STRUCTURAL STABILITY OF PROTEINS IN DEEP EUTECTIC SOLVENTS AND THEIR AQUEOUS SOLUTIONS

Samal Kaumbekova, Bachelor in Chemistry

Submitted in fulfillment of the requirements
for the degree of Master of Science
in Chemical Engineering

School of Engineering
Department of Chemical Engineering
Nazarbayev University

53 Kabanbay Batyr Avenue,
Astana, Kazakhstan, 010000

Supervisors: Dr. Dhawal Shah and
Dr. Mehdi Torkmahalleh

December 2018

DECLARATION

I hereby, declare that this manuscript, entitled "Structural stability of proteins in Deep Eutectic Solvents and their aqueous solutions", is the result of my own work except for quotations and citations which have been duly acknowledged.

I also declare that, to the best of my knowledge and belief, it has not been previously or concurrently submitted, in whole or in part, for any other degree or diploma at Nazarbayev University or any other national or international institution.

Name: Samal Kaumbekova
Date: 13.12.2018

Abstract

Deep Eutectic Solvents (DESs) are eutectic mixtures composed of different cations and anions of Lewis and Bronsted acids and bases. With their favorable properties such as low volatility, biodegradability, low cost, simple synthesis, etc., they are considered as green media and sustainable replacements of ionic liquids (1). One amongst many promising applications of DESs is maintaining the structure of protein in the absence of water. Experimental studies conducted on different proteins in DESs and their aqueous solutions have shown that the secondary and most of the tertiary protein structure is maintained in the DESs and that refolding can also be achieved in aqueous DESs (2, 3). However, the experiments are limited and the intermolecular interactions between the solvents and protein are obscured. To this end, we perform molecular dynamics simulations of lysozyme in its native structure in the presence of DES composed of urea and choline chloride, known as Reline, and its aqueous solutions. The simulations are carried to analyze (i) protein structure at room temperature, (ii) protein structure at high temperatures, when unfolding occurs and (iii) refolding of thermally unfolded protein structure at room temperature in reline and its aqueous solutions. The interactions between solvents and protein are evaluated in terms of radial distribution function (RDF) and hydrogen bonding patterns. Protein structure is studied in the form of its deviation from the native structure. The results suggest stability of the

secondary structures of lysozyme in the presence of reline.

Acknowledgements

Firstly, I would like to give thanks to my supervisor Dr. Dhawal Shah for sharing with me with his knowledge, giving me support, guidance and valuable feedback during my MSc program. Secondly, I would like to give my acknowledgements to my husband for believing in me and motivating me towards new academic achievements. Finally, I would also like to thank my parents for showing enormous support and respect towards my studies.

Table of Contents

List of Figures

Chapter 1 – Introduction

Solvents are widely used in industry for various processes. Most of the chemical syntheses and separation processes done in the laboratory or in the industry, as well as the cellular biochemical reactions in the living organisms, need solvents to proceed successfully and efficiently. Over the last three decades, focus has moved from using conventional organic solvents to development of non-conventional solvents such as Ionic Liquids (ILs). The usages of ILs have exploded in recent years; however, they have been around for long. In fact, one of the first ILs were described in 1914, when $[(EtNH_3)(NO_3)]$ – ethylammonium nitrate was synthesized (4). IL, as the name suggests, consist of an anionic (An^-) and a cationic (Cat^+) species, which is liquid at room condition, and is represented as $[(Cat^+)(An^-)]$ as it was shown by giving an example of ethylammonium nitrate compound (5).

More recently, novel category of non-traditional solvents has been developed, called as Deep Eutectic Solvents (DESs) (6). DESs are considered as ionic liquid analogous because of their certain similar physico-chemical properties, which we describe below. By definition, a *eutectic mixture* has melting point less than the melting points of its components (5). A *deep eutectic* term implies strong depression in the melting point of the mixture, represented on Figure 1.1.

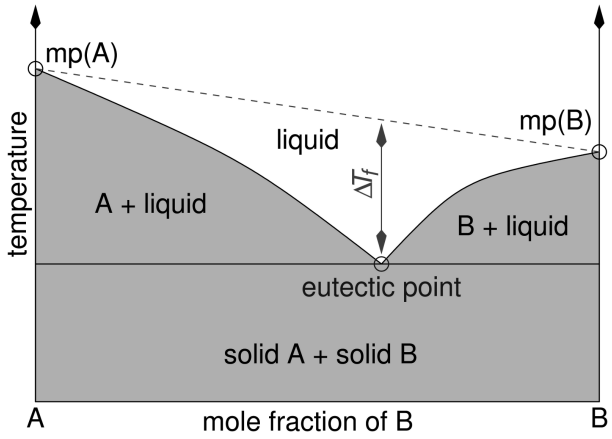

Figure 1.1. Formation of Deep Eutectic Solvent (5)

One of common examples of DESs is a mixture of choline chloride and urea (ChCl:urea) in 1:2 molar ratio, known as Reline solution. The initial melting points of the components are 303°C (ChCl) and 134°C (urea). Once they are mixed in 1:2 molar ratio, a DES with a eutectic point of 12°C is obtained (5). Such a unique behavior in the lowering of melting point of DES is generally caused by the presence of hydrogen bonding network prevalent in the mixture (7).

There are several advantages of using Deep Eutectic Solvents instead of Ionic Liquids. Most importantly, synthesis of DESs is simple as it requires only mixing the reagents. Furthermore, DESs are environmentally benign and are generally less expensive as compared to ILs. On the basis of favorable physicochemical properties and tunable components, several applications of DESs are being developed. For example, DESs are currently being used commercially in metal processing, metal dissolution and metal deposition areas by incorporating metal ions into the solvents (5).

