-hydroxymethylfurfural using heterogeneous catalysts and biomass-derived solvents. *Green Chemistry* **15**, 85–90 (2013).

126. Mushrif, S. H., Caratzoulas, S. & Vlachos, D. G. Understanding solvent effects in the selective conversion of fructose to 5-hydroxymethyl-furfural: A molecular dynamics investigation. *Physical Chemistry Chemical Physics* **14**, 2637–2644 (2012).

127. Chew, A. K., Jiang, S., Zhang, W., Zavala, V. M. & Van Lehn, R. C. Fast predictions of liquid-phase acid-catalyzed reaction rates using molecular dynamics simulations and convolutional neural networks. *Chem Sci* **11**, 12464–12476 (2020).

128. Pronk, S. *et al.* GROMACS 4.5: A high-throughput and highly parallel open-source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).

129. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* **118**, 11225–11236 (1996).

130. Pomata, M. H. H., Sonoda, M. T., Skaf, M. S. & Dolores Elola, M. Anomalous dynamics of hydration water in carbohydrate solutions. *Journal of Physical Chemistry B* **113**, 12999–13006 (2009).

131. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J Chem Phys* **81**, 511–519 (1984).

132. Humphrey, W., Dalke, A. & Schulten, K. Sartorius products. *J Mol Graph* **14**, 33–38 (1996).

133. Laio, A. & Gervasio, F. L. Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry, and material science. *Reports on Progress in Physics* **71**, (2008).

134. Pavan, G. M., Barducci, A., Albertazzi, L. & Parrinello, M. Combining metadynamics simulation and experiments to characterize dendrimers in solution. *Soft Matter* **9**, 2593–2597 (2013).

135. Hoang, T. M. C., Lefferts, L. & Seshan, K. Valorization of humin-based byproducts from biomass processing - A route to sustainable hydrogen. *ChemSusChem* **6**, 1651–1658 (2013).

136. Dee, S. J. & Bell, A. T. A study of the acid-catalyzed hydrolysis of cellulose dissolved in ionic liquids and the factors influencing the dehydration of glucose and the formation of humins. *ChemSusChem* **4**, 1166–1173 (2011).

137. Mellmer, M. A. *et al.* Solvent-enabled control of reactivity for liquid-phase reactions of biomass-derived compounds. *Nat Catal* **1**, 199–207 (2018).

138. Mellmer, M. A. *et al.* Solvent-enabled control of reactivity for liquid-phase reactions of biomass-derived compounds. *Nat Catal* **1**, 199–207 (2018).

139. Constable, D. J. C., Jimenez-Gonzalez, C. & Henderson, R. K. Perspective on solvent use in the pharmaceutical industry. *Org Process Res Dev* **11**, 133–137 (2007).

140. Gallezot, P. Conversion of biomass to selected chemical products. *Chem Soc Rev* **41**, 1538–1558 (2012).

141. Lenihan, P. *et al.* Dilute acid hydrolysis of lignocellulosic biomass. *Chemical Engineering Journal* **156**, 395–403 (2010).

142. Song, B., Yu, Y. & Wu, H. Solvent effect of gamma-valerolactone (GVL) on cellulose and biomass hydrolysis in hot-compressed GVL/water mixtures. *Fuel* **232**, 317–322 (2018).

143. Velasco Calderón, J. C., Jiang, S. & Mushrif, S. H. Understanding the Effect of Solvent Environment on the Interaction of Hydronium Ion with Biomass Derived Species: A Molecular Dynamics and Metadynamics Investigation. *ChemPhysChem* **22**, 2222–2230 (2021).

144. Sabin, J. R., Brändas, E. J., Kawai, J., Kim, Y. S. & Adachi, H. *Advances in Quantum Chemistry: DV-Xá for Industrial-Academic Cooperation*. (Elsevier Science, 2011).

145. Chew, A. K. *et al.* Effect of Mixed-Solvent Environments on the Selectivity of Acid-Catalyzed Dehydration Reactions. *ACS Catal* **10**, 1679–1691 (2020).

146. Beuzekom, S. Van & Waal, J. C. Van Der. Effect of the Solvent in Enhancing the Selectivity to Furan Derivatives in the Catalytic Hydrogenation of Furfural. *ACS Sustainable Chem. Eng.* **6**, 16235–16247 (2018).

147. Hazra, M. K. & Bagchi, B. Non-equilibrium solvation dynamics in water-DMSO binary mixture: Composition dependence of non-linear relaxation. *Journal of Chemical Physics* **149**, (2018).

148. Tran, B., Milner, S. T. & Janik, M. J. Kinetics of Acid-Catalyzed Dehydration of Alcohols in Mixed Solvent Modeled by Multiscale DFT/MD. *ACS Catal* **12**, 13193–13206 (2022).

149. Mellmer, M. A. *et al.* Effects of chloride ions in acid-catalyzed biomass dehydration reactions in polar aprotic solvents. *Nat Commun* **10**, 1–10 (2019).

150. Caratzoulas, S. & Vlachos, D. G. Converting fructose to 5-hydroxymethylfurfural: A quantum mechanics/molecular mechanics study of the mechanism and energetics. *Carbohydr Res* **346**, 664–672 (2011).

151. Nikbin, N., Caratzoulas, S. & Vlachos, D. G. A First Principles-Based Microkinetic Model for the Conversion of Fructose to 5-Hydroxymethylfurfural. *ChemCatChem* **4**, 504–511 (2012).

152. Mushrif, S. H., Varghese, J. J. & Krishnamurthy, C. B. Solvation dynamics and energetics of intramolecular hydride transfer reactions in biomass conversion. *Phys Chem Chem Phys* **17**, 4961–4969 (2015).

153. Zhang, Q. & Asthagiri, A. Solvation effects on DFT predictions of ORR activity on metal surfaces. *Catal Today* **323**, 35–43 (2019).

154. Morais, R. F. De, Kerber, T., Calle-vallejo, F. & Sautet, P. Capturing Solvation Effects at a Liquid / Nanoparticle Interface by Ab Initio Molecular Dynamics: Pt 201 Immersed in Water. *small* **12**, 5312–5319 (2016).

155. Stabilization, S. Theoretical Studies of Enzymic Reactions: *J. Mol. Biol.* **103**, 227–249 (1976).

156. Padmanathan, A. M. D. & Mushrif, S. H. Pyrolytic activation of cellulose: energetics and condensed phase effects. *React Chem Eng* **7**, 1136–1149 (2022).

157. Varghese, J. J. & Mushrif, S. H. Origins of complex solvent effects on chemical reactivity and computational tools to investigate them: A review. *React Chem Eng* **4**, 165–206 (2019).

158. Schwinn, K., Ferre, N. & Huix-Rotllant, M. E ffi cient Analytic Second Derivative of Electrostatic Embedding QM / MM Energy: Normal Mode Analysis of Plant Cryptochrome. *J. Chem. Theory Comput* **16**, 3816–3824 (2020).

159. Arora, J. S., Chew, J. W. & Mushrif, S. H. Influence of Alkali and Alkaline-Earth Metals on the Cleavage of Glycosidic Bond in Biomass Pyrolysis: A DFT Study Using Cellobiose as a Model Compound. *Journal of Physical Chemistry A* **122**, 7646–7658 (2018).

160. Pascual-ahuir, J. L., Silla, E. & Tuñon, I. GEPOL: An improved description of molecular surfaces. III. A new algorithm for the computation of a solvent-excluding surface. *J Comput Chem* **15**, 1127–1138 (1994).

161. Padmanathan, A. M. D. & Mushrif, S. H. Pyrolytic activation of cellulose: energetics and condensed phase effects. *React Chem Eng* **7**, 1136–1149 (2022).

162. Bordwell, F. G., Cheng, J. P. & Harrelson, J. A. Homolytic Bond Dissociation Energies in Solution from Equilibrium Acidity and Electrochemical Data. *J Am Chem Soc* **110**, 1229–1231 (1988).

163. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926–935 (1983).

164. Kollman, P. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem Rev* **93**, 2395–2417 (1993).

165. Qian, X., Nimlos, M. R., Davis, M., Johnson, D. K. & Himmel, M. E. Ab initio molecular dynamics simulations of β-D-glucose and β-D-xylose degradation mechanisms in acidic aqueous solution. *Carbohydr Res* **340**, 2319–2327 (2005).

166. Prodans, E. & Kohn, W. Nearsightedness of electronic matter. *Proc Natl Acad Sci U S A* **102**, 11635–11638 (2005).

167. Hartwigsen, C., Goedecker, S. & Hutter, J. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **58**, 3641–3662 (1998).

168. Laio, A. & Gervasio, F. L. Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry, and material science. *Reports on Progress in Physics* **71**, (2008).

169. Chakraborty, P., Agrawal, K. & Kishore, N. Kinetic Modeling of Conversion of Levulinic Acid to Valeric Acid in Supercritical Water Using the Density Functional Theory Framework. *Ind. Eng. Chem. Res* **59**, 18683−18692 (2020).

170. Carolina, N. & Hill, C. Electronegativity Equalization Method for the Calculation of Atomic Charges in Molecules. *J. Am. Chem. Soc.* **108**, 4315–4320 (1986).

171. Lamoureux, G. A simple polarizable model of water based on classical Drude oscillators. **5197**, 5185–5197 (2017).

172. Davis, M. E. & Mccammon, J. A. Electrostatics in Biomolecular Structure and Dynamics. *Chem. Rev* **90**, 509–521 (1990).

173. Tucker, M. H. *et al.* Sustainable solvent systems for use in tandem carbohydrate dehydration hydrogenation. *ACS Sustain Chem Eng* **1**, 554–560 (2013).

174. Motagamwala, A. H., Huang, K., Maravelias, C. T. & Dumesic, J. A. Solvent system for effective near-term production of hydroxymethylfurfural (HMF) with potential for long-term process improvement. *Energy Environ Sci* **12**, 2212–2222 (2019).

175. Qi, X., Watanabe, M., Aida, T. M. & Smith, R. L. Catalytic dehydration of fructose into 5-hydroxymethylfurfural by ion-exchange resin in mixed-aqueous system by microwave heating. *Green Chemistry* **10**, 799–80 (2008).

176. Velasco Calderón, J. C., Jiang, S. & Mushrif, S. H. Understanding the Effect of Solvent Environment on the Interaction of Hydronium Ion with Biomass Derived Species: A Molecular Dynamics and Metadynamics Investigation. *ChemPhysChem* **22**, 2222–2230 (2021).

177. Kubař, T. & Elstner, M. Solvent reorganization energy of hole transfer in DNA. *Journal of Physical Chemistry B* **113**, 5653–5656 (2009).

178.   Mushrif, S. H., Varghese, J. J. & Krishnamurthy, C. B. Solvation dynamics and energetics of intramolecular hydride transfer reactions in biomass conversion. *Phys Chem Chem Phys* **17**, 4961–4969 (2015).

179.   Koverga, V. *et al.* Local Structure of DMF-Water Mixtures, as Seen from Computer Simulations and Voronoi Analysis. *Journal of Physical Chemistry B* **126**, 6964–6978 (2022).

180.   Marcel Schlaf, Z. C. Z. *Reaction Pathways and Mechanisms in Thermocatalytic Biomass Conversion II*. (Springer, 2016).

181.   Frank Jensen. *Introduction to Computational Chemistry*. (Wiley, 2007).

182.   Samir H. Mushrif. Molecular Modeling Methods. 194–275.

183.   Li, Y., Yang, F. H. & Yang, R. T. Kinetics and mechanistic model for hydrogen spillover on bridged metal-organic frameworks. *Journal of Physical Chemistry C* **111**, 3405–3411 (2007).

184.   Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J Chem Phys* **81**, 511–519 (1984).

185.   Groenhof, G. *Biomolecular Simulations: Methods and Protocols*. *Biomolecular Simulations: Methods and Protocols* vol. 924 (2013).

186.   Yu, H. & Gunsteren, W. F. Van. Accounting for polarization in molecular simulation. **172**, 69–85 (2005).

187.   Jensen, W. *Introduction to Computational Chemistry*. *Direct Nerve Stimulation for Induction of Sensation and Treatment of Phantom Limb Pain* (Wiley, 2019).

188.   Curtiss, L. A., Jones, C., Trucks, G. W., Raghavachari, K. & Pople, J. A. Gaussian-1 theory of molecular energies for second-row compounds. *J Chem Phys* **93**, 2537–2545 (1990).

189.   Curtiss, L. A., Raghavachari, K., Trucks, G. W. & Pople, J. A. Gaussian-2 theory for molecular energies of first- and second-row compounds. *J Chem Phys* **94**, 7221–7230 (1991).

190.   Sousa, S. F., Fernandes, P. A. & Ramos, M. J. General performance of density functionals. *Journal of Physical Chemistry A* **111**, 10439–10452 (2007).

191.   Cohen, A. J., Mori-Sánchez, P. & Yang, W. Challenges for density functional theory. *Chem Rev* **112**, 289–320 (2012).

192.   Krishnan, R., Schlegel, H. B. & Pople, J. A. Derivative studies in configuration-interaction theory. *J Chem Phys* **72**, 4654–4655 (1979).

193.   Car, R. & Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **55**, 2471–2474 (1985).

194.   Pastore, G., Smargiassi, E. & Buda, F. Theory of ab initio molecular-dynamics calculations. *Phys. Rev. A* **44**, 6334–6347 (1991).

195.   Valleau, J. P. & J.M., T. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J Comput Phys* **23**, 187–199 (1977).

196.   Laio, A. & Gervasio, F. L. Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry, and material science. *Reports on Progress in Physics* **71**, (2008).

197.   Iannuzzi, M., Laio, A. & Parrinello, M. Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics. *Phys. Rev. Lett.* **90**, 238302 (2003).

198.   Ensing, B., Laio, A., Parrinello, M. & Klein, M. L. A recipe for the computation of the free energy barrier and the lowest free energy path of concerted reactions. *Journal of Physical Chemistry B* **109**, 6676–6687 (2005).

Appendices

A.1 Computational methods

A.1.1 Classical Molecular Dynamics

In forcefield calculations, the electronic motion is omitted, and only the nuclei degrees of freedom are considered following the Born-Oppenheimer approximation suggesting that the differentiation between the nuclei and electrons is valid since both occur at different timescales. Thus, the energy depends only on the nuclear position. In this section, nuclear or atomic coordinates will be used indifferently. In molecular mechanics or forcefield methods, the atom is considered the fundamental unit in the system. Molecules are composed of atoms; hence using this modelling approach, we can understand the conformation of a molecule by thinking of it as a "balls and springs" model, as shown in Figure A.1. [180]

.
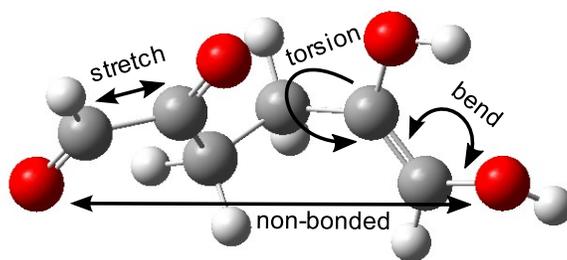
Figure A.1 Molecular bonded and non-bonded interactions (Adapted from Frank Jensen181) © 2007 John Wiley & Sons, Inc.