One of the other possible applications of DESs is their use in the field of protein structure stability and protein refolding. This field of study is important in terms of drug design, cell cycle, and medicine (8). All of these are strongly related to the structure and unique folding of a particular protein, which should not undergo any changes for proper functioning. It is suggested that the factors, such as Hydrogen bonds, van der Waals interactions, electrostatic interactions, hydrophobicity, bond angles are responsible for keeping the protein in its native structure (9, 10).

In addition, a protein can be unfolded and lose its folded native structure during denaturation process due to changes in temperature, pH, pressure or other conditions (11). The reverse process of protein unfolding is known as protein refolding (12). This process is challengeable in the way of achieving proper and complete refolding of a protein back to its native structure in small amount of time.

While the experiments are important, the results of the experimental studies on the protein unfolding and refolding in different solvents such as DES, do not fully show the information about the interaction patterns between molecules, essential to understand the main concepts of protein stability on molecular level. To address this problem, we herein performed Molecular Dynamics (MD) simulations to look at the interactions between molecules, formation of Hydrogen bonds, and deviations in protein structure.

In what follows, we present a detailed literature review for protein structures, DES, and their applications in protein refolding. Furthermore, we discuss the simulations methodology, followed by our results from analyzing the native structure of protein, particularly Hen Egg Lysozyme, in reline and its aqueous solutions. In addition, we performed MD study on protein structure at high temperature, when protein unfolding occurs. Finally, we analyzed the rate and effectiveness of refolding of thermally unfolded lysozyme back to its compact structure in reline and its aqueous solutions. Nevertheless, all simulations were performed in 50 ns, so that the results of this study showed the initial patterns of molecular interactions involved in protein structure analysis.

Chapter 2 - Literature Review

2.1 Protein structure stability and refolding

Proteins are important chemicals found in living organisms. Proteins have unique structure and particular folding necessary for their activity (13). Most of the proteins are found in globular conformation necessary for the biological functions. The forces stabilizing the protein conformation are mainly hydrophobicity and hydrogen bonding (9). Protein structure stability and the rate of its denaturation depends on the various factors such as pH, pressure, temperature or the chemical composition of the environment in which the protein is kept (11). During denaturation, protein lose its native structure through unfolding because of a change in the peptide chain or loss of hydrogen bonds (Figure 2.1) (13).

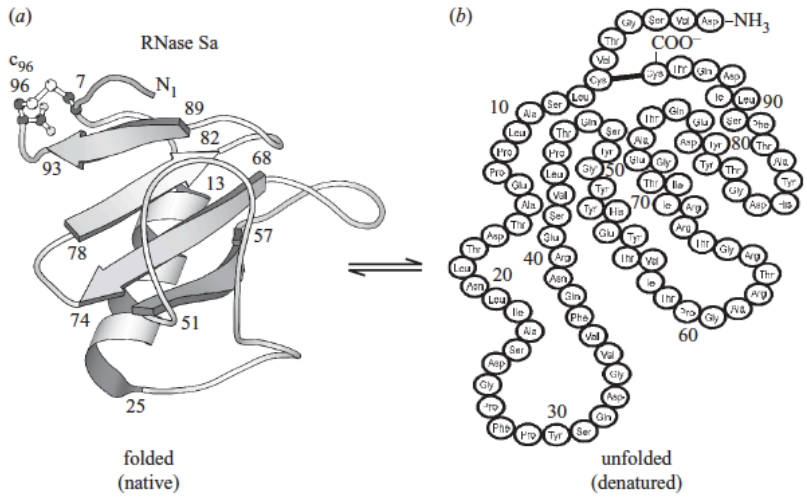

Figure 2.1 Denaturation of protein RNAse (13)

Nowadays, the number of commercially available products synthesized on the basis of proteins increases, which makes it more significant to study protein structure stability and instability (14). The stability of the protein structure is also important in terms of its influence on the mechanism of the cell cycle, involved in the cancer disease (8). Furthermore, the development of medicines containing proteins strongly depends on the chemical and physical stability of proteins, essential to indicate their biological, physical and chemical properties, including the factors preventing the protein denaturation and degradation of the drugs (15).

Refolding of protein structure, on the other hand, is an important field of study in terms of practical applications in genome research, protein design, and pharmaceuticals. In human genome, there are approximately 10^6 proteins, which need to have a stable folded structure in order to be biologically active. On molecular level, it is challengeable to understand how a polypeptide chain of aminoacids can obtain a three-dimensional structure. The refolding can happen spontaneously even for the proteins with broken disulfide bridges (12).

According to Dill and MacCallum (10), the main factors driving the folding of proteins are Hydrogen bonds, van der Waals interactions, backbone bond angles, chain entropy, as well as electrostatic and hydrophobic interactions. In computer modeling and simulations, these forces are indicated by “forcefield” parameters (10). The details of the protein refolding mechanism are not fully determined, so that the number of theoretical, computational and

experimental studies on protein stability and refolding increased for the last years (10).

Thus, the stability of the native globular structure and refolding of the protein is an important concept of its unique functionality and participation in chemical reactions in living cells. Up to date, many different solvents including DESs were used as an environment for protein structures to investigate the stability of proteins in different media that will be reviewed in the next sections.

2.2 Protein stability in organic solvents

As it was mentioned above, the structure of protein is important in its functioning. Several studies, both experimental and Molecular Dynamics simulations have been performed to analyze the structure of protein in different solvents under varying conditions (temperature, concentrations, etc.).