By assuming a multimolecular system, the potential energy can be distinguished into two main contributions: bonded and nonbonded interactions. Bonded interactions can be, in turn, subdivided into bonds (stretch), angles (bend) and dihedrals (torsion) as shown in equation 2. In contrast, nonbonded interactions account for Van der Waals and Coulomb electrostatic interactions (equation 3). In addition, more sophisticated forcefield implementations account for atomic polarizabilities or more complex coupling terms, such as cross-coupling between angles and bonds; however, those last considerations are not used in this work. [182]

$$E_{MM} = E_{bonded} + E_{non-bonded}$$

1

$$E_{bonded} = E_{stretch} + E_{bend} + E_{torsion}$$

2

$$E_{non-bonded} = E_{vdW} + E_{electrostatic}$$

3

Molecular stretching and bending terms can be written using harmonic oscillator form (considering balls and spring model) as shown in equation 4 & 5. However, torsion energy must ensure periodicity because, after rotating 360 degrees, it comes to the initial value as indicated in equation 6. [183]

$$E_{stretch}^{12} = \frac{1}{2!}\frac{d^2E}{dl^2}(l^{12} - l^0)^2 = K_{stretch}(l^{12} - l^0)^2$$

4

where $K_{stretch}$ is the force constant.

$$E_{bend}^{123} = K_{bend}(\theta^{123} - \theta^0)^2$$

5

$$E_{torsion}^{1234} = K_{torsion}(1 - cos(\eta\omega))$$

6

where $K_{torsion}$ is the constant, ω is angle of rotation and η for periodicity. Van der Waals energy is expressed using Lennard-Jones potential accounting for repulsion and attraction between each pair of atoms (equation 7), while the well-known Coulomb potential gives the electrostatic interactions (equation 8).

$$E_{vdW}^{12} = 4E_{min}^{LJ}\left[(\frac{R^0}{R^{12}})^{12} - (\frac{R^0}{R^{12}})^6\right]$$

7

where E_{min}^{LJ} is the depth of minimum potential and R^0 distance at zero potential

$$E_{electrostatic}^{12} = \frac{Q^1Q^2}{\varepsilon_{dielec}R^{12}}$$

8

where Q^1 and Q^2 are atomic charges and ε_{dielec} the dielectric constant

The validation of force fields depends on a detailed comparison with experimental data. Force fields are parameterized using a particular set of properties, but they should be able to reproduce or predict quantities that can accurately be measured but were not used during the parameterization All previous expressions depend on the atoms' coordinates (spatial location). Hence, the system's potential energy is calculated as a function of the atomic coordinates. Since the energy is a function of the position of the atoms, the minimum energy corresponds to the most stable atomic/molecular configuration. Physically, minimizing E_{MM} involves rearranging the atoms to reach the state of minimum energy. The dynamics of the atoms in molecular mechanics obey classical Newton's laws of motion. For a multiatomic system $\vec{R}(R_1, R_2, \ldots, R_N)$, the potential energy at a specific configuration $\vec{R}$ is $E(\vec{R})$, and the differential form of that is shown in equation 9.

$$-\frac{dE}{d\vec{R}} = m\frac{d^2\vec{R}}{dt^2}$$

9

Newton's second law of motion conserves the total energy including the kinetic and potential energy of the system. An iterative process is performed to obtain the minimum energy as a function of the change of atomic coordinates with time. The velocity verlet molecular dynamics algorithm (shown in Figure A.2) is a

numerical method that, given an initial set of atomic coordinates, minimizes E $(\vec{R})$ by reconfiguring the positions of the atoms by updating their associated velocities.
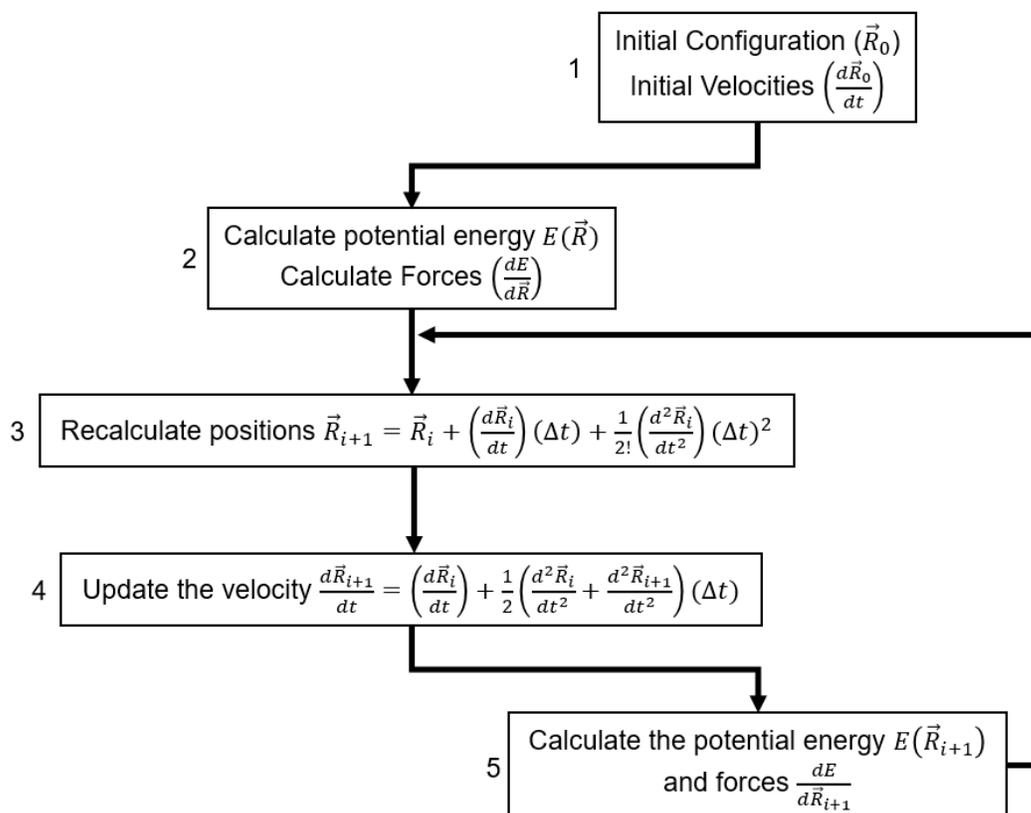

Figure A.2 Verlet algorithm diagram explaining classical molecular dynamics.

Recalculating the atomic positions at time step Δt is achieved using Taylor series expansion, as shown in step 3 in Figure A.2. Additionally, different ensembles can be performed for molecular dynamics, such as microcanonical ensemble NVE, where the number of particles N, volume V and total energy E are kept constant throughout the simulation. Other options of ensembles, such as canonical NVT or isobaric-isothermal NPT, are selected depending on the type or properties that need to be evaluated. A barostat is used for ensembles where the pressure P needs to be constant. Similarly, for isothermal ensembles such as NVT, thermostats are implemented (i.e., Nose-Hoover or Andersen) [184].

Forcefield methods have excellent performance at a moderate computational cost to provide an understanding of molecular motion at even mesoscale model systems. There is a wide variety of applications where this methods have contributed significantly, such as biochemistry modelling proteins biopolymers or studying condensed phase system phenomena (physical interaction between the solvent molecules and the substrates).[185] On the other hand, in the catalysis space, there is an intrinsic need to model chemical reactions, which generate an electronic charge deformation (bond formed/broken or

altering the electronic orbitals within the atoms); hence, forcefield calculations fail to model a system in which chemical reactions occur. [186]

A.1.2 Overview of electronic structure calculations

The Schrödinger equation, one of the most critical equations in quantum mechanics, refers to the spatial and temporal evolution of the wave function of a particle in each potential. Since this is initially a partial differential equation, separating the wavefunction into spatial and temporal parts under Born-Oppenheimer approximation, the Schrödinger equation can be expressed as follows in equation 10.[187]

$$H_e \Psi_e(r_e, R_e) = E_e \Psi_e(r_e, R_e)$$

10

However, modelling poly electronic systems has limitations in solving this equation analytically. It is possible to approximate an accurate solution using valence bond theory and representing each wavefunction/orbital as Slater determinants. During a chemical reaction (bond formation), the individual atomic orbitals AOs are overlapped, forming molecular orbitals MOs. Each molecule has a set of molecular orbitals that can be mathematically described as combining the individual AOs The MOs are expressed as the linear combination of atomic orbitals (LCAO) as shown in equation 11, obeying Pauli's exclusion principle.

$$\chi = C_1 \varphi_1 + C_2 \varphi_2 + \ldots + C_n \varphi_n$$

11

where χ is the molecular orbital, φ_i is an atomic orbital C_i is the coefficient associated with the atomic orbital.

To ensure Pauli's exclusion principle by representing the total wave function for N electrons, then the molecular orbitals χ are expressed as Slater determinants (equation 12).

$$\Psi_{(1,2,3\ldots N)} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) & \ldots & \chi_N(1) \\ \chi_2(1) & \chi_2(1) & \ldots & \chi_N(2) \\ \chi_1(N) & \chi_2(N) & \ldots & \chi_N(N) \end{vmatrix} = |\chi_1(1) \quad \chi_2(2) \quad \chi_N(N)|$$

12

To define the AOs in polyelectronic systems, AOs are reconstructed in the form of Slater orbitals (SOT) or Gaussian orbitals (GTO) (STOs prefer over GTOs because of the more accurate description of the AOs). The functional form of Slater orbitals is shown in equation 13.

$$\varphi_{\zeta,n,l,m}(r, \theta, \phi) = N Y_{l,m}(\theta, \phi) r^{n-1} e^{-\zeta r}$$

13

where N is a normalization constant, ζ represent the effective charge of the nucleus and $Y_{l,m}$ are spherical harmonic functions.

Then, based on the variational principle, the wave function is determined calculating the set of coefficients associated with the AOs. The objective is to find the combination of coefficients those results in the minimum energy of the system (the minimum energy of the system is minimized as a function of the atomic coefficients).

A.1.3 Ab Initio Methods: Hartree-Fock and Post Hartree-Fock Approximations

The Hamiltonian associated for N electrons is given in equation 14.[180]

$$\mathbf{H} = \sum_i^N -\frac{1}{2}\nabla^2 - \sum_i^N \sum_\alpha^N \frac{Z_\alpha}{|\mathbf{r_i} - \mathbf{r_\alpha}|} + \sum_i^N \sum_j^N \frac{1}{|\mathbf{r_i} - \mathbf{r_j}|}$$

14

On the other hand, the energy of the system can be written as follow in equation 15.

$$\mathbf{E} = \langle \psi | \Sigma_i^N \mathbf{h_i} | \psi \rangle + \langle \psi | \Sigma_i^N \Sigma_j^N \mathbf{g_{ij}} | \psi \rangle$$

15

Considering $\mathbf{h_i} = \Sigma_i^N -\frac{1}{2}\nabla^2 - \Sigma_i^N \Sigma_\alpha \frac{Z_\alpha}{|\mathbf{r_i}-\mathbf{r_\alpha}|}$ and $\mathbf{g_{ij}} = \Sigma_i^N \Sigma_j^N \frac{1}{|\mathbf{r_i}-\mathbf{r_j}|}$

But taking considering again Pauli's exclusion principle, the wavefunction is written in equation 16 in terms of Slater's determinant.

$$E = \sum_I^N \langle \chi_1 | -\frac{\nabla_I^2}{2} | \chi_I \rangle + \sum_I^N \langle \chi_1 | -\frac{Z_\alpha}{r_{i\alpha}} | \chi_I \rangle + \sum_i^N \sum_j^N (J_{ij} - K_{ij})$$

16

J_{12} is the coulomb integral, it is define as $\mathbf{J_{12}} = \langle \chi_1(1)\chi_2(2) | \frac{1}{r_{12}} | \chi_1(1)\chi_2(2) \rangle$, while K_{12} is the exchange integral derived from the antisymmetric constraint as $\mathbf{K_{12}} = \langle \chi_1(1)\chi_2(2) | \frac{1}{r_{12}} | \chi_2(1)\chi_1(2) \rangle$. Since in the molecular system every orbital is doubly occupied, then the energy considering N=2n can be calculated as shown in equation 17.

$$\mathbf{E} = \sum_i^n 2\mathbf{h_{ii}} + \sum_{ij}^n (2\mathbf{J_{ij}} - \mathbf{K_{ij}})$$

17

Since the variational principle is used to guess the initial set of coefficients associated to the AOs to calculate the minimum energy for the system. Lagrange multipliers are introduced to perform the minimization of the energy E as a function of the coefficient associated with AOs $(\mathbf{c_1}, \mathbf{c_2}, \ldots, \mathbf{c_n})$. To obtain the minimum energy, the derivate of the Lagrangian function with respect to the coefficients must be zero. The calculation process is written as shown in equation 18.

$$\xi[\chi] = \sum_i^n 2h_{ii} + \sum_{ij}^n (2J_{ij} - K_{ij}) - \sum_i^n 2\lambda_{ij}[\langle\chi_i|\chi_j\rangle - \delta_{ij}]$$

18

The derivative of equation 18 is given in equation 19.

$$\rightarrow \frac{\delta\xi}{\delta\chi_i} = 0 \rightarrow \left[h_i + \sum_j^n (2J_{ij} - K_{ij})\right]\chi_i = \sum_j^n \lambda_{ij}\chi_j$$

19

Lagrange multipliers (λ_{ij}) are introduced under constraint where the MOs remain orthogonal and normalized. Given that the wavefunction is defined as a determinant, can be diagonalized to obtain energies (ε_i) leading to the Hartree-Fock equation 20.

$$\left[h_i + \sum_j^n (2J_j - K_j)\right]\chi_i = \mathbb{F}_i\chi_i = \varepsilon_i\chi_i$$

20

As a recall, MOs are defined as LCAO, as follows

$$\chi_i = \sum_v^M c_{iv}\varphi_v$$

21

Therefore, writing the Hartree-Fock equation explicitly in terms of AOs and their associated coefficients we get equation 22.

$$\mathbb{F}_i \sum_v^M c_{iv}\varphi_v = \varepsilon_i \sum_v^M c_{iv}\varphi_v$$

22

where M is the number of basis functions

Specifying the number of basis functions to the previous equation we obtain the following expression.

$$\mathbb{F}c = \mathbb{S}c\varepsilon$$

23

where $\mathbb{F}$ is the Fock Matrix and $\mathbb{S}$ is the overlap matrix. The Hartree Fock algorithm needs an initial set of AOs coefficients. The Hartree-Fock algorithm is an iterative process that converges when the coefficients entered to are the same as those emanating from the diagonalization. By increasing the number of basis sets, the accuracy of the electronic calculation increases; however, the electron-electron interaction is not captured during the Hartree-Fock calculations.[188] This is because the electrons' motion is not correlated in the Hartree-Fock procedure. [187]

Post Hartree-Fock methodologies include additional Slater determinants as excited state orbitals to the initial "ground states" orbitals as written in equation 24 to calculate the electrons' motion at certain extend.[189]

$$\Psi_{correlation} = c^{-0}\Psi^{HF} + c^{-1}\Psi^1 + c^{-2}\Psi^2 + \ldots + c^{-n}\Psi^n$$

24

A.1.4 Density Functional Theory (DFT) Calculations

DFT is an alternative methodology for electronic structure calculations that enhances the computational performance due to more efficient calculations compared to ab initio Hartree-Fock and Post Hartree-Fock methods. DFT approach considers the electron density as the variable targeted, reducing the degrees of freedom of the many-electron wave functions used in HF and Post HF methods.[180] Hohenberg-Kohn theorem suggests that the ground-state wave function of the entire system and the energy change are functional of the electron density as shown in equation 6.25.[190]

$$\Psi_0 = \Psi[\rho_0]$$

25

The total energy of the system as a function of the electron density can be written as shown in equation 6.26 and 6.27.

$$E_0 = E[\rho_0] = \langle\Psi[\rho_0]|\hat{H}|\Psi[\rho_0]\rangle$$

26

$$E[\rho] = T[\rho] + V_{Ne}[\rho] + U_{ee}[\rho] = \int \rho(r)V_{Ne}(r)dr + F_{HK}[\rho]$$

27

$V_{Ne}[\rho]$ corresponds to electron-nuclei interaction, and $F_{HK}[\rho]$ is the universal functional containing all the unknown terms and, in turn is defined as shown in equation 28.

$$F_{HK}[\rho] = T[\rho] + J[\rho] + U_{ncl}$$

28

The universal functional is described based on the kinetic energy of electrons $(T[\rho])$, classical coulombic electron-electron interaction $(J[\rho])$ and all the non-classical interactions, exchange, and Coulombic correlations (U_{ncl}). While $V_{Ne}[\rho]$ is relatively easy to obtain, the challenge of this theorem is to calculate $F_{HK}[\rho]$. To approximate a solution, it is proposed that each electron experiences an average local potential produced by all the electrons (electron density), which in turn, is constructed by a set of Kohn-Sham spin orbitals (K-S spin orbitals $\emptyset_i$). This K-S spin orbitals are analogous to the Slater determinants used in HF theory. Using a non-interacting electron system as a blank, $T[\rho]$ can be written as follows

93

$$T_s = -\frac{1}{2}\sum_i^N \langle \phi_i | \nabla^2 | \phi_i \rangle$$

29

However, the energy using the previous form for a non-interacting electron-electron deviates substantially from the real case. Defining the electron density for the real system as $\rho(\mathbf{r}) = \sum_i^N |\phi_i|^2 = \rho_s(\mathbf{r})$, then the energy of the real system is described as follows:

$$E[\rho] = T_s[\rho] + V_{Ne}[\rho] + J[\rho] + E_{XC}[\rho]$$

30

Using equation 30, all the unknown terms are grouped into the functional $E_{XC}[\rho]$, known as exchange correlation energy. $E_{XC}[\rho]$ contains the kinetic energy difference between the real and the blank system (used as a correction term). The aim of correctly expressing the form of $E_{XC}[\rho]$ is first (i) to find the set of K-S spin orbitals to determine the ground-state density and (ii) to minimize the energy written in equation 30. The single electron K-S is shown in equation 31.

$$\left(-\frac{1}{2}\nabla^2 - \sum_A^M \frac{Z_A}{r_{1A}} + \int \frac{\rho_{(r_2)}}{r_{12}} dr_2 + V_{XC}(r_1) \right) \emptyset_i = \varepsilon_i \emptyset_i$$

31

The exchange correlation potential is for definition the exchange-correlation energy as a function of the electron density $V_{XC}(\mathbf{r}) = \frac{\delta E_{XC}}{\delta \rho}$. The solution would yield the ground state of the system, however, getting the exact solution is not possible and the accuracy of the method depends on the efficiency of $V_{XC}(\mathbf{r})$. There are two main methodologies to approximate $E_{XC}[\rho]$ (i) Local Density Approximation (LDA) and (ii) General Gradient Approximation (GGA).[191] LDA considers a homogeneous electron gas which means constant charge density. LDA is suitable for systems where the electron density barely changes with time such as bulk metals or ionic crystals but to predict long range interactions systems, usually deviate from experiments. $E_{XC}[\rho]$ under LDA scheme is written as $E_{XC}^{LDA}[\rho] = \int \rho(\mathbf{r}) E_{XC}[\rho(\mathbf{r})]d\mathbf{r}$. GGA approach accounts for nonhomogeneous electron density, in contrast to LDA which considers a local region of the electron density, GGA consider the first derivative of the electron density (how the electron density changes as a function of the position). $E_{XC}[\rho]$ using GGA can be written in the form of $E_{XC}^{GGA}[\rho] = \int \rho(\mathbf{r}) E_{XC}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})]d\mathbf{r}.$

Analogous to the wavefunction description in HF methods, K-S MOs are constructed as a linear combination of a set of basis functions centered on different atoms, better known as basis sets. As mentioned, AOs are represented using STOs; however, to enhance computing performance, STOs are described as linear combinations of GTOs, known as CGTOs. Some examples of CGTO are double, triple

or quadruple-split. The quantifier added corresponds to the number of basis sets employed to describe an AO. The basis sets can be enabled with polarization and diffusion functions. Polarization accounts for the distortion of AOs from the original symmetry, while diffusion is added to precisely describe the AO far away from the nucleus, i.e., valence AOs.