For example, 1-methyl-3-octyleimmidazolium chloride ([OMIM][Cl]) was used to study the stability of lysozyme and its activity by experimental and molecular dynamics approach (16). Different solutions with various amounts of [OMIM][Cl] were prepared and protein fluorescence intensity was measured, for each of them to understand the protein structure. It was observed that with increasing solvent concentration, the intensity decreased, indicating protein unfolding. The reason, author stated, is the formation of intermolecular hydrogen bonds in the presence of the solvent and losing of intramolecular bonds and secondary structure of protein. The results from the molecular

dynamics simulations performed in 20 ns (20000 ps) are shown on Figure 2.2 in terms of root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), and solvent accessible surface area (SASA). The purpose of this analysis was to compare the deviations in the native structure of pure protein in the presence of solvent and without it (16).

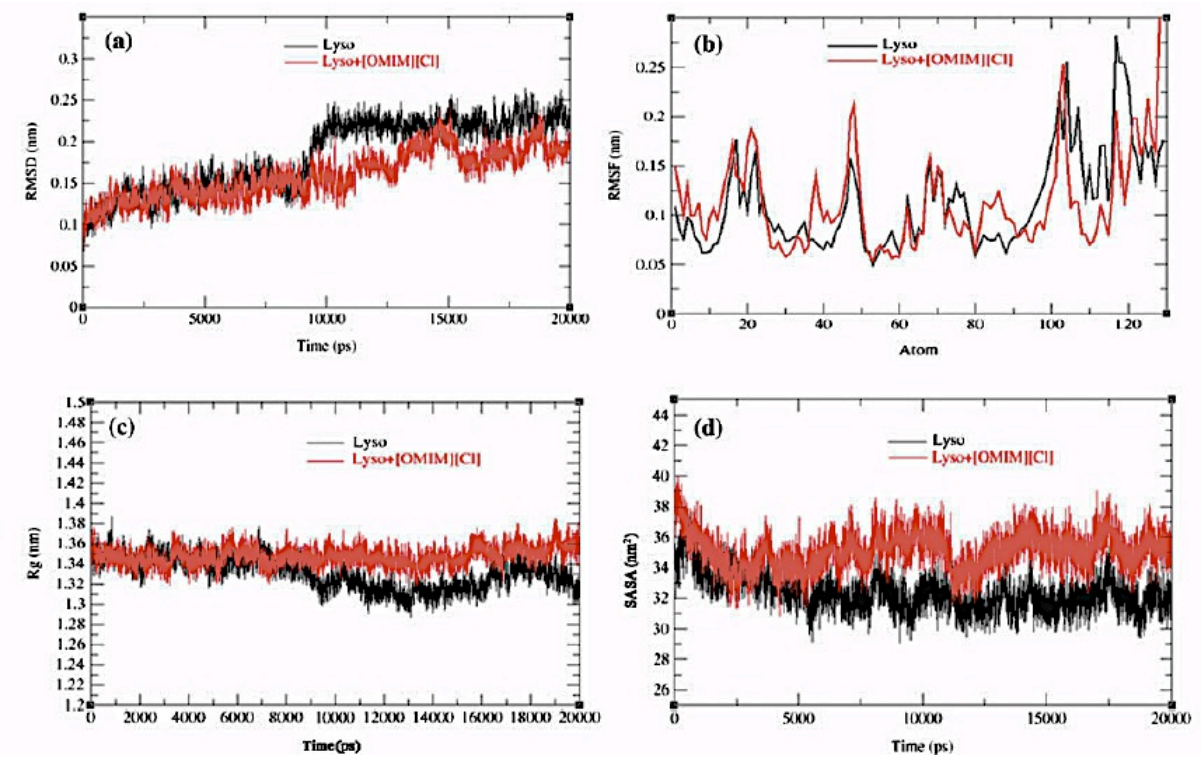

Figure 2.2. Molecular Dynamics simulations results of lysozyme stability in [OMIM][Cl]: a) RMSD, b) RMSF, c) Rg and d) SASA (16)

To be more detailed, according to the results of Figure 2.2, the values of RMSD of lysozyme in the solvent (0.1 – 0.22 nm) deviated in the range close to the RMSD of pure lysozyme (0.1 - 0.25 nm) with the major deviations after 10 ns of simulation in both cases. In addition, the differences observed between RMSD of the lysozyme structures indicate the binding of the solvent with the

protein. In addition, according to RMSF analysis, the fluctuations of 129 residues of the lysozyme during 20 ns are in the similar range (0.06 - 0.30 nm) with the major fluctuations of 10 residues, responsible for the binding of [OMIM][Cl] to the lysozyme. The results also showed the change of protein secondary structure in the solvent in terms of R_g and SASA, indicating the unfolding pattern in the presence of the solvent during 20 ns of the simulation. R_g of lysozyme deviated more significantly (1.36 ± 0.02 nm) in the presence of the solvent, in comparison to the pure protein (1.31 ± 0.02 nm) (16).

Moving to another example, in 2017, Sedov and Magsumov analyzed through molecular dynamics simulations thermal stability of lysozyme protein in the presence of water and dimethyl sulfoxide (DMSO) (11). Different solvent dilutions with water were used to study the effect of DMSO concentrations on thermal unfolding of the protein at 450 K. It was observed that, the organic solvent accelerates denaturation by exposing protein hydrophobic side chains to dimethyl sulfoxide, indicated by increased values of hydrophobic SASA (solvent accessible surface area) and RMSD (Root Mean Square Deviation) (Figure 2.3). Moreover, it was found that time required for unfolding was less for the solvent with higher concentration of DMSO (11).