On the other hand, basis sets combined with pseudopotentials extend the description of period systems. Unlike the CGTO, plane-wave basis sets are not associated with atoms but are constructed by a finite number of plane-wave functions describing the valence electrons explicitly. According to Bloch's theorem, to model periodic systems is not necessary to account for an infinite number of electrons but to use a finite number of electrons in a periodically repeated unit cell. The accuracy of this approach lies in the cutoff energy of those plane wave basis sets to converge the system's total energy.

A.1.5 Car-Parinello Molecular Dynamics

Combining classical molecular dynamics using force field and Ab Initio molecular dynamics, it is possible to model complex chemical systems where the electronic arrangement and bonds configuration changes because of the chemical reaction (classical molecular dynamics drawback), but at the same time, the accessible timescale is significantly larger than for Ab Initio methods. The combination of the two approaches is known as Born-Oppenheimer molecular dynamics, where the electronic structure calculations are performed after every molecular dynamic timestep.[192] Here, the electronic motion is calculated using quantum mechanics, but the nuclear motion is treated classically. The electronic structure calculations can be performed using Hartree-Fock, Post Hartree-Fock, or Kohn Sham DFT approaches. However, optimizing the wave function or the electronic density after every molecular dynamic step is still computationally intense. Car-Parinello molecular dynamics differentiate from Born Oppenheimer methods because the wavefunction or density is optimized throughout the molecular dynamics' calculations avoiding the repetitive and intensive electronic optimization.[193] The rationale behind the CPMD methodology is propagating the optimized electron density using classical formalism. The classical propagation of the wavefunction is achieved by assigning a fictitious mass to the electronic wavefunction. Therefore, the total energy using Lagrangian is written as follows in equation 32.

$$\xi_{CP} = \sum_i m_i \dot{R}_i + \sum_j \frac{1}{2} \mu_i \langle \dot{\chi}_j | \dot{\chi}_j \rangle - \langle \Psi | H | \Psi \rangle + \text{constraint}$$

32

where χ_j is MO of the electron j^{th} electron, μ is the fictitious mass associated to the electron motion and the constraints can be some external or internal constraints like orthonormality. The kinetic energy associated with the wave function in the Lagrangian formulation is fictitious and not linked to the actual quantum kinetic energy. The fictitious mass μ associated with the kinetic energy of the wavefunction is selected to aim at a very low fictitious temperature.[194]

A.1.6 Metadynamics

Even though significant computational enhancement is obtained using ab initio molecular dynamics, some reaction systems occur at much larger timescales and thus become difficult to access. In some cases, the system is stuck in a local minimum requiring longer computational time to overcome the barrier associated with the reaction to reach the "global" minima. Therefore, it is required to accelerate the dynamics for the reaction to occur faster.[195] Metadynamics implementations accelerated the events in an accessible computational simulation time. This technique can be implemented in classical or ab initio molecular dynamics.[196] The metadynamics approach consists of adding potential to fill up the energy surface. The potential drops are tracked to reconstruct the energy surface to quantify the energy landscapes and barriers. In this way, the system crosses the energy barrier associated with the reaction to reach the global minima by adding the potentials seen in Figure A.3.
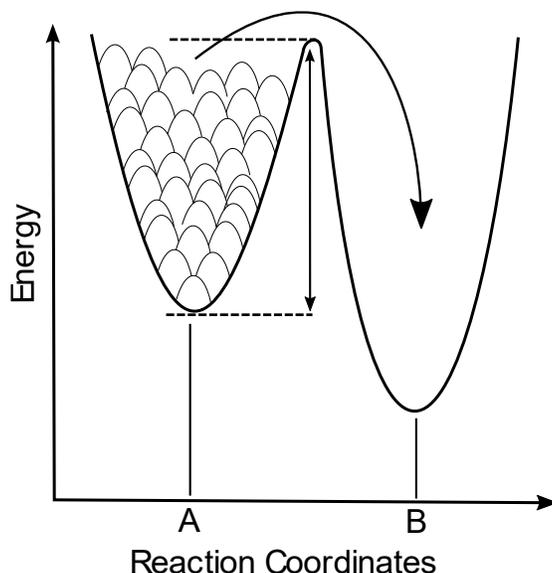
Figure A.3 Wells filling by the addition of the potential on the free energy surface during the metadynamics simulation

The free energy space is explored as a function of selected collective variables. These collective variables CV are chosen according to the reaction or phenomena to analyze; some examples of CVs are bond distances, coordination numbers, or dihedral angles. ς is taken as the vector associated with the CVs, the CPMD Lagrange is extended as follows in equation 33.[197]

$$\xi_{MTD} = \xi_{CP} + \sum_{CV} m_{CV}\dot{\varsigma}_{CV} - \frac{1}{2}\sum_{CV} k_{CV}\left[\varsigma_{CV}(\vec{R}_{CV}) - \varsigma_{CV}\right]^2 + \upsilon_{CV}(t,\varsigma)$$

33

ξ_{CP} is the lagrangian associated to the CPMD calculation followed by the kinetic energy of the collective variables $\sum_{CV} m_{CV}\dot{\varsigma}_{CV}$ and the harmonic restraining potential $\frac{1}{2}\sum_{CV} k_{CV}\left[\varsigma_{CV}(\vec{R}_{CV}) - \varsigma_{CV}\right]^2$. The last term corresponds to energy associated to the potential dropped in the energy space defined by the collective

variables ς at a certain time interval. Additionally, the potential used to fill the energy well is given in equation 34.[196]

$$\upsilon_{CV}(t, \varsigma) = \sum_{t_i < t} H_{MTD} \exp\left[-\frac{(\upsilon - \upsilon^i)^2}{2(\Delta w^{\mathcal{I}})}\right] \exp\left[-\frac{\{(\varsigma - \varsigma^i)(\varsigma^{i+1} - \varsigma^i)\}^2}{2\Delta w^{\mathcal{I}4}}\right]$$

34

where $\mathbf{H_{MTD}}$ is the height and $\Delta\mathbf{w}^{\mathcal{I}}$ is the width of the potential in a form of a Gaussian. The height and width of the Gaussian need to be specified in such a way that the free energy surface is faithfully reproduced. Also, it can be observed that the expression for the potential added force is the potential to be added as a function of the trajectory; hence the potentials are added close to each other to map accurately the energy surface in the space of interest. [198]

Well-tempered metadynamics is an improvement upon standard metadynamics designed to address its convergence issues, where a bias potential is still added to the simulation, but it is modulated by a boosting factor that decreases over time. The boosting factor is initially designed to be close to 1, allowing for rapid exploration of the free energy landscape. As the simulation progresses, the boosting factor decreases, effectively tempering the bias potential. The tempering helps achieve better convergence by reducing the bias at later simulation stages, allowing the system to explore deeper energy minima. In summary, while both metadynamics and well-tempered metadynamics aim to enhance sampling in molecular dynamics simulations, well-tempered metadynamics introduces a self-adjusting bias potential that reduces over time. [133, 134]

A.2 Appendix to Chapter 3

A.2.1 Atomic charges and simulation parameters.

Table A.1 Atomic charges and simulation parameters implemented in the present work for water, DMSO, 5-HMF and fructose. Blue balls indicate carbon atoms, red balls indicate oxygen atoms, white balls indicate hydrogen atoms and yellow ball indicates sulfur atom.
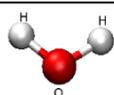
Molecule		Charges (in parentheses) and parameters, e/kJ mol-1, σ/Å
water		O(-0.834)-0.31,6.36 H(0.417)-0.00,0.00
hydronium ion		O(-0.626)-0.31,6.36 H(0.542)-0.00,0.00
DMSO		S(0.282)-0.35,1.65 O(-0.515)-0.29,1.17 C(-0.406)-0.38,6.69 H(0.174)-0.25,2.09
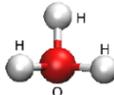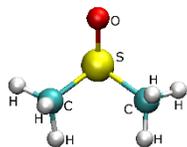
5-HMF
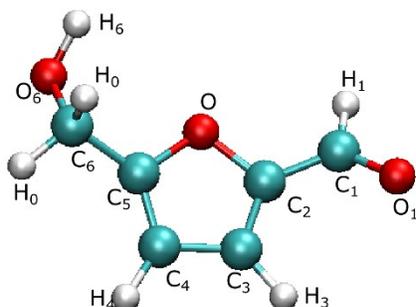
$O_1(-0.504)-0.29,8.78$
$C_1(0.412)-0.37,4.39$
$H_1(0.022)-0.24,6.27$
$C_2(0.020)-0.35,3.17$
$O(-0.210)-0.29,5.85$
$C_5(0.210)-0.35,3.17$
$C_4(0.303)-0.35,3.17$
$H_4(0.196)-0.24,1.25$
$C_3(-0.086)-0.35,3.17$
$H_3(0.165)-0.24,1.25$
$C_6(0.096)-0.39,4.93$
$H_0(0.079)-0.25,1.25$
$O_6(-0.601)-0.31,7.11$
$H_6(0.416)-0.00,0.00$

fructose
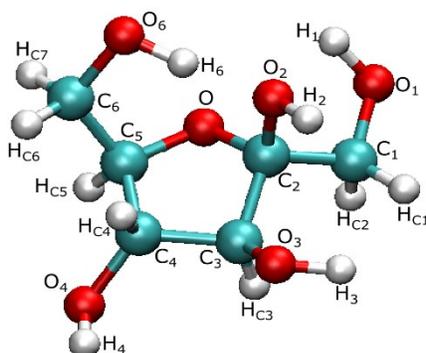
$H_6(0.452)-0.00,0.00$
$O_6(-0.699)-0.31,7.11$
$C_6(0.159)-0.35,2.76$
$H_{C6}(0.041)-0.25,1.25$
$H_{C7}(0.019)-0.25,1.25$
$C_5(0.325)-0.35,2.76$
$H_{C5}(0.045)-0.25,1.25$
$C_4(0.040)-0.35,2.76$
$H_{C4}(0.140)-0.25,1.25$
$O_4(-0.619)-0.31,7.11$
$H_4(0.406)-0.00,0.00$
$C_3(0.268)-0.35,2.76$
$H_{C3}(0.085)-0.25,1.25$
$O_3(-0.636)-0.31,7.11$
$H_3(0.420)-0.00,0.00$
$C_2(0.195)-0.35,2.76$
$O(-0.523)-0.30,7.11$
$O_2(-0.651)-0.31,7.11$
$H_2(0.479)-0.00,0.00$
$C_1(0.185)-0.35,2.76$
$H_{C1}(0.018)-0.25,1.25$
$H_{C2}(0.097)-0.25,1.25$
$O_1(-0.650)-0.31,7.11$
$H_1(0.407)-0.00,0.00$

Table A.2 Number of molecules and simulation cell size for fructose 1 % wt. at different concentrations of DMSO. For all systems, 3 molecules of fructose, 4 hydronium ions and 4 Cl^-.

DMSO wt. %	No. water	No. DMSO	Cell size (length) nm
0	2100	0	4.10674
5	2027	25	4.13317
20	1700	100	4.11765
50	1200	300	4.30766
80	425	460	4.16689

Table A.3 Number of molecules and simulation cell size for HMF 1 % wt. at different concentrations of DMSO. For all systems, 3 molecules of fructose, 4 hydronium ions and 4 Cl^-.

DMSO wt. %	No. water	No. DMSO	Cell size (length) nm
0	2135	0	4.13136
5	2027	25	4.12828
20	1700	100	4.11950
50	1200	300	4.31870
80	425	460	4.15276

A.2.2 RDFS for fructose and 5-HMF systems.
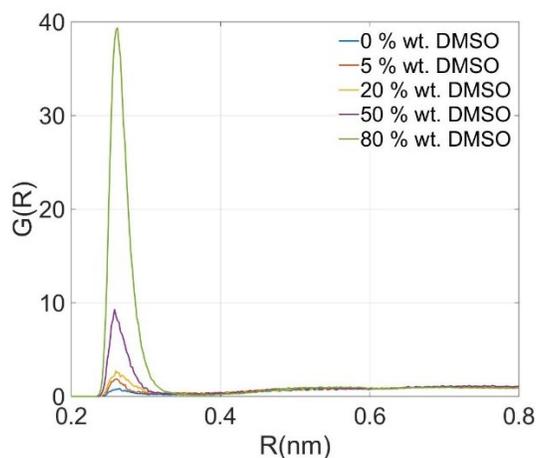
Figure A.4 Radial pair distributions for the fructose oxygen (O2)–hydronium ion oxygen (Ohyd) at all DMSO concentration studied. Fructose concentration of 10% wt.
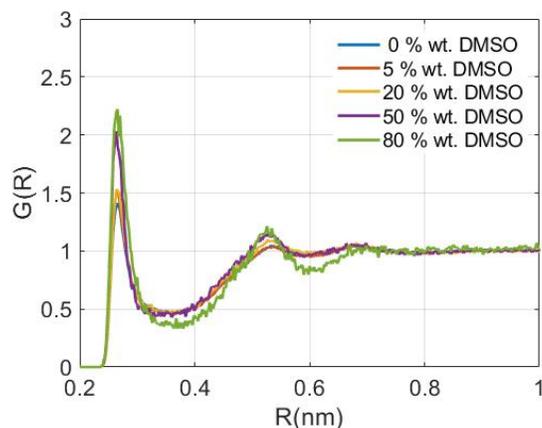
Figure A.5 Radial pair distributions for the fructose oxygen (O2)–water oxygen (OW) at all DMSO concentration studied. Fructose concentration of 1% wt.

99
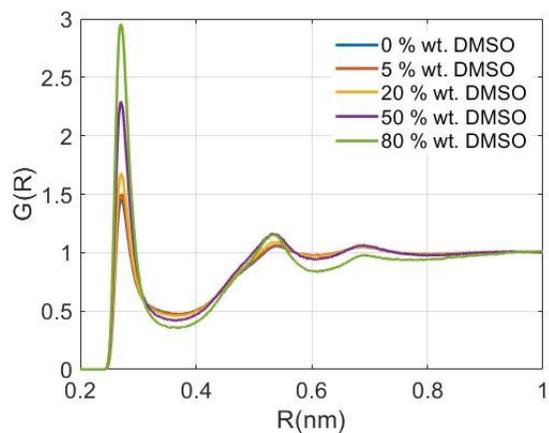
Figure A.6 Radial pair distributions for the fructose oxygen (O2)–water oxygen (OW) at all DMSO concentration studied. Fructose concentration of 10% wt.
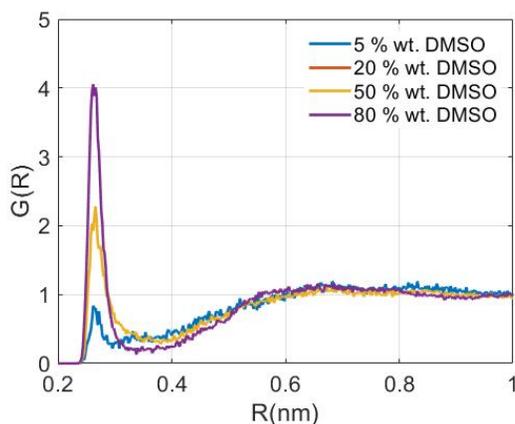
Figure A.7 Radial pair distributions for the oxygen (O2)–DMSO oxygen (OD) at all DMSO concentration studied. Fructose concentration of 1% wt.
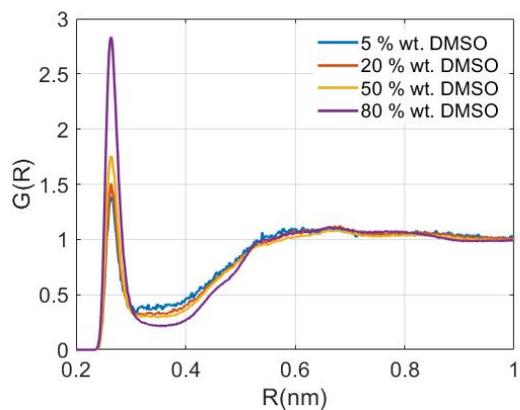
Figure A.8 Radial pair distributions for the fructose oxygen (O2)–DMSO oxygen (OD) at all DMSO concentration studied. Fructose concentration of 10% wt.
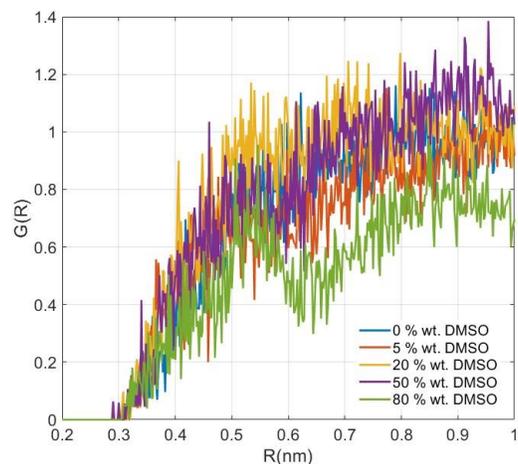
Figure A.9 Radial pair distributions for the 5-HMF C5 carbon atom–hydronium ion oxygen atom pairs, at all DMSO concentration studied. 5-HMF concentration of 1% wt.
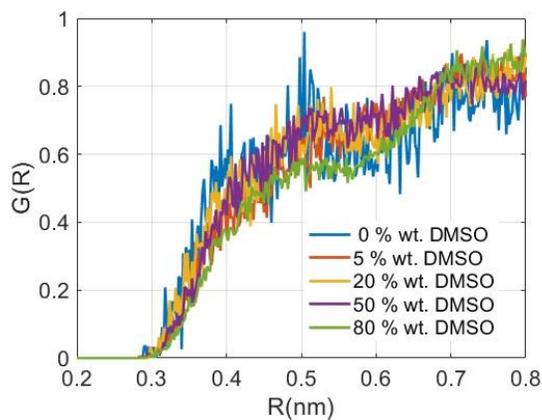
Figure A.10 Radial pair distributions for the 5-HMF C5–hydronium ion oxygen (Ohyd) at all DMSO concentration studied. 5-HMF concentration of 10% wt.
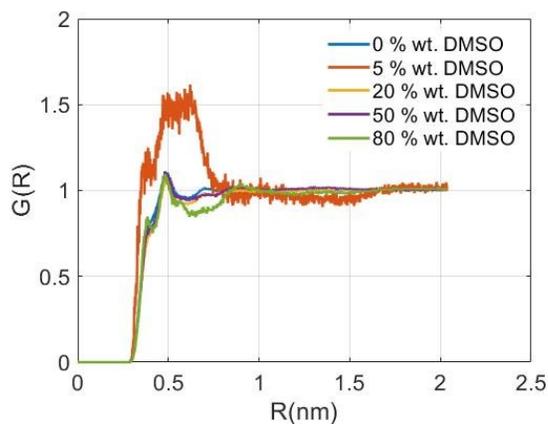
Figure A.0.11 Radial pair distributions for the 5-HMF C5– water oxygen (OW) at all DMSO concentration studied. HMF concentration of 1% wt.
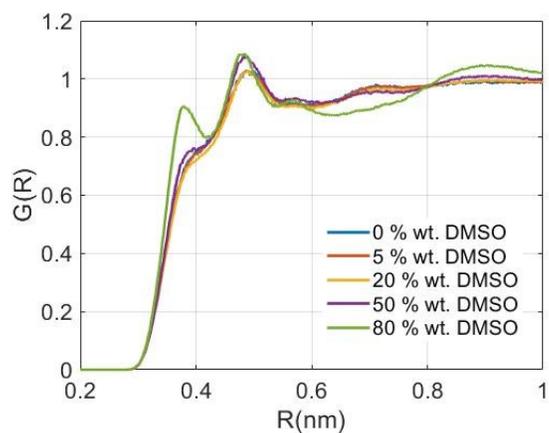
Figure A.12 Radial pair distributions for the 5-HMF C5 – water oxygen (OW) at all DMSO concentration studied. HMF concentration of 10% wt.
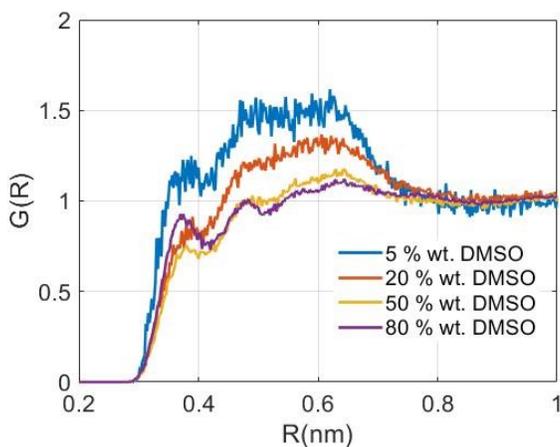
Figure A.13 Radial pair distributions for the 5-HMF C5– DMSO oxygen (OD) at all DMSO concentration studied. 5-HMF concentration of 1% wt.
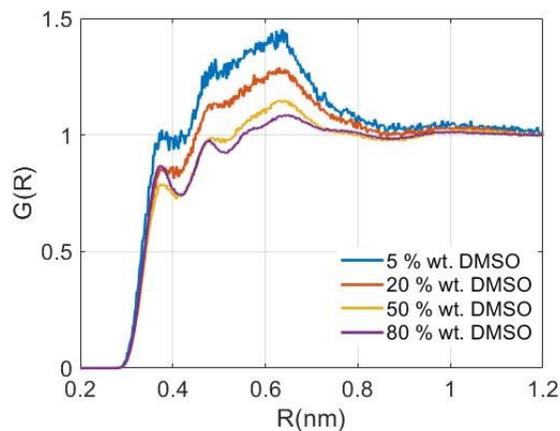
Figure A.14 Radial pair distributions for the 5-HMF C5– DMSO oxygen (OD) at all DMSO concentration studied. 5-HMF concentration of 10% wt.
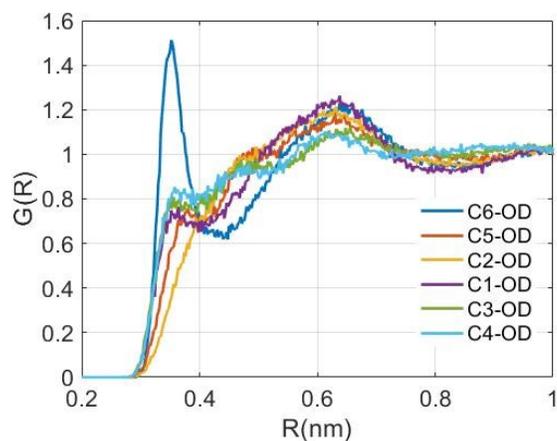
Figure A.15 Radial pair distributions for the 5-HMF carbon atoms– DMSO oxygen (OD) at all DMSO concentration studied. 5-HMF concentration of 1% wt.
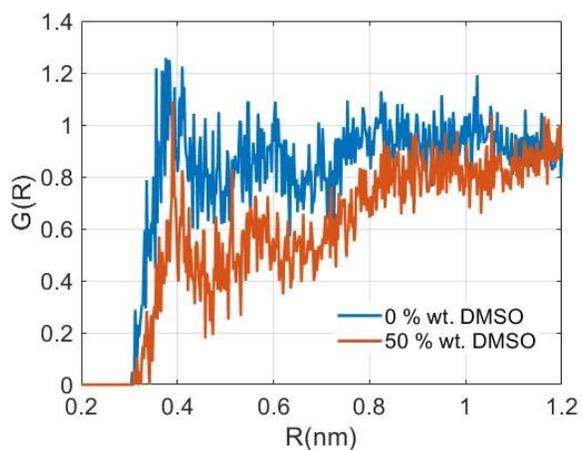
Figure A.16 Radial pair distributions for the 5-HMF C1– DMSO oxygen (OD) at DMSO concentration of 0 and 50 % wt. 5-HMF concentration of 1% wt.

A.2.3 H-bonds

Table A.4 Number of hydrogen bonds between DMSO and water molecules with Fructose/5-HMF. The H-Bonds were calculated at 3 Å Donor-Acceptor distance and 20 degrees of angle cutoff.

DMSO wt. %	No. H-Bonds fructose system	No. H-Bonds HMF system
0	9.7	0
5	9.8	0
20	9.5	0
50	8.5	0
80	6.9	0

A.2.4 Isovalues

Table A.5 Isovalues for systems with 1% wt. fructose/5-HMF. Constant isovalue of hydronium ions of 0.01.

DMSO wt. %	water isovalue	DMSO isovalue
0	0.55	------
5	0.60	0.05
20	0.50	0.15
50	0.40	0.50
80	0.15	0.55

A.2.5 Free energy surfaces
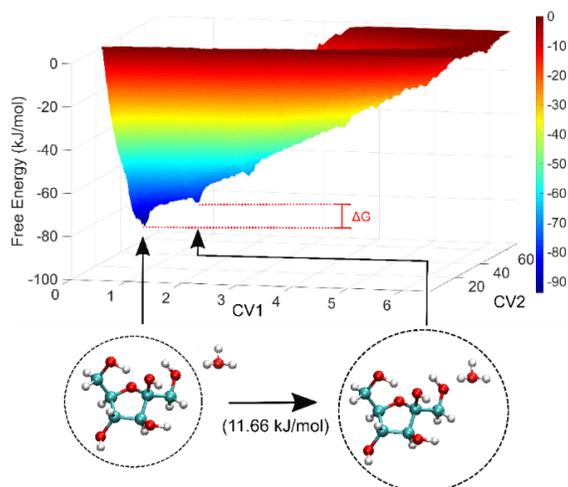
Figure A.17 Free energy surfaces as function of coordination number of water and hydronium ion with 1 % wt. fructose.
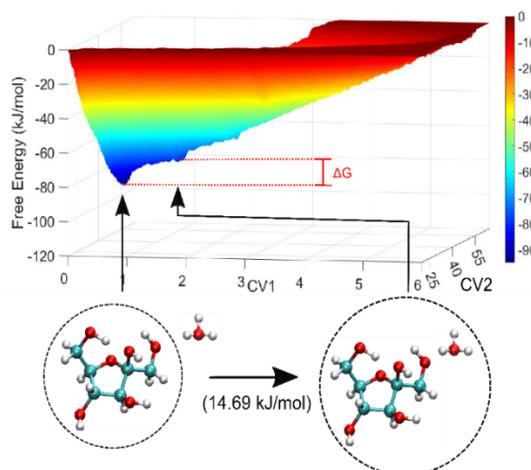
Figure A.18 Free energy surfaces as function of coordination number of water and DMSO 5 % wt. and hydronium ion with 1 % wt. fructose.
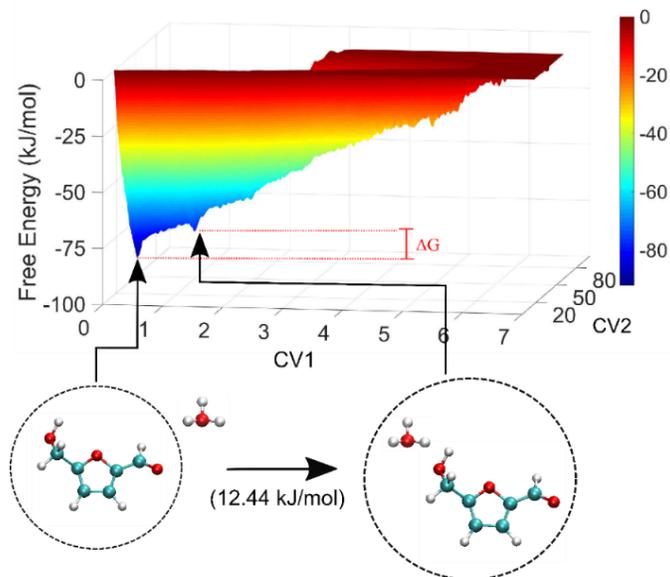
Figure A.19 Free energy surfaces as function of coordination number of water and DMSO 20 % t. and hydronium ion with 1 % wt. fructose.
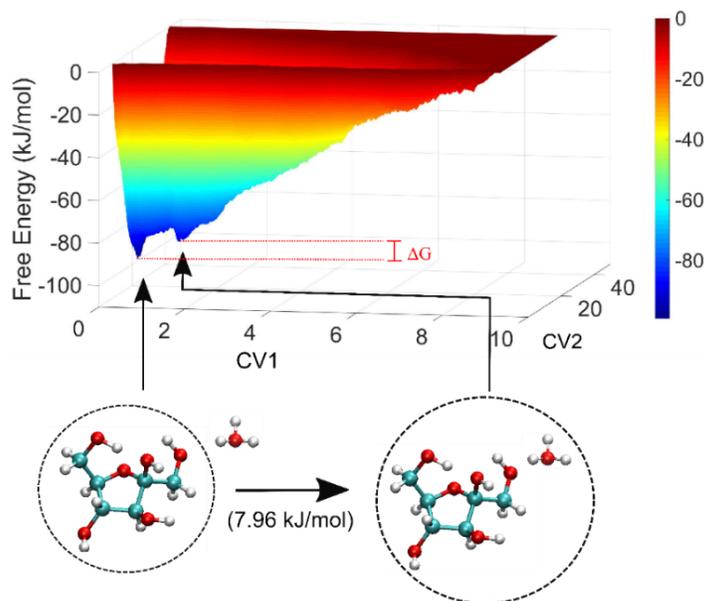
Figure A.20 Free energy surfaces as function of coordination number of water and DMSO 80 % wt. and hydronium ion with 1 % wt. fructose.
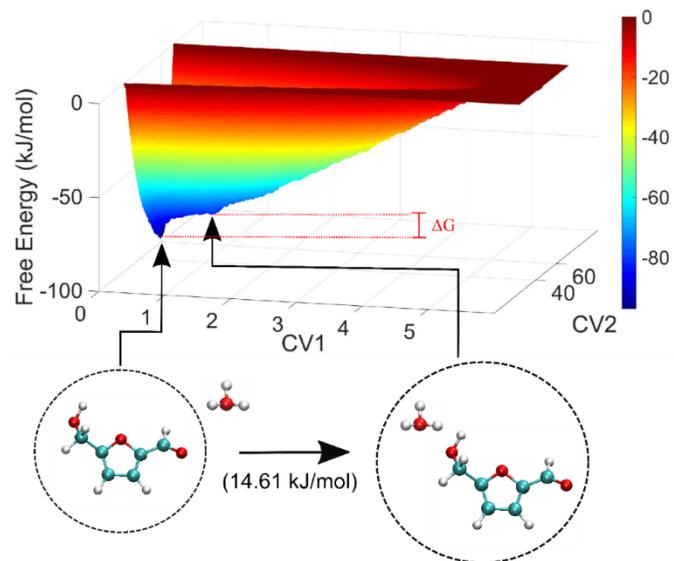
Figure A.21 Free energy surfaces as function of coordination number of water and hydronium ion with 1 % wt. 5-HMF.
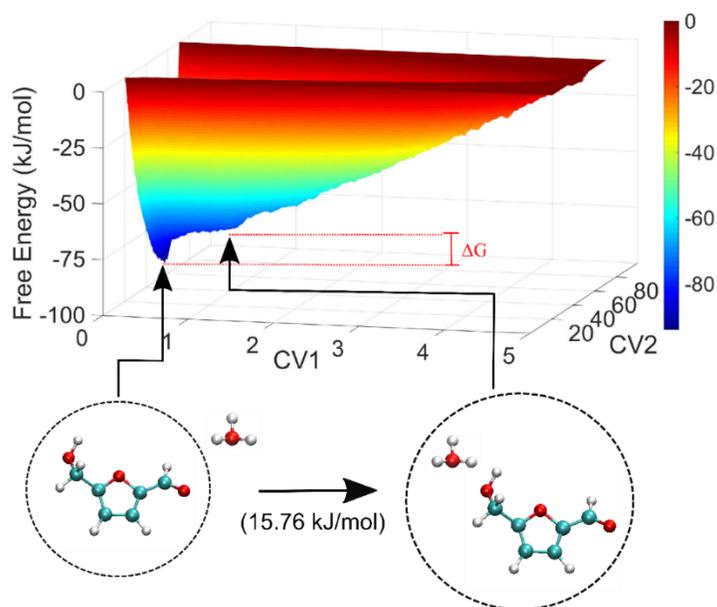
Figure A.22 Free energy surfaces as function of coordination number of water and DMSO 5 % wt. and hydronium ion with 1 % wt. 5-HMF.
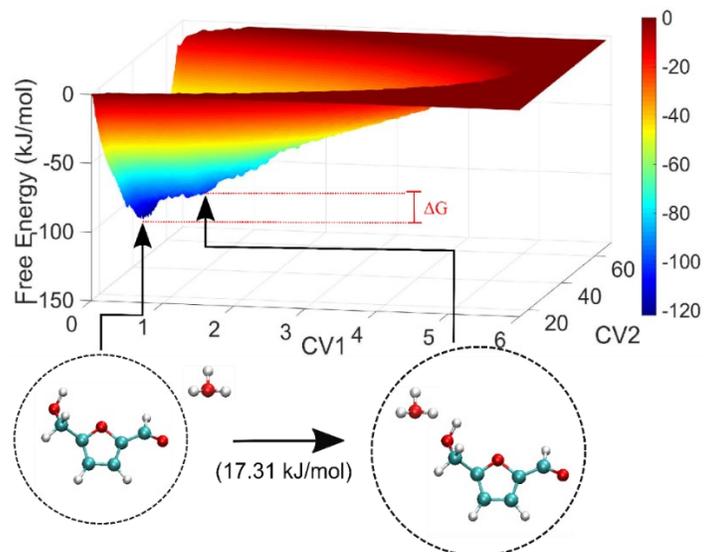
Figure A.23 Free energy surfaces as function of coordination number of water and DMSO 20 % wt. and hydronium ion with 1 % wt. 5-HMF.
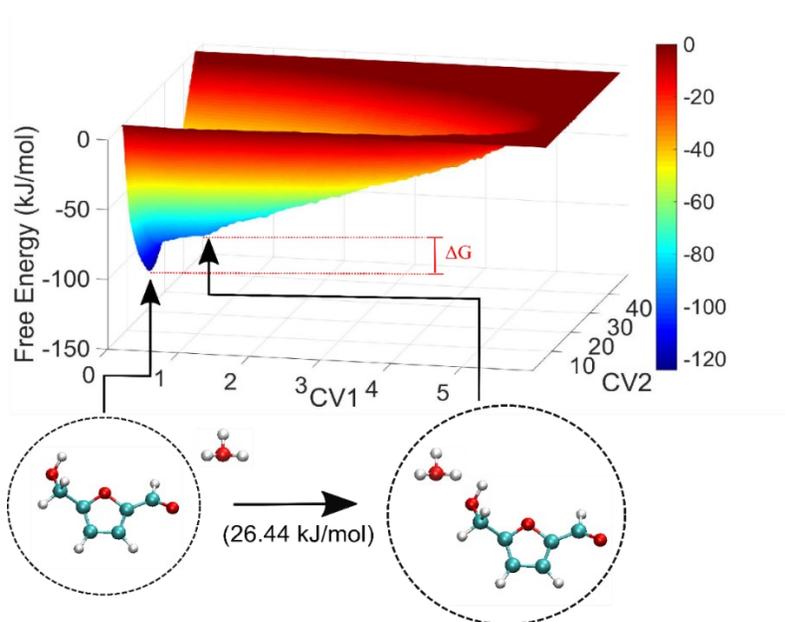
Figure A.24 Free energy surfaces as function of coordination number of water and DMSO 80 % wt. and hydronium ion with 1 % wt. 5-HMF.
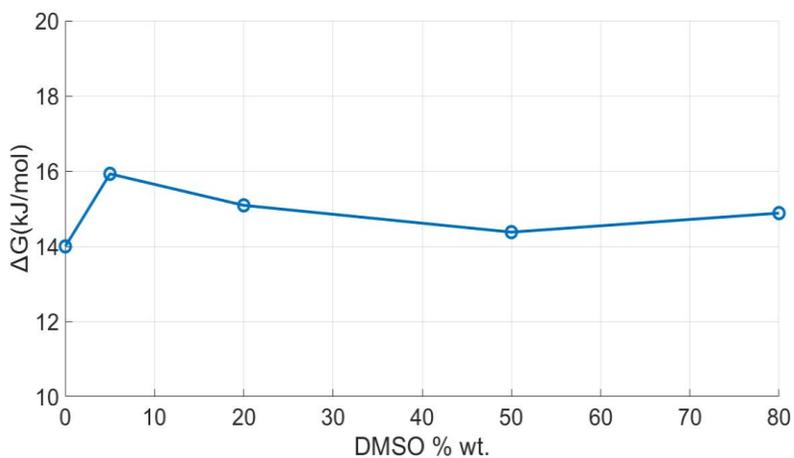
Figure A.25 Comparison of the barrier associated with the movement of proton to the immediate vicinity of fructose, due to the solvent reorganization at all DMSO concentration studied. Fructose 1 % wt.