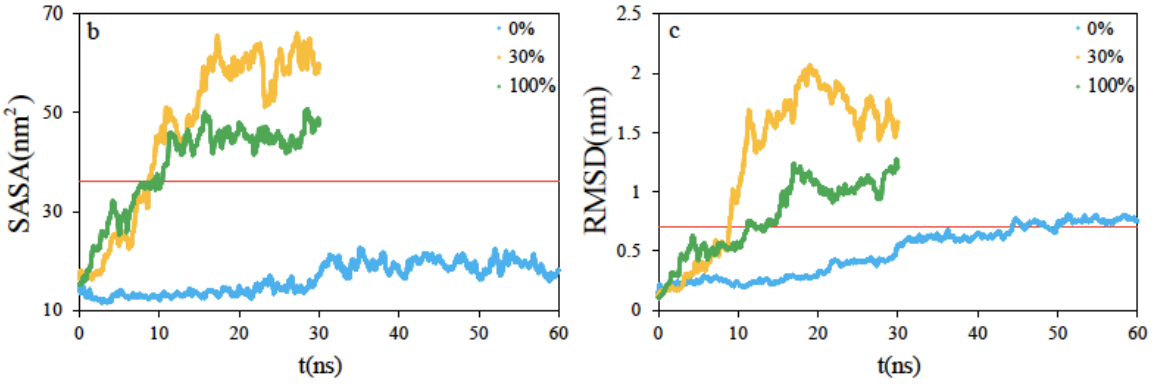

Figure 2.3 SASA and RMSD analysis of lysozyme in 0%, 30% and 100% DMSO solutions at 450 K (11)

According to Figure 2.3, in 30 ns of MD simulation, hydrophobic SASA values significantly increased in 30% and 100% DMSO solutions from 12 nm^2 to approximately 60 nm^2 and 50 nm^2, respectively, with minor changes of SASA in the presence of water in 60 ns of simulation. Similarly, RMSD analysis showed more deviations in the structure of protein in DMSO solutions, indicating enhanced unfolding pattern (11).

Another Molecular Dynamics study was performed on folding and thermal unfolding of protein Engrailed Homeodomain (En – HD) (17). During 60 ns of the simulation, the stability of the native structure of the protein was studied at 298 K and 498 K. In addition, partially unfolded structure obtained after 5 ns of thermal unfolding was used to study the protein refolding at 298 K. The results of this study are represented on Figure 2.4 in terms of C_α RMSD analysis.

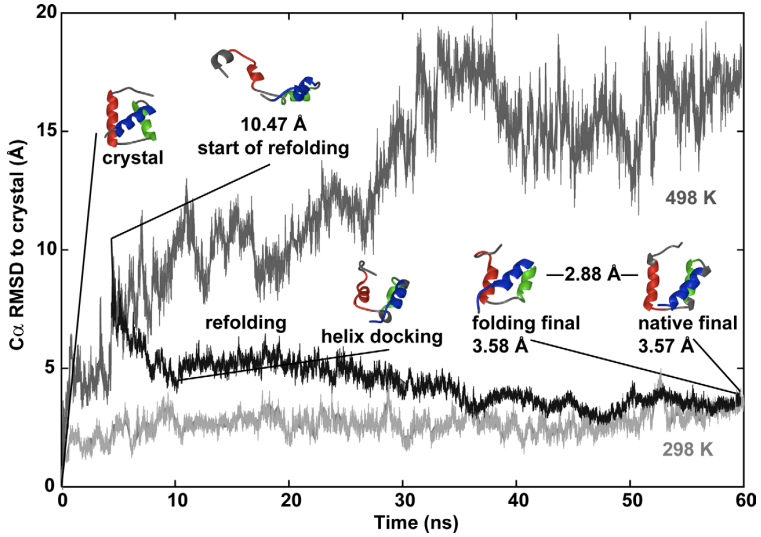

Figure 2.4 MD simulation of Engrailed Homeodomain protein (17)

According to the results of C_α RMDS analysis, represented on Figure 2.4, the protein has clear unfolding pattern at 498 K indicated by enhanced deviations on the plot. In addition, it was found that changing the temperature from 498 K to 298 K after 5 ns of MD simulation can bring partially unfolded En - HD protein to the structure similar to the protein native structure in 55 ns of simulation (17).

2.3. Classification and Preparation of Deep Eutectic Solvents (DESs)

As mentioned above, Deep Eutectic Solvents are mixtures of two or more components, with a strong depression in the melting point as compared with the individual components. The components of a particular DES are mixed in a specific ratio that results in the lowest final melting point of the mixture. DESs are generally represented as "Cat^+X^-zY", where "Cat^+" is a cation, "X^-" stands for a Lewis base, "Y" represents a Lewis or Bronsted acid and "z" is a number

of interacting “Y” molecules (5). In general, sulfonium, phosphonium and ammonium cations are used, whereas “X^-” is usually a halide anion. Deep Eutectic Solvents can be classified to several groups according to the complex agent “Y” they are made from (5, 18). In Type I DESs “Y” is a metal chloride (MCl_x). In Type II DESs metal halide hydrates ($MCl_x \times yH_2O$) are used as “Y”. Type III DESs contains alcohols, carboxylic acids, amides, etc. in their structures. DESs of Type IV are the eutectics composed by metal salts (5).

The preparation of DES is based on making a mixture of a HBD (Hydrogen Bond Donor) – Lewis or Bronsted acid and a HBA (Hydrogen Bond Acceptor) – Lewis or Bronsted base at a particular temperature (7). Two solid materials combined together finally produce a liquid solvent with eutectic properties. One of the first DESs was obtained by mixing choline chloride (ChCl) and urea in 1:2 ratio by Abbott et al. in 2001 (19). Both reactants, ChCl and urea, are solids at room temperature with the melting points 302°C and 133°C, respectively. The solvent produced from the mixing of these components was eutectic liquid at room temperature with the comparatively low melting point (12°C). Strong hydrogen bonds and weak van der Waals forces present in the obtained mixture prevent the starting materials from the solidification (7).