A.2.6 WT-MTD converge criteria

i) Evolution of free energy surface with respect to time.

Taking the example of 50 % wt. of DMSO for fructose and HMF systems, following are the ways in which WT-MTD convergence was ensured.
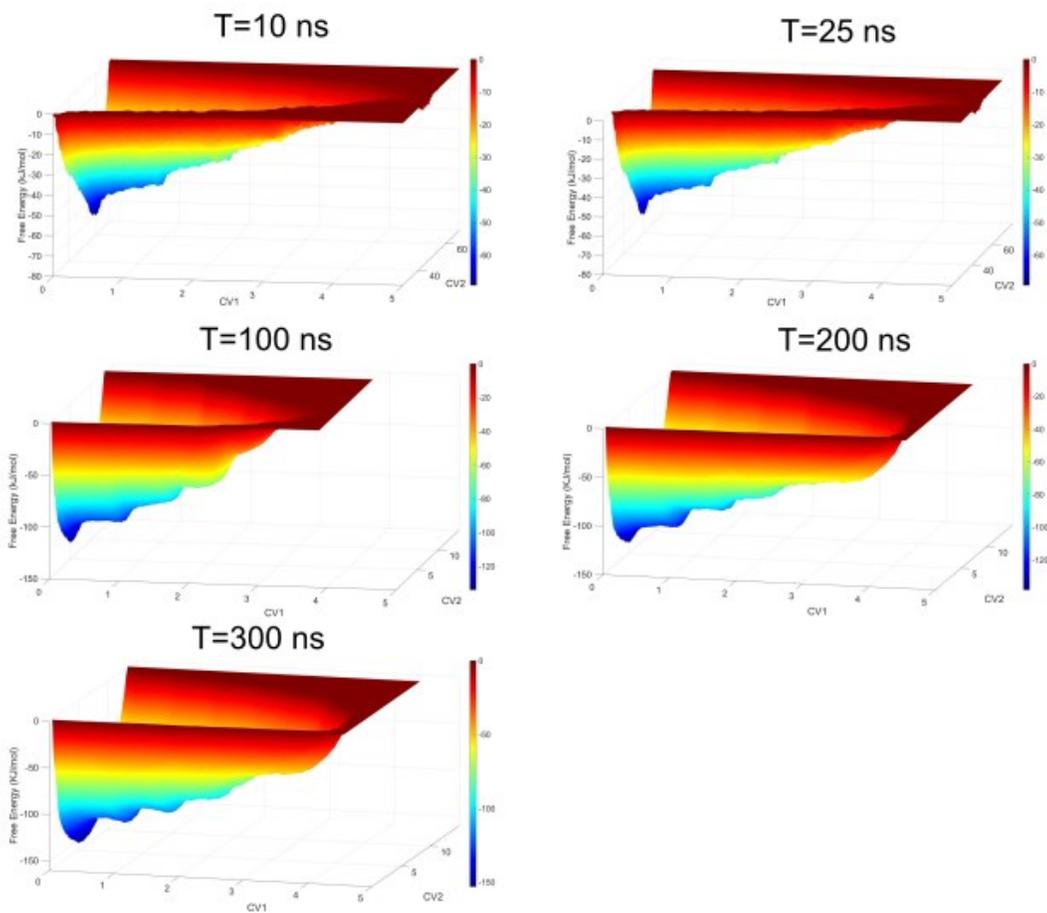
Figure A.26 Evolution of free energy surface for fructose 1 % wt. 50 % wt. of DMSO with respect to time.
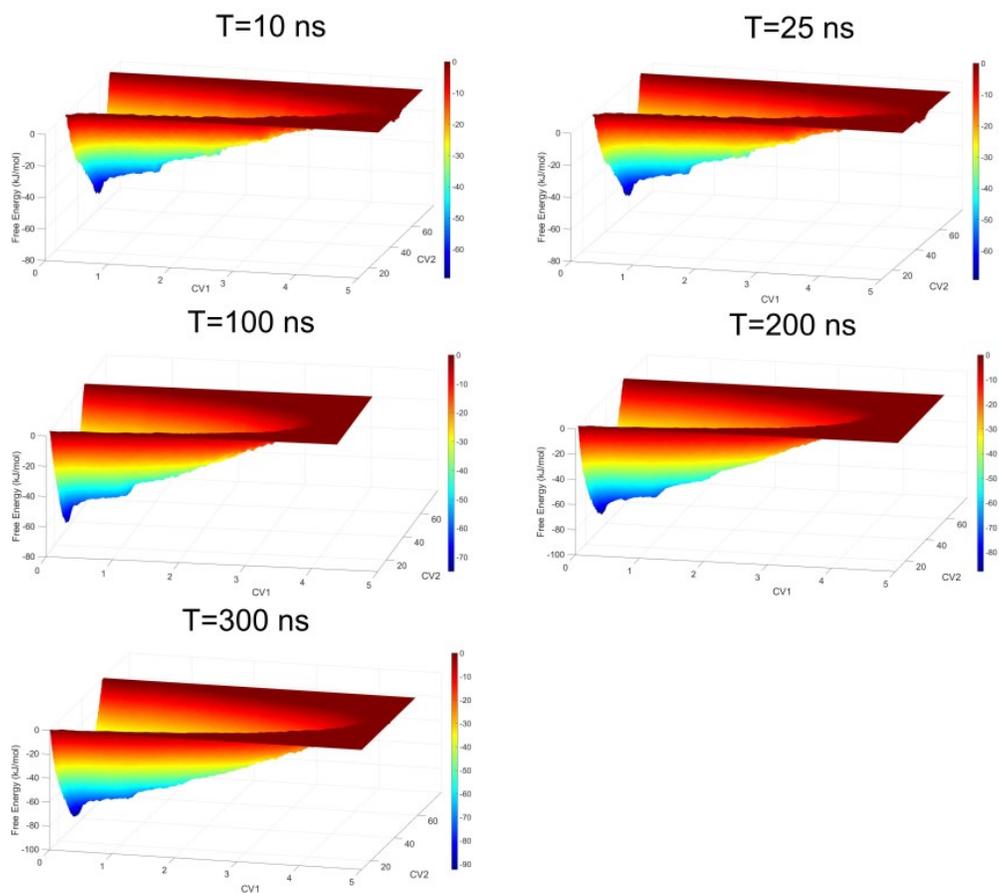
Figure A.27 Evolution of free energy surface for HMF 1 % wt. 50 % wt. of DMSO with respect to time.

ii) Collective variable CV1 with respect to time.
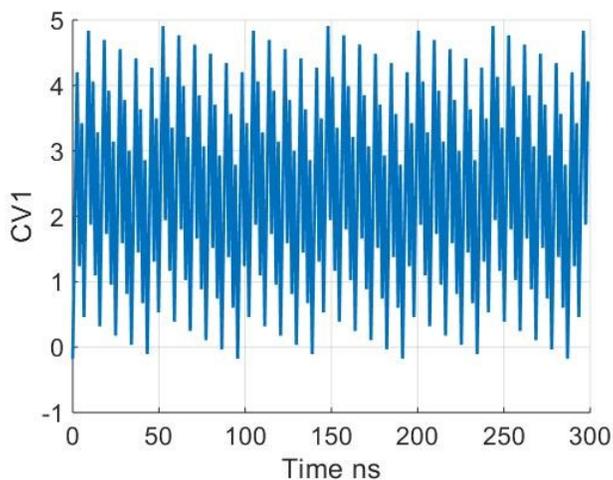
Figure 6.28 Change in coordination number between fructose molecules and hydronium ions taken as the collective variable (CV1), with respect to time.
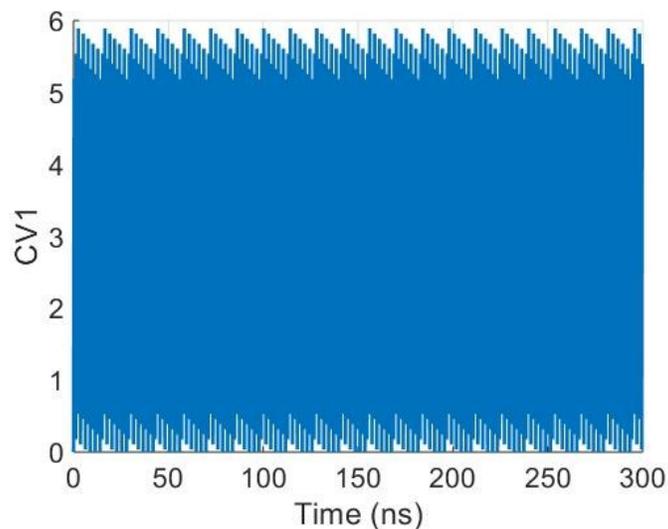
Figure A.29 Change in coordination number between HMF molecules and hydronium ions taken as the collective variable (CV1), with respect to time.

iii) Change in free energy difference with respect to time.

The ΔG (FES minimum corresponding to hydronium in bulk solvent – FES minimum corresponding to hydronium in the vicinity of fructose/HMF) for different DMSO concentrations are plotted. It was observed that when the simulations are converged, the fluctuations in the free energy difference significantly reduce.
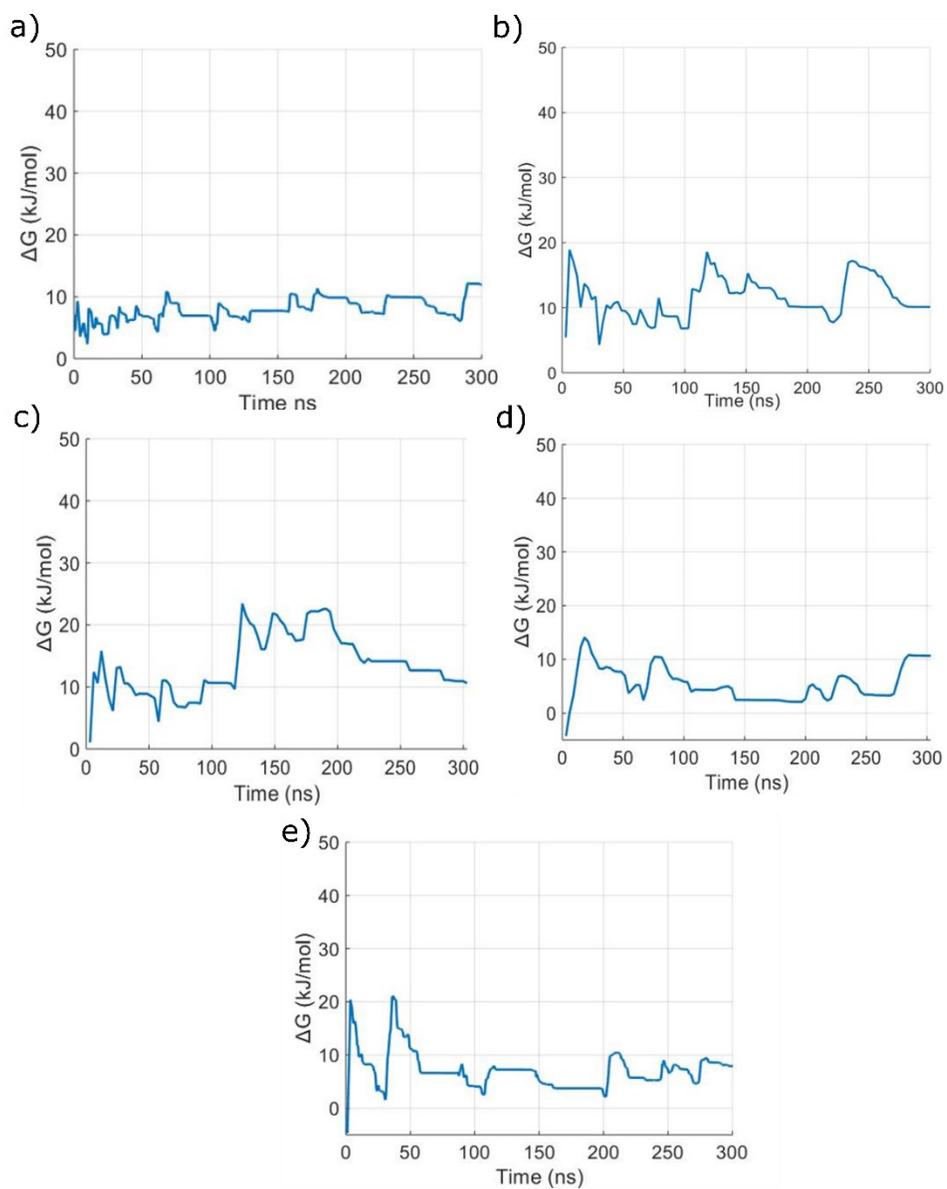
Figure A.30 Change in free energy difference for 1 % wt. fructose a) 0 wt. % DMSO b) 5 wt. % DMSO c) 20 wt. % DMSO d) 50 wt. % DMSO e) 80 wt. % DMSO.
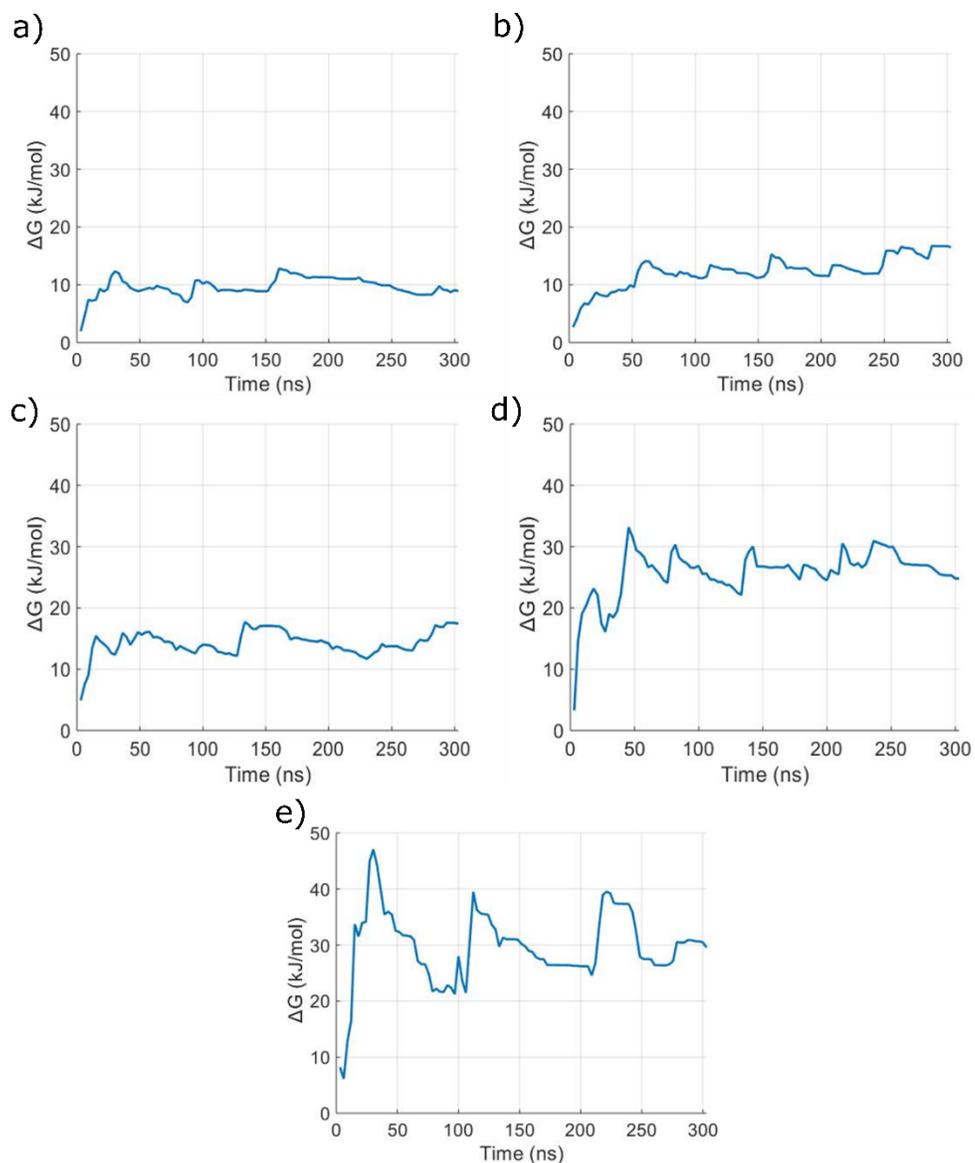
Figure A.31 Change in free energy difference for 1 % wt. HMF a) 0 wt. % DMSO b) 5 wt. % DMSO c) 20 wt. % DMSO d) 50 wt. % DMSO e) 80 wt. % DMSO.

A.3 Appendix to Chapter 4

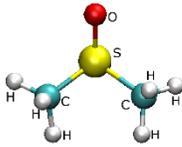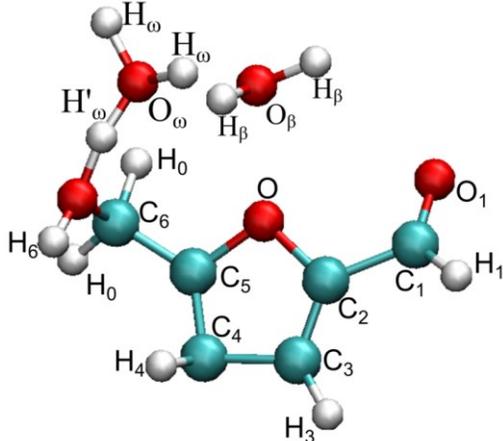**A.3.1 First principles DFT calculations – Implisolv method.**

Table A.6 Free energy values for all the intermediate species and transition state for protonation of 5-HMF. Values are reported in kJ/mol.

	Gas Phase	**DMSO**	**water**
Reactant	0	0	0
TS	92.8	94.5	93.9
Product	66.1	82.7	82.0

A.3.2 Explicit equilibrated solvation non polarizable without solvent dynamics - Resolv method.

A.3.2.1 Atomic charges and simulation parameters.

Table A.7 Atomic charges and simulation parameters implemented in the present work for water, DMSO, 5-HMF and fructose. Blue balls indicate carbon atoms, red balls indicate oxygen atoms, white balls indicate hydrogen atoms and yellow ball indicates sulfur atom.

Molecule	Charges (in parentheses) and parameters, e/kJ mol-1, σ/Å
Water	O (-0.834)-3.1,0.64 H(0.417)-0.0,0.00
Hydronium ion	O(-0.626)- 3.1,0.64 H(0.542)-0.0,0.00
DMSO	S(0.282)- 3.5,1.65 O(-0.515)-2.9,1.17 C(-0.406)- 3.8,0.27 H(0.174)- 2.5,0.21
Reactant	O_1(-0.504)- 2.90,0.88 C_1(0.412)- 3.70,0.44 H_1(0.022)- 2.40,0.63 C_2(0.020)- 3.50,0.32 O(-0.210)- 2.90,0.58 C_5(0.210)- 3.50,0.32 C_4(0.303)- 3.50,0.32 H_4(0.196)- 2.40,0.13 C_3(-0.086)- 3.50,0.32 H_3(0.165)- 2.40,0.13 C_6(0.096)- 3.90,0.49 H_0(0.079)- 2.50,0.13 O_6(-0.601)- 3.10,0.71 H_6(0.416)-0.00,0.00 $O\omega$(-0.834)-3.10,0.64 O_β(-0.834)-3.10,0.64 $H'\omega$(0.542)-3.10,0.64 $H\omega$(0.542)-0.00,0.00 H_β(0.417)- 0.00,0.00
TS	O_1(-0.504)-2.90,0.88 C_1(0.412)-3.70,0.44 H_1(0.022)-2.40,0.69 C_2(0.020)-3.50,0.32 O(-0.210)-2.90,0.58 C_5(0.210)-3.50,0.32 C_4(0.303)-3.50,0.32 H_4(0.196)-2.40,0.13 C_3(-0.086)-3.50,0.32 H_3(0.165)-2.40,0.16 C_6(0.096)-3.90,0.49 H_0(0.079)-2.50,0.13
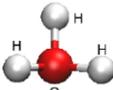
O$_6$(-0.601)-3.10,0.71
H$_6$(0.416)-0.00,00
Oω(-0.834)-3.10,0.64
O$_\beta$(-0.834)-3.10,0.64
H'ω(0.542)-3.10,0.64
Hω(0.542)-0.00,0.00
H$_\beta$(0.417)- 0.00,0.00

Product

O$_1$(-0.434)-2.9,0.88
C$_1$(0.310)-3.7,0.44
H$_1$(0.218)-2.4,0.69
C$_2$(0.155)-3.5,0.32
O(-0.392)-2.9,0.58
C$_5$(0.665)-3.5,0.32
C$_4$(-626)-3.5,0.32
H$_4$(0.350)-2.4,0.13
C$_3$(-0.151)-3.5,0.32
H$_3$(0.322)-2.4,0.13
C$_6$(-0.299)-3.9,0.49
H$_0$(0.288)-2.5,0.15
O$_2$(-0.763)-3.1,0.71
H$_6$(0.326)-0.0,0.00
Oω(-1.052)-3.1,0.64
O$_\beta$(-1.033)-3.1,0.64
H'ω(0.382)-3.1,064
Hω(0.518)-0.0,0. 00
H$_\beta$(0.533)-0.0,0. 00

Table A.8 Number of molecules and simulation cell size for reactant, TS, and product. at different concentrations of DMSO. For all systems, 1 hydronium ions.

DMSO wt. %	No. water	No. DMSO	Cell size (length) nm
0	1400	0	3.711
25	1050	80	3.643
50	700	162	3.559
75	239	350	3.489

A.3.3 RDF$_S$ between reactant, TS, and product with solvent molecules.

A.3.3.1 Reactant
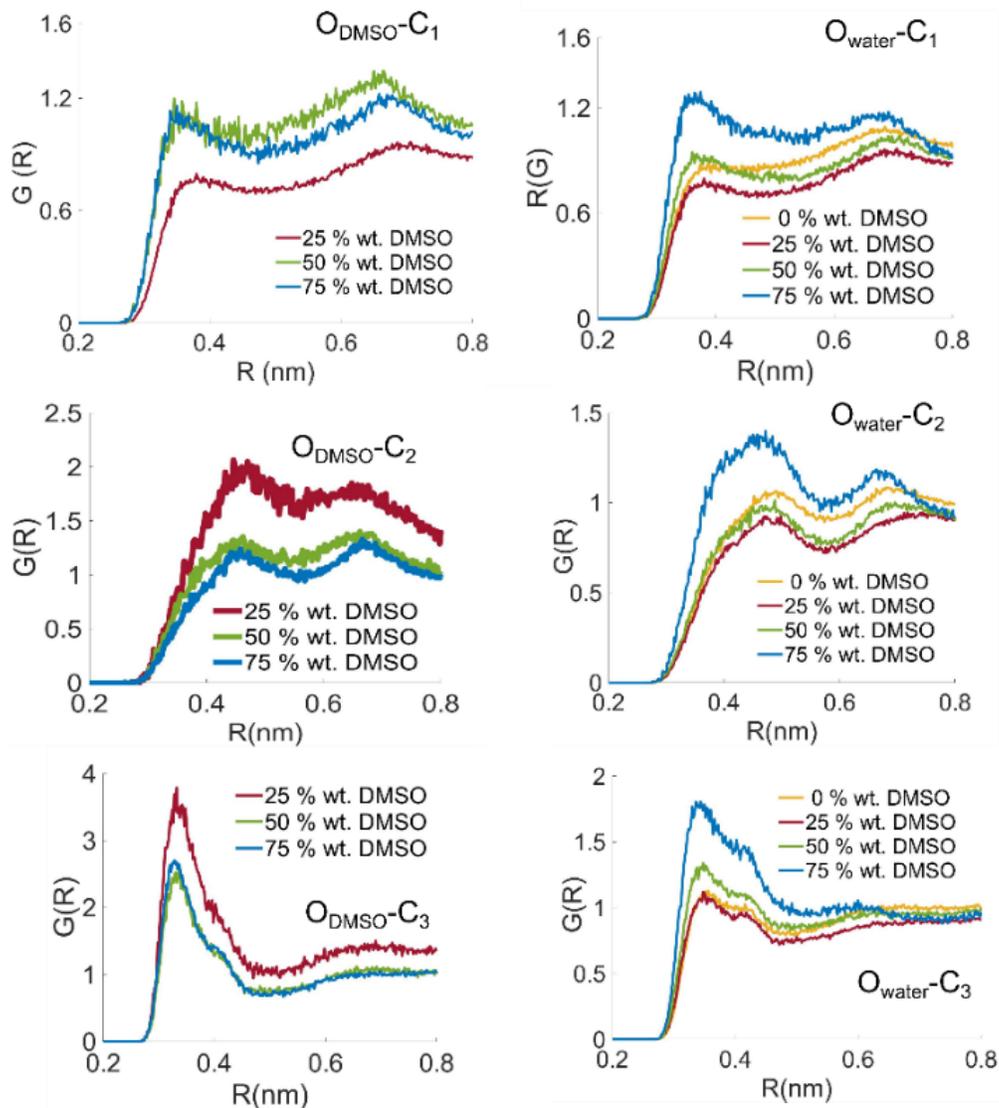
Figure A.32 Radial pair distribution Reactant - C$_1$, C$_2$ C$_3$ with solvent oxygen atom pairs, at all DMSO concentration studied.
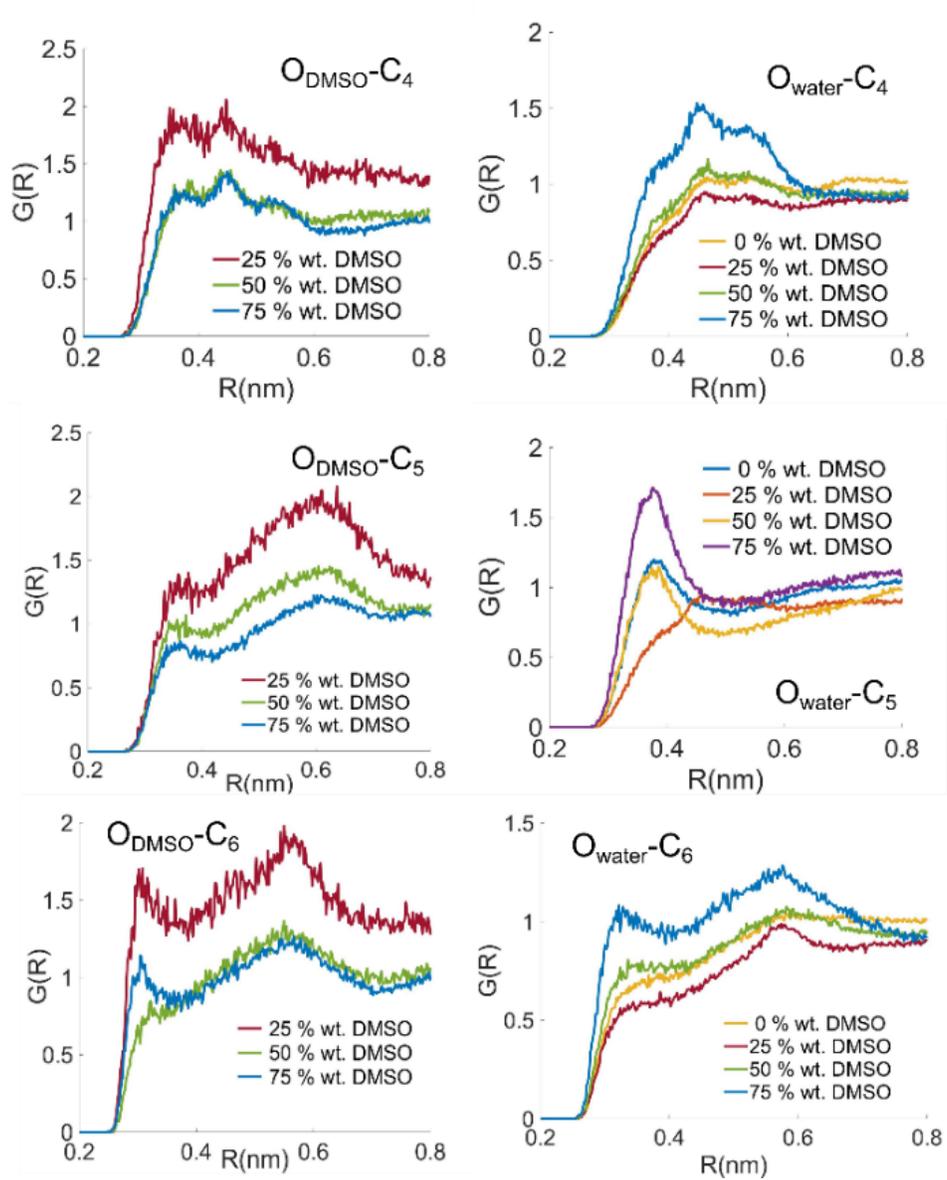
Figure A.33 Radial pair distribution Reactant - C_4, C_5 C_6 with solvent oxygen atom pairs, at all DMSO concentration studied.
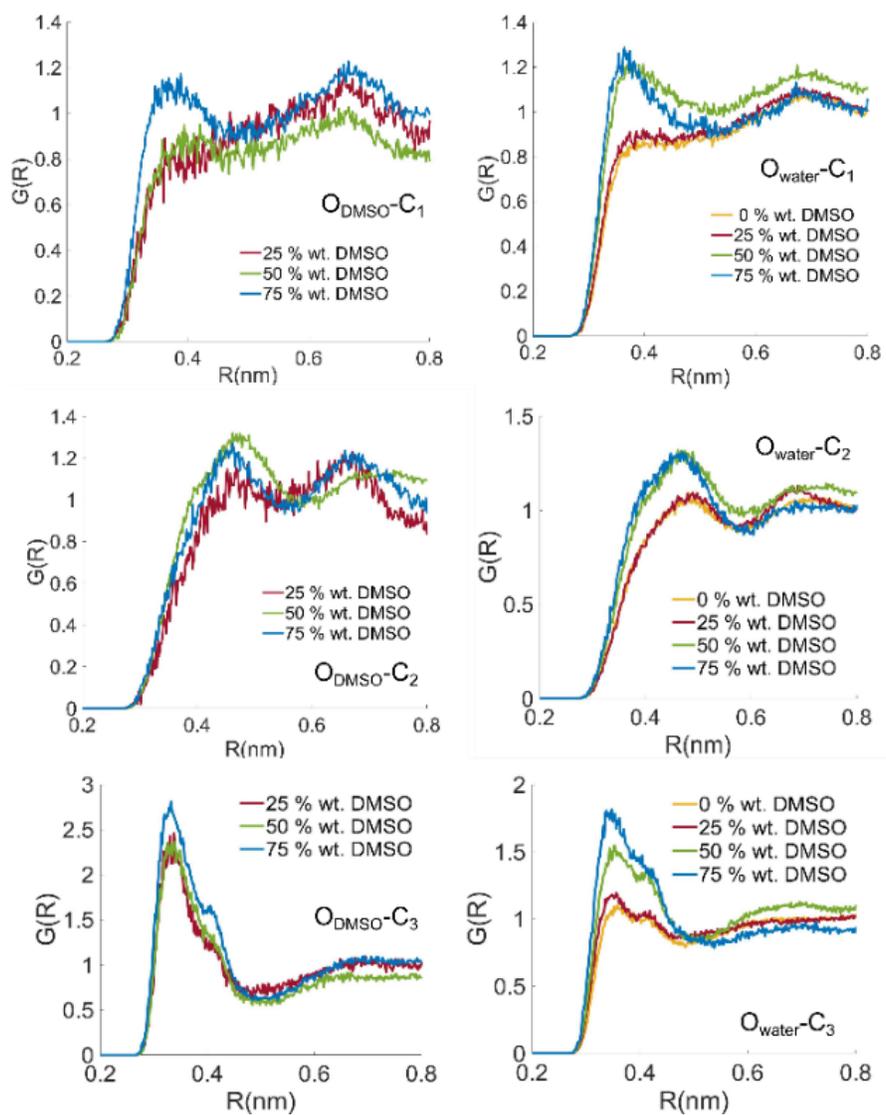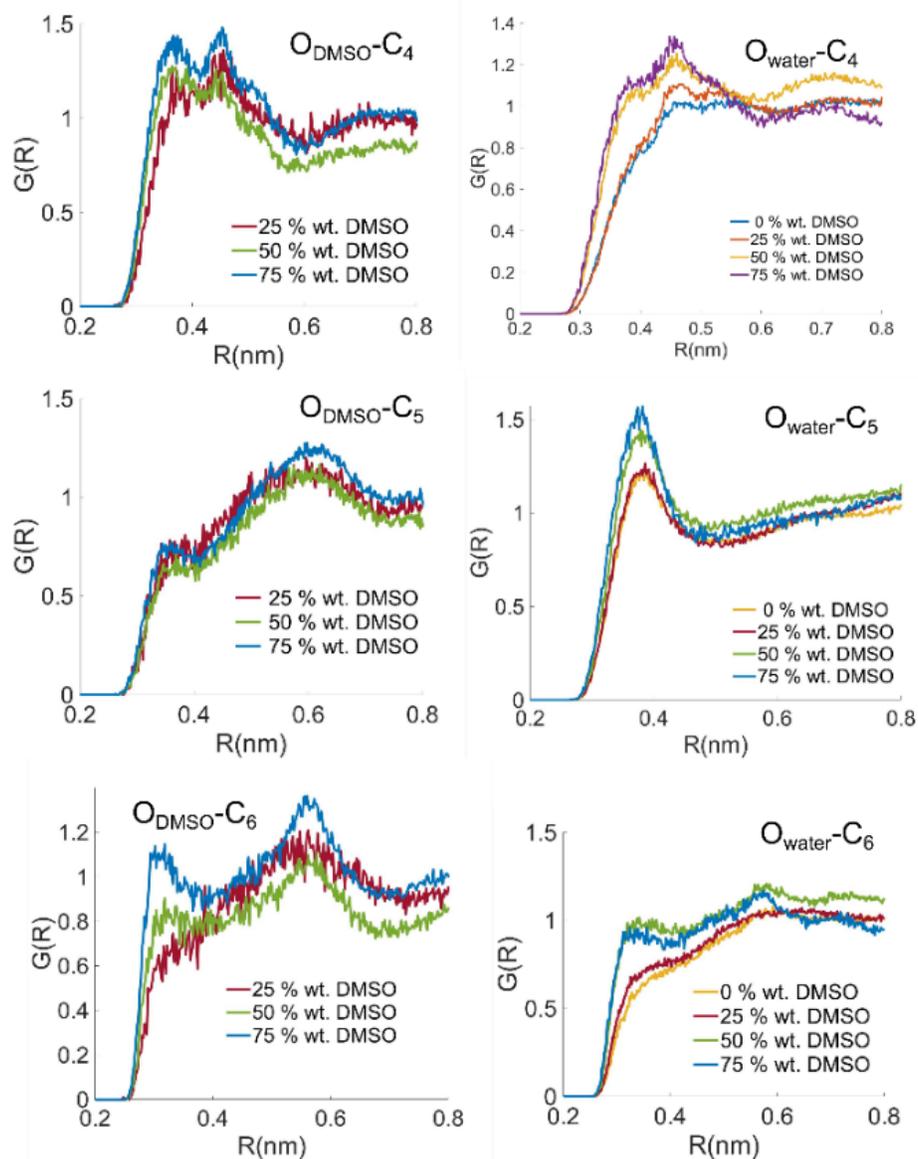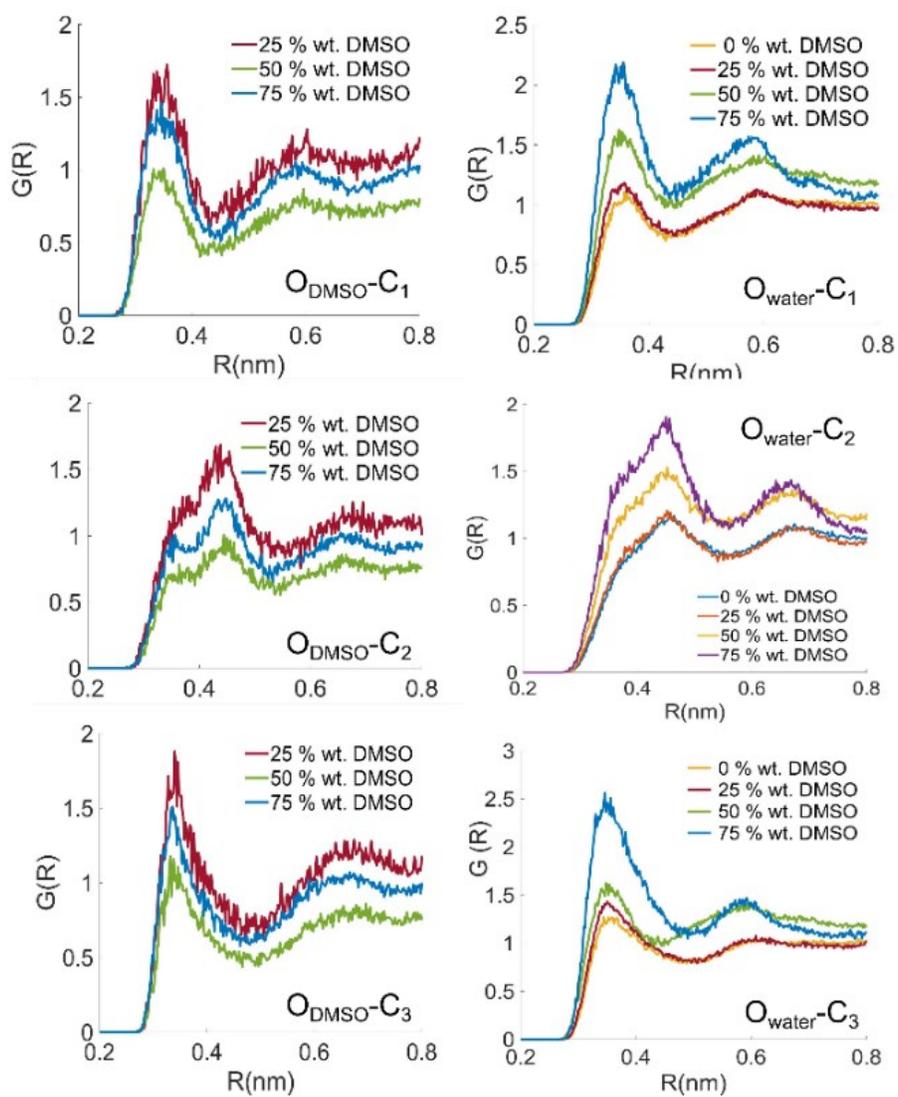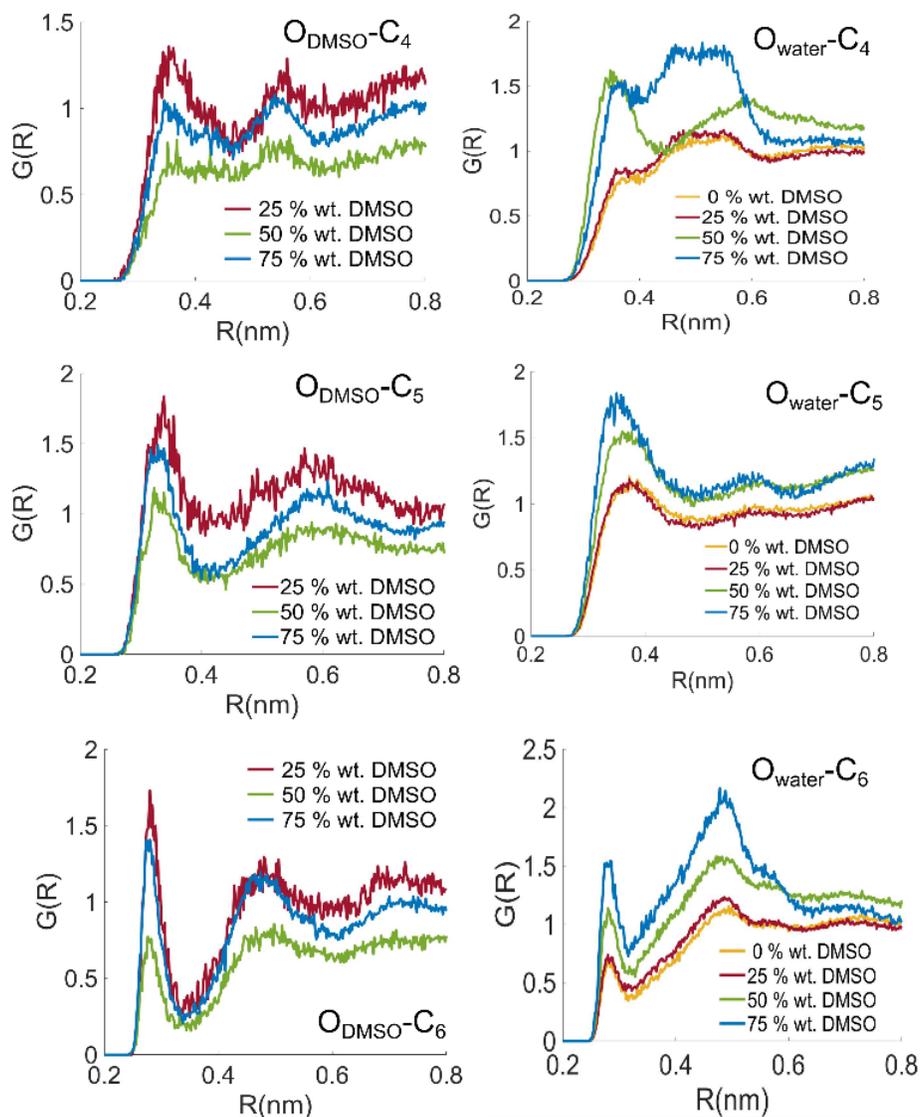
A.3.3.2 TS