Commonly used HBA in DESs are organic compounds like choline chloride, carboxylic acids, glucose, betaine, and inorganic compounds like metal chlorides ($FeCl_3$ and $ZnCl_2$). Urea, succinic acid, glycerol, aminoacids,

glycols, and citric acid are typical HBD in DESs (7). In addition, Choi et al. proposed a new type of DES called NADESs (Natural Deep Eutectic Solvents) in 2011 (1). They synthesized 30 NADESs using different combinations of ChCl, sugars, water and naturally occurring carboxylic acid.

2.4 Physicochemical properties of DESs

It is very important to investigate the properties of DESs in order to explore their potential applications. To begin with, ILs and DESs have certain similarities and differences in their properties. According to definition, ILs are fluids with boiling point less than 100ºC consisting of ions only (5). Both ILs and DESs have similarities in their physical properties, such as low vapor pressure and nonflammability. The advantages of DESs over ILs are easy synthesis, availability and low cost of production. In addition, DESs have higher viscosity, but lower conductivity (5).

In 2016 Al-Omar et al., for example, have synthesized and investigated physical properties, including freezing point, surface tension, density, conductivity and viscosity of several different DESs (20). Specifically, glycerol molecule was chosen as the hydrogen bond donor in all DESs. Six different salts were used as hydrogen bond acceptors, which are choline chloride (ChCl), tetra-n-butylammonium bromide (TBAB), N,N-diethylethanolammonium chloride (DAC), and three phosphonium bromides (PB), each with a distinct radical; methyl triphenyl (MTPB), benzyl triphenyl (BTPB) and allyl triphenyl

(ATPB). During the experiment 70 different solvents were prepared by changing the molar ratios of a HBD and a particular HBA. The obtained mixtures were observed in different phases, including solid, crystal, semi–solid and liquid phases. In addition, some of the liquids were recrystallized after some time. The experiment showed that the eutectic solvents can be produced only at specific ratios of HBA and HBD (20). According to the results, DESs showed the freezing point values lower than the freezing points of the starting materials. Furthermore, densities of DESs decreased with increasing temperature, due to the increasing of molar volume, thus enhancing the activity and mobility of the molecules (20).

Furthermore, DESs generally have a high viscosity due to hydrogen bonds, which binds and immobilize the free species of the solvent (21). However, the viscosity decreases as temperature increases. The electrical conductivity of DES is generally low, due to their high viscosity. Consequently, increasing the temperature increased the DES conductivity. Finally, studying of DES surface tension properties showed that, in general, its values decreases as temperature increases due to the breaking of hydrogen bonds (21). Overall, the investigation of the physical properties of DES and their dependence on the temperature can help to understand the nature of these novel solvents, which further can assist in designing more DESs with tailored properties.

2.5 Applications of DESs

2.5.1 Using DESs in the desulfurization of fuels

One of the most important ecological issues we face these days is air pollution all around the world. One of the reasons of this environmental problem is the presence of organic sulfides in the fuel. Burning of the fuel with high amount of sulfur leads to the emissions of harmful gaseous sulfur oxides (SO_x) to the atmosphere and results in acid rains, corrosion of constructions, and degradation of catalysts (22). Deep Eutectic Solvents are turning up to be cheap, effective, and environmentally friendly methods to remove sulfur from fuels.

One of the first research was conducted in 2013 on desulfurization of fuel by means of DESs (23). Specifically, all of the synthesized and analyzed DES samples contained coordinated metal ions, which increased the efficiency of removing sulfur from fuels. Experimentally, the researchers came up to the conclusion that the efficiency of the process depends on different factors like operating temperature, sulfur concentration in the fuel, and mass ratio of fuels to DESs. It was found that increasing the ratio of the DES mass to the fuel mass increases the efficiency of the desulfurization. Nevertheless, the studied solvents showed efficient recyclability. Finally, it was suggested that the mechanism of the successful desulfurization process by DES is based on the interactions driven by hydrogen bonding, as well as coordination bonding of metals (24).

2.5.2 Using DESs for gas absorption

One of the ways to reduce the amount of the greenhouse gas is capturing CO_2 molecules, as they are being produced in the power plants (25). The traditional way of using aqueous amine-based solutions as an absorbent has the disadvantages like high cost, consumption of energy, corrosion of the equipment and degradation of amine (26).

Deep Eutectic Solvents have been proposed as an alternative solvent, which is recyclable, biodegradable, sustainable, and efficient for CO_2 capturing (26). It was found that the composition of DESs must be adjusted by using different cations and anion in order to increase the affinity for the gas absorption. Moreover, it is important to consider pressure, temperature, gas solubility in DES, and the selectivity of the absorbent for successful gas capturing (26).

2.5.3 Using DESs in cellular metabolism

According to Choi Y. H. and co – workers, living organisms have Natural DESs (NADESs) inside their cells (1). It is suggested, that different combinations of compounds like water, citric acid, choline chloride, malic acid, maleic acid, aconite acid and other chemicals found inside a cell can result in the formation of NADES. Experimentally, it was investigated that some of the water – soluble and water – insoluble natural products have higher solubility in NADES rather than in water. For instance, rutin (a citrus flavonoid, plant

metabolite) has the solubility in the tested NADES for 50 - 100 times higher than the value of its solubility in water. This group of researchers hypothesize, that NADESs can play an essential role in the biochemical reactions inside the cells and can serve as a third type of cellular solvents, in addition to two other known liquids like water and lipids (1).

The topic is still opened for further investigations and more computational modeling as well as experimental data must be collected to make final statement. Consequently, this theory can help to solve different problems in understanding the biochemical reactions inside the cell in the near future.

2.5.4 Using DESs in extraction and separation techniques

There are several other applications of Deep Eutectic Solvents in analytical chemistry. For example, DESs can be used in (micro)extractions to improve the extraction efficiency of nanoparticles and different sorbents by their modifications, dissolution of solids, and changing the mobile phase used in the chromatography techniques (27). In addition, NADESs can be used as extraction solvents due to the “donor - acceptor mechanism” of hydrogen bond, which consequently favors the processes of the dissolution (7). Deep Eutectic Solvents also can be implemented in the purification of biofuels.

For example, in 2010, Shahbaz and co-workers described the ability of DESs to successfully remove glycerol from the synthesized biodiesel (28). The research team optimized the composition of the solvents (ChCl:trifluracetamide,

ChCl:ethylene glycol) in order to maximize the separation of glycerol from the biodiesel by liquid – liquid extraction technique based on the solvation properties of the glycerol molecules in DES. According to the analysis, as low as 0.199 wt% glycerol in the biodiesel sample was achieved by using DESs.

2.5.5 Using DES as a catalyst

Moving to other applications of DESs, they can be used to catalyze chemical reactions, such as alcoholysis, aminolysis and hydrolysis. The advantages of using DES as a catalyst are its better solvation properties due to hydrogen bonding, thermal stability, low volatility, low cost, accessibility, easy synthesis in one step reactions and biodegradability. Certainly, the main advantage of using DES is determined by its eutectic property of having the melting point lower than the melting points of the compounds combined to make a particular Deep Eutectic Solvent. This gives an opportunity to use DESs under milder conditions (29).

DESs have also been used as solvents for enhancing activities of biochemical reactions. The research conducted by Durand et al., in particular, studied the ability of enzymes to be dissolved and exhibit their activity in different DESs, such as ChCl:Urea and ChCl:Glycerol (30). Transesterification reaction between ethyl valerate and 2 – butanol at 60°C catalyzed by lipase in these solvents showed the effectiveness same with the results obtained from the reaction in toluene as a solvent. Furthermore, the aminolysis reaction between

ethyl valerate and butylamine in the presence of the enzyme also resulted in the similar conversion in DESs, as in toluene. It is also important to note that the lipases are more stable (only 1% loss of activity) in DESs than in the presence of its component urea, where the enzyme lost 75% of the catalytic activity. It was suggested, that the reason of the enhanced stability of DES is the network of hydrogen bonds, which keeps the components of the solvent together and decrease their reactivity and prevent the degradation of the enzyme. Consequently, it was pointed out that further investigations are necessary to study the possible applications and modifications of DESs for the catalytic reactions with the substrates capable of breaking the hydrogen bonds network in the solvent (30).

2.5.6 Protein conformation in DESs

Finally, protein stability, thermal stability and refolding of lysozyme in Deep Eutectic Solvents were studied experimentally by Esquembre et al. in 2013 (3). Two different DESs were synthesized by using choline chloride as hydrogen bond acceptor and glycerol or urea as hydrogen bond donor. Fluorescence spectroscopy and circular dichroism (CD) measurements were used to study the protein stability (Figure 2.5), possible refolding (Figure 2.6), as well as to check the activity of lysozyme in pure and aqueous DESs (86, 43, 10 wt% DES).

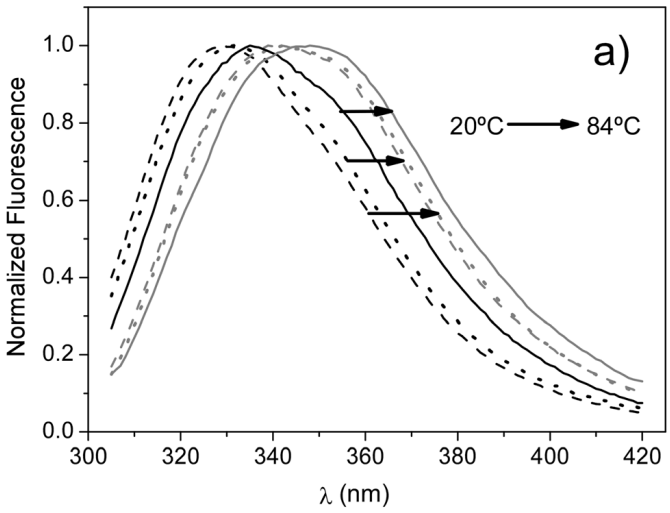

Figure 2.5. Fluorescence emission spectra of tryptophan of lysozyme at (i) 20ºC and (ii) 84ºC, in buffered solution (solid line), Urea-ChCl DES (dotted line) and Glycerol – ChCl DES (dashed line) (3)

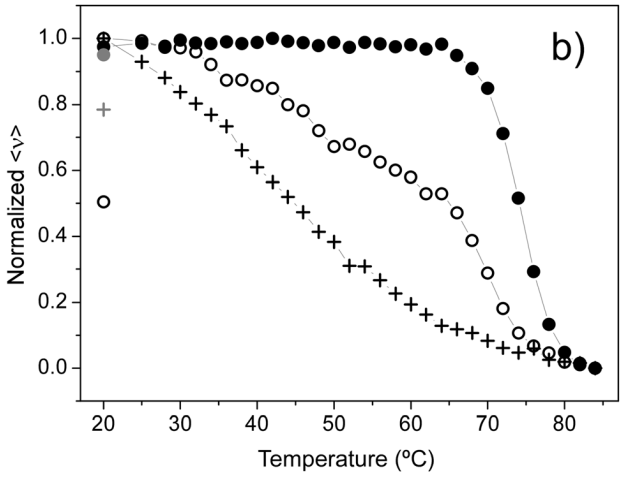

Figure 2.6 Mean fluorescence energy of the protein in buffered solution (•), Urea-ChCl DES (+) and Glycerol – ChCl DES (o); after heat treatment (black) and at 20ºC (grey) (3)

On the basis of the results represented on Figure 2.5, it was found that at room temperature there are minimal differences in the stability of lysozyme in buffer solution and DESs. However, after thermal treatment at 84 ºC, the changes in the structure of a protein changed the fluorescence patterns of two tryptophan residues (Trp 62 and Trp 108) of lysozyme. Moreover, study of the

protein folding in the presence of DESs after thermal unfolding showed the partial refolding through formation of folding intermediates (Figure 2.6). Furthermore, it was found that more diluted DESs showed thermal unfolding pattern more close to the unfolding in buffer (3).