Figure A.34 Radial pair distribution TS - C_1, C_2 C_3 with solvent oxygen atom pairs, at all DMSO concentration studied.

Figure A.35 Radial pair distribution TS - C_4, C_5 C_6 with solvent oxygen atom pairs, at all DMSO concentration studied.

A.3.3.3 Product

Figure A.36 Radial pair distribution Product - C_1, C_2 C_3 with solvent oxygen atom pairs, at all DMSO concentration studied.

Figure A.37 Radial pair distribution Product - C_4, C_5 C_6 with solvent oxygen atom pairs, at all DMSO concentration studied.

A.3.4 Ab initio molecular dynamics (CPMD scheme)-metadynamics implementation – converge criteria for Dynasolv method

In this section, we present the converge criteria to determine the validity of the data reported using Dynasolv for 0, 25, and 50 % wt. DMSO proportions. This analysis determines until which simulation time the states of interest (5-HMF before and after protonation) converge. For each DMSO proportion, the following converge criteria are reported.

(i) Activation and reaction-free energy values change as a function of simulation time.

(ii) Free energy surfaces at different simulation times to compare their evolution and the possible exploration of other regions according to the collective variables specified.

(iii) Key bond distances change as a function of simulation time.

A.3.4.1 System - 0 % wt. DMSO
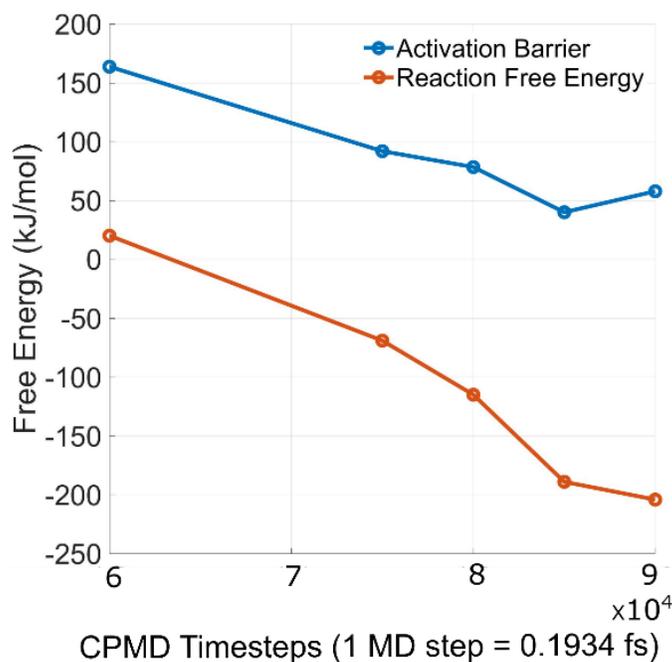
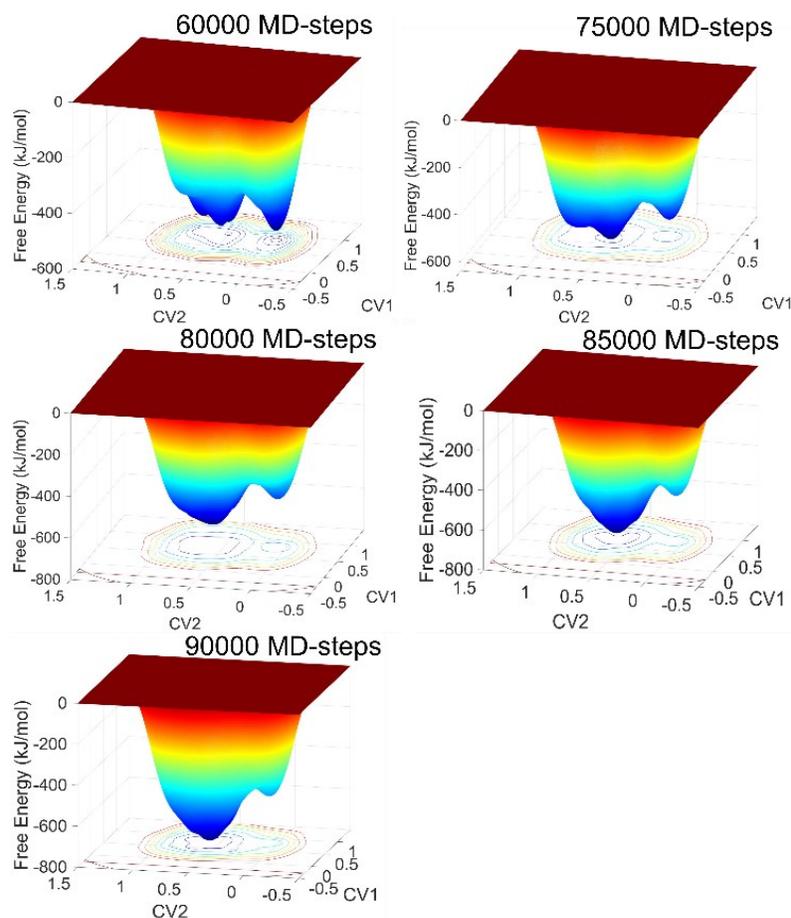**Figure A.38 Change of Activation Barrier and Reaction Free Energy in function of simulation time at 0 % wt. DMSO.**

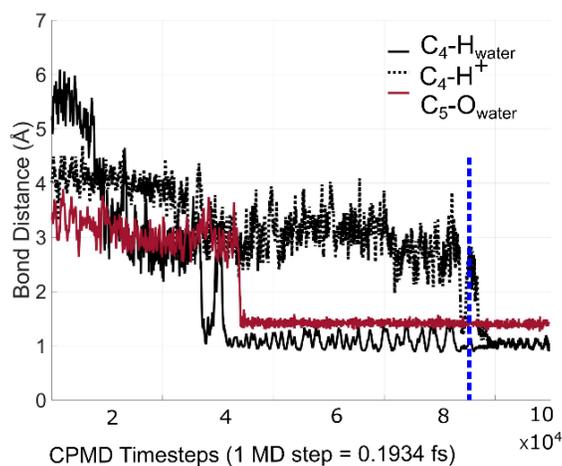**Figure A.39 Evolution of Free Energy Surface as a function of simulation time at 0 % wt. DMSO.**

Figure A.40 Variation in the bond distances during simulation time at 0 % wt. DMSO.