In addition, the activity of the protein was studied by looking at the rate of hydrolysis of 1, 4 – β linkages present in the cell walls of Gram – positive bacteria. It was found that the activity of lysozyme is lower in DESs in comparison to buffered solution environment (Figure 2.7). Finally, the authors suggested studying particular interactions between protein and DES molecules before their usage in the biocatalytic reactions (3).

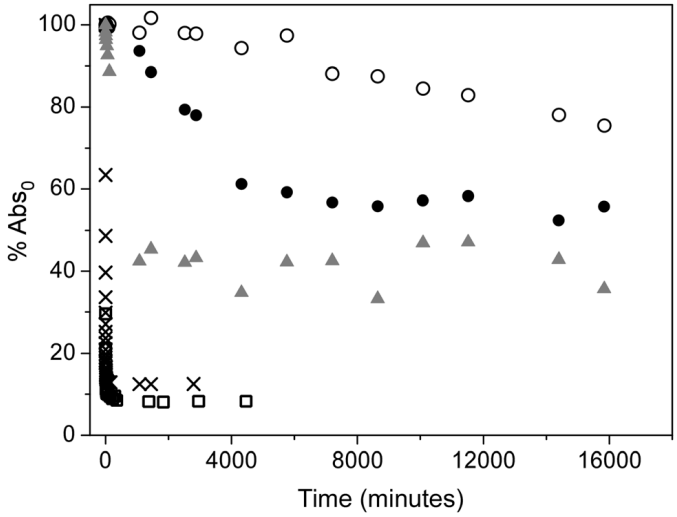

Figure 2.7 Optical density plot in buffered solution (□), Glycerol – ChCl DES (○), 80 wt% DES (•), 60 wt% DES (▲) and 10% wt DES (×) (3)

Another experimental work was done by Sanchez – Fernandez et al. in 2017 using pure DESs and their hydrated mixtures to study the stabilities of two

different proteins: bovine serum albumin (BSA) and lysozyme (2). DESs used to study BSA were the aqueous mixtures of choline chloride (ChCl) and glycerol. And for lysozyme study, aqueous mixtures of both DESs, ChCl:glycerol and ChCl:urea were used. The analysis of protein stability was performed by using CD. Figure 2.8 represents the experimental results from the studies of BSA and lysozyme stabilities (2).

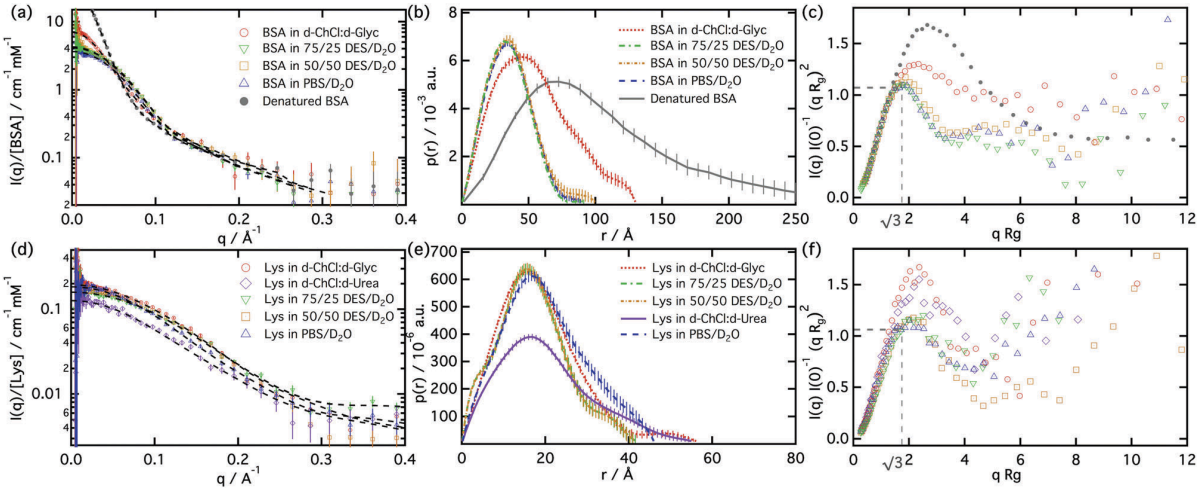

Figure 2.8. Experimental results of BSA (a, b, c) and lysozyme (d, e, f) stabilities in pure and hydrated DESs: a) and d) SANS data, b) and e) pair distance distribution, c) and f) Normalized Kratky plot (2)

According to the results, the secondary structure of BSA did not change too much in different solvents of pure DES and its aqueous mixtures, while there were small changes in its tertiary structure in the hydrated DESs (2). In addition, lysozyme showed partial folding in pure DES, but, maintained stability in hydrated DESs. However, the solvation mechanism that directly affects the stabilities of these proteins remains unclear.