Activation barrier and reaction-free energy changes during the simulation time are reported in Figure A.38. It is observed that both the activation barrier and reaction-free energy converge over time. Additionally, Figure A.39 shows how the wells within the FES evolve along the entire CPMD-metadynamics simulation. The first well is at CV1 ≈ 1 and CV2 ≈ 0 (5-HMF before protonation), and the second well is at CV1 ≈ 1 and

CV2 ≈ 1 (5-HMF after protonation). Finally, the bond distance between C_4 - H^+, C_4 - H_{water}, and C_4- O_{water} is reported in Figure A.40. It is observed that the protonation occurs after 30 000 MD steps (~6 ps), and O_{water} instantaneously stabilizes thh electro deficient C_5. After 80 000 MD steps (~16 ps), H^+ forms a bond with C_4, which is still bonded to H_{water}. This observation is enough to ensure that the two wells corresponding to 5-HMF before and after protonation have been filled with the potential added during the CPMD-metadynamics run allowing the system to explore new regions in the energy space bounded by the CVs.
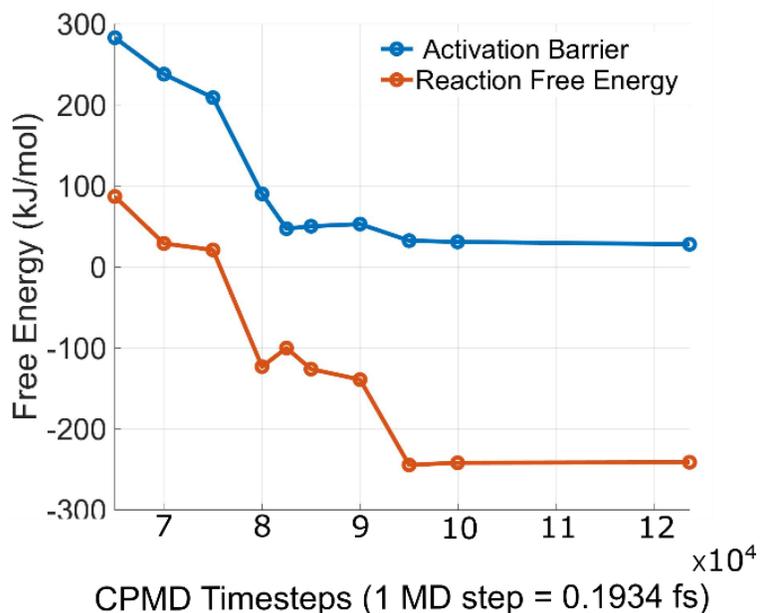
6.3.4.2 System - 25 % wt. DMSO

Figure A.41 Change of Activation Barrier and Reaction Free Energy in function of simulation time at 25 % wt. DMSO.
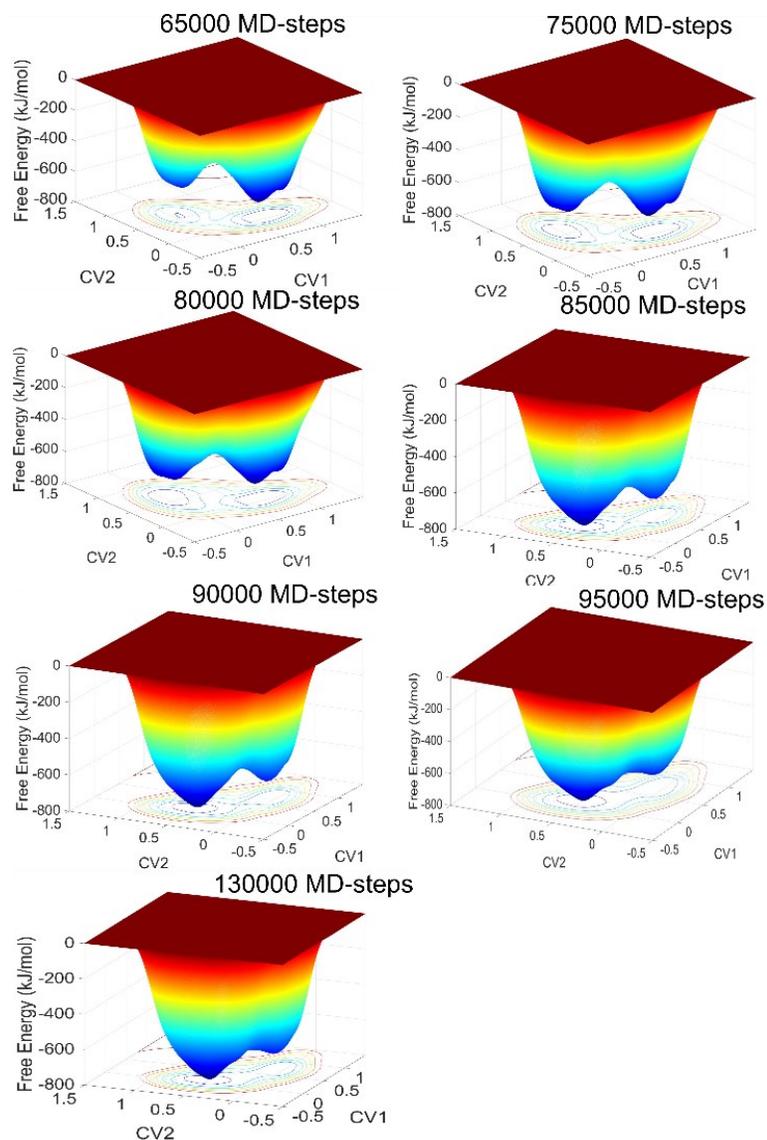
Figure A.42 Evolution of Free Energy Surface as a function of simulation time at 25 % wt. DMSO.
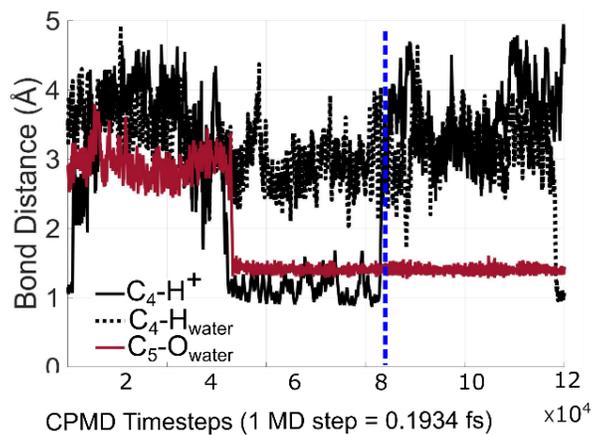

Figure A.43 Variation in the bond distances during simulation time at 25 % wt. DMSO.

Activation barrier and ΔG_R during the simulation time are reported in Figure A.41. It is observed that both the activation barrier and ΔG_R converge after 80 000 MD steps (~16 ps). It is determined that the simulation converges at ~16 ps. According to Figure A.42, after 80 000 MD steps, bond C_4 - H^+ is broken leading to the reactant state. Afterward, it migrates to the product well and later it returns to the reactant state. This means that both the reactant and product wells have been filled with the potential added during the CPMD-metadynamics run. If the simulation continues, the FES may get deformed; therefore, the simulation stops after 16 ps. Similar to pure water, O_{water} stabilizes C_5^+ after protonation occurs. In this case, as can be seen in Figure A.43, the first well corresponding to the reactant is at CV1 ≈ 1 and CV2 ≈ 0 (5-HMF before protonation), and the second well is at CV1 ≈ 0 and CV2 ≈ 1 (5-HMF after protonation).

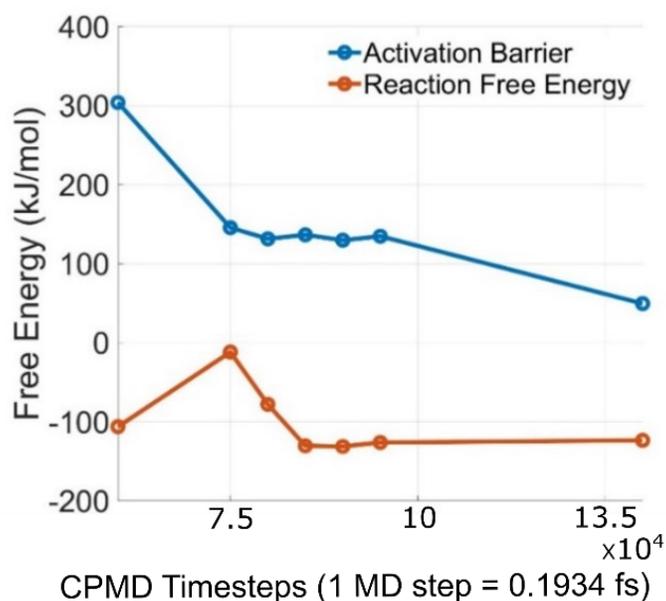### 6.3.4.3 System - 50 % wt. DMSO

Figure A.44 Change of Activation Barrier and Reaction Free Energy in function of simulation time at 50 % wt. DMSO.
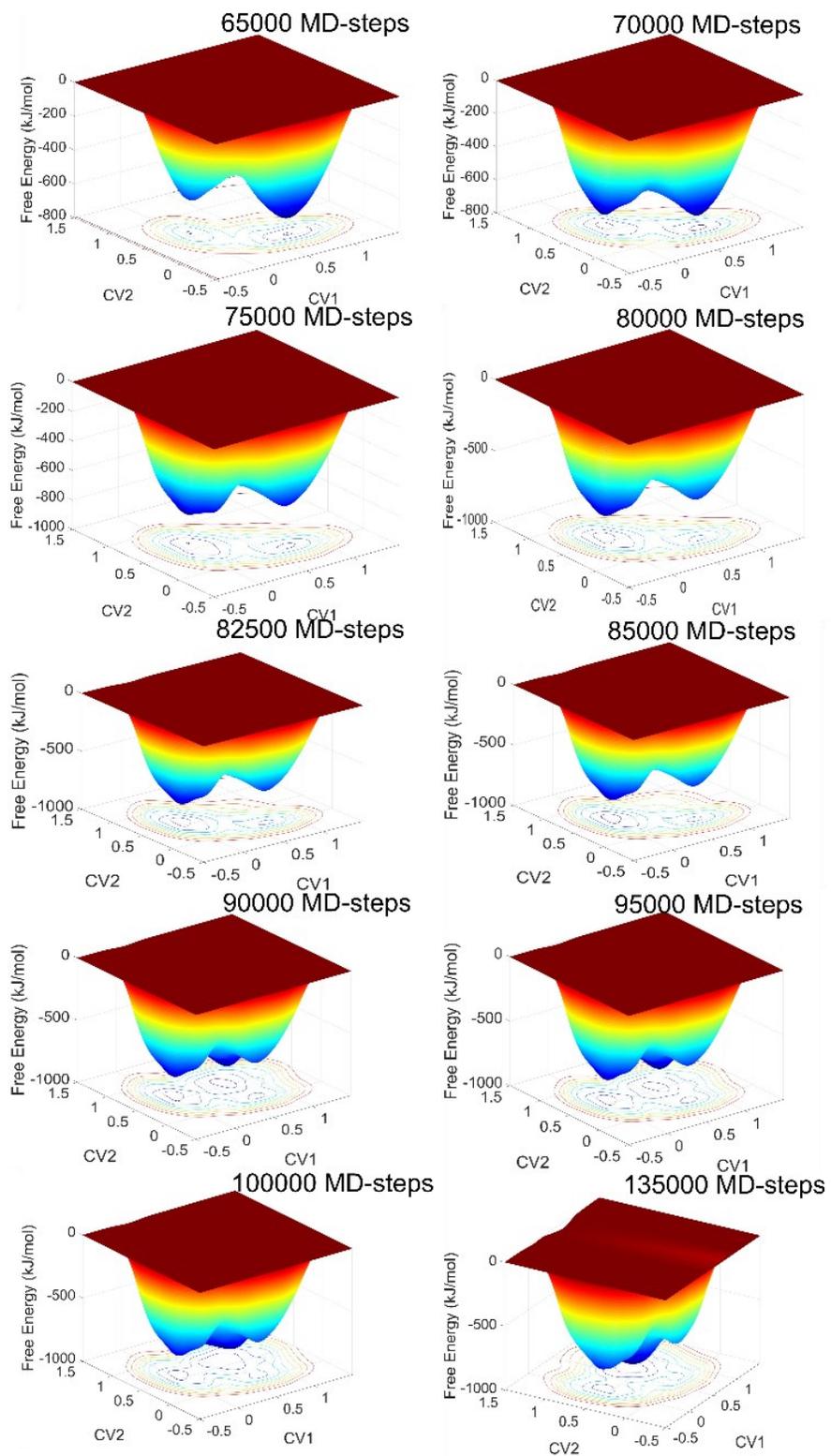
Figure A.45 Evolution of Free Energy Surface as a function of simulation time at 50 % wt. DMSO.
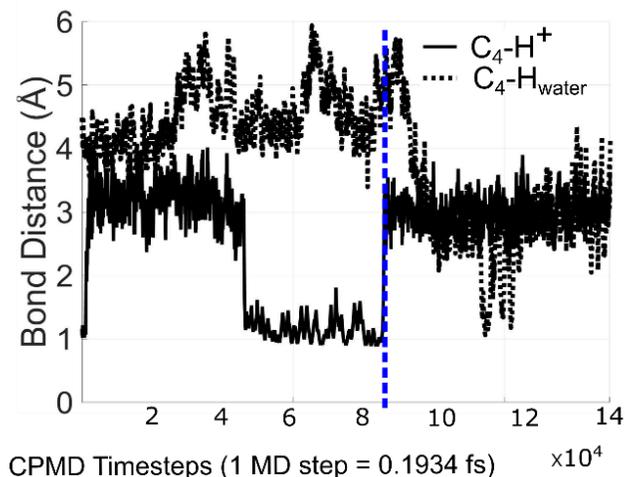
Figure A.46 Variation in the bond distances during simulation time at 50 % wt. DMSO.

Finally, the activation barrier and reaction-free energy for 50 % wt. DMSO is shown in Figure A.44. It is observed how both the activation barrier and ΔG_R converge after 80 000 (~16 ps). Figure A.45 shows the FES evolution over time, like 25 % wt. DMSO, the first well corresponding to the reactant, is at CV1 ≈ 1 and CV2 ≈ 0 (5-HMF before protonation), and the second well is at CV1 ≈ 0 and CV2 ≈ 1 (5-HMF after protonation). It is also observed how a third well at 90 000 MD steps is formed. The well is at CV1 ≈ 1 and CV2 ≈ 1, meaning H_{water} has replaced the proton to be bonded to C_4. The third well formation also suggests that the previous wells corresponding to reactant and product (before and after 5-HMF protonation) have been filled with the potential added during CPMD-metadynamics. In addition, Figure A.46 shows that after 80 000, the bond between $C_4 - H^+$ is broken, which that the system has moved from product to reactant, suggesting that both reactant and product wells have been successfully filled.

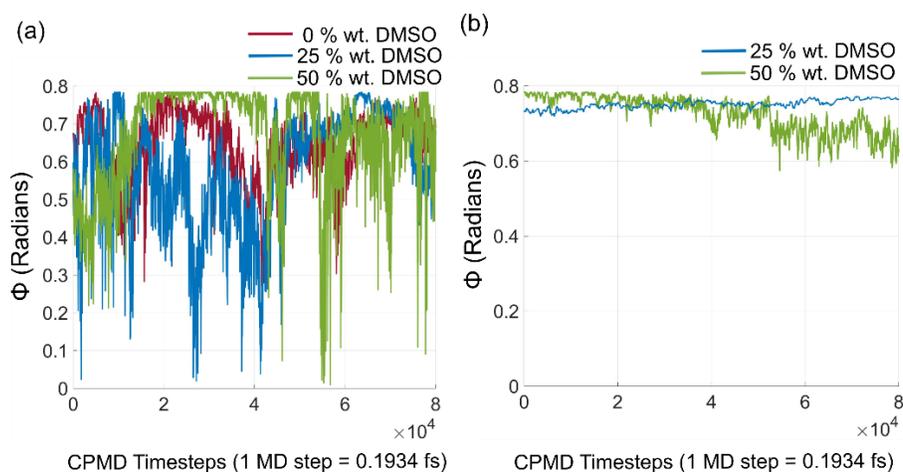**A.3.5 Dipole orientational vectors for water and DMSO along the trajectory**

Figure A.47 Variation in the orientation of the solvent dipole vector for the water and DMSO molecules separately with respect to the z-axis (or the angle that the solvent dipole vector makes with the z-axis of the simulation cell) during the protonation step in the CPMD-metadynamics trajectory. (a) water molecules (b) DMSO molecules at different DMSO proportions.
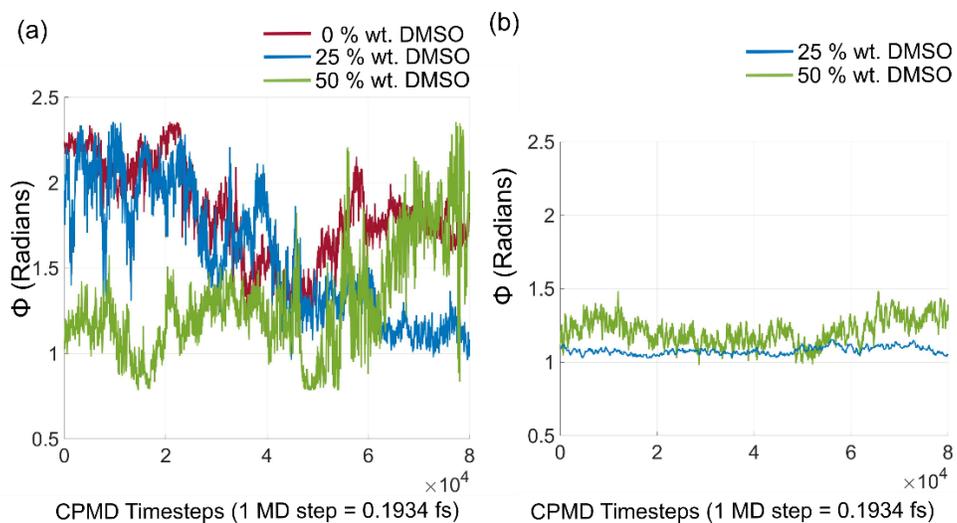
128

Figure A.48 Variation in the orientation of the solvent dipole vector for the water and DMSO molecules separately with respect to the y-axis (or the angle that the solvent dipole vector makes with the y-axis of the simulation cell) during the protonation step in the CPMD-metadynamics trajectory. (a) water molecules (b) DMSO molecules at different DMSO proportions.