A molecular dynamics simulation was performed by Monhemi et al. to study the stability of Candida Antarctica Lipase B (CALB) in DES of urea and choline chloride (31). RMSD, Rg, and RMSF were used to analyze the stability of initial structure of CALB based on C_α and simulated structure of enzyme. On the basis of the results of this study, it was shown that there are strong H-bond interactions between urea molecules and Cl-atoms of DES. Moreover, RMSD and Rg values of CALB remained constant in DES, but increased in the presence of 8 M urea at 300 K after 50 ns of simulation (Figure 2.9). This shows the structural stability of enzyme in DES. In addition, it was found that there is no diffusion of urea molecules of DES into the enzyme, and hence no enzyme denaturation (31).

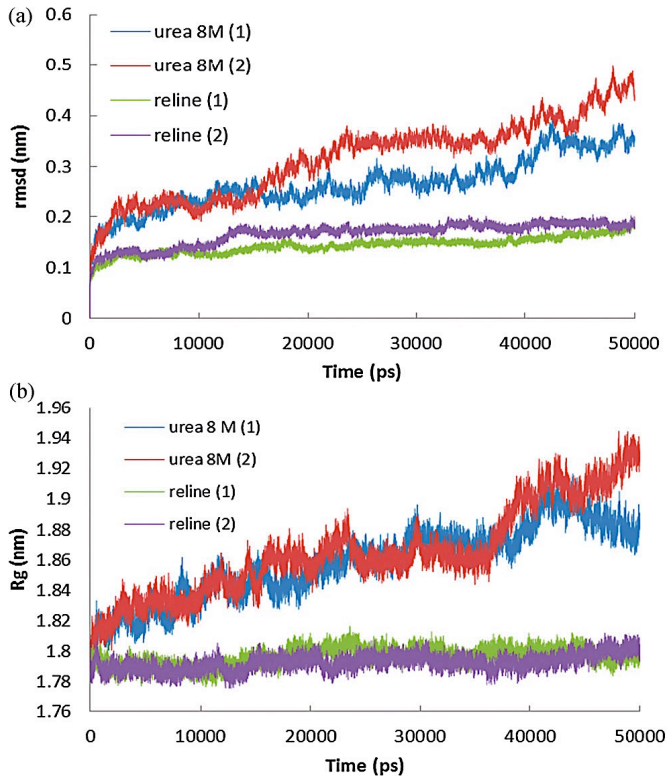

Figure 2.9. MD simulations results of CALB stability in reline and urea at 300 K (31)

2.6 Conclusion and Objective of this work

To sum up, on the basis of the performed literature review, it was concluded that the importance of the researches done on protein unfolding and refolding increases in the last years due to their application in protein design, drug design and medicine. So that, taking in consideration the importance of the protein stability, as well as promising usage of DES as a new solvent, it was suggested to explore the applications of DESs in the field of protein stability. Furthermore, it was decided to make a research based on Molecular Dynamics (MD) simulations to look more deeply to the solvation mechanisms and to the interaction patterns between protein and the solvent components.

Consequently, the main objectives of this work were to perform Molecular Dynamics study on the application of DES and its aqueous solutions in analyzing (i) protein structure at room temperature, (ii) protein structure at high temperature, when unfolding occurs and (iii) refolding patterns of thermally unfolded protein structure at room temperature. To be more detailed, in this study, Hen Egg Lysozyme was chosen as a protein and a mixture of Choline Chloride and urea in 1:2 molar ratio, known as Reline was chosen as DES.

Chapter 3 - Methodology

3.1 Molecular Dynamics Simulation

Molecular Dynamics (MD) Simulation is a methodology used to study molecular movement and interactions via computational approach. The global algorithm of MD consists of several steps. The first step involves creating a box and randomly placing the desired number of molecules in the system. After that, the forces applied on atoms due to bonded, non-bonded and columbic interactions are computed. Furthermore, atomic movements are simulated on the basis of Newton's law of motion considering the interactions between atoms (32).

$$\nu_i = \frac{dr_i}{dt} \qquad (3.1)$$

$$\frac{d\nu_i}{dt} = \frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i} \qquad (3.2)$$

3.2 GROMACS Tool

During this study GROMACS 2018.1 package was used for molecular dynamics simulations. Hen Egg Lysozyme was chosen as a protein to investigate its stability in DES media. The structure of the protein (hen-egg lysozyme, 1AKI.pdb) was downloaded from the Protein Data Bank (PDB) (33).

A mixture of Choline chloride (ChCl) and urea in 1:2 molar ratios, known as Reline, was chosen as DES, and later water with TIP4P model was chosen for making aqueous reline samples. The forcefield parameters and structures of

the mentioned molecules were taken from gromos54a7 database and Automated Topology Builder databases. The forcefield was validated by measuring the density of pure reline solution and comparing it with experimental data (34).

The simulation was started making a low-density box with the dimensions of 15 × 15 × 15 nm^3 and inserting a particular number of molecules for each system under study according to the following table (Table 3.1) along with 8 Cl^- anions to neutralize the positive charge of a protein.

Table 3.1: Details of different systems simulated

№	Lysozyme molecule	ChCl molecules	Urea molecules	Water molecules	Description (wt%)	Lysozyme concentration
1	1	250	500	0	Reline	15.7 mM
2	1	250	500	401	10% water	14.4 mM
3	1	250	500	1 546	30% water	11.2 mM
4	1	250	500	3 608	50% water	7.92 mM
5	1	250	500	8 418	70% water	4.7 mM
6	1	250	500	32 468	90% water	1.5 mM
7	1	0	0	3 566	100% water	13.3 mM
8	1	0	2 040	10 668	39% Urea (8M Urea)	3.4 mM

To be more detailed, according to Table 3.1, 401 water molecules were added to 250 reline molecules, consisting of 250 ChCl and 500 Urea molecules, to make 10% water dilution by weight. Likewise, 1546 molecules of water were added to reline solution to make 30% dilution, 3608 water molecules - for 50% diluted sample, 8418 water molecules - for 70% ww water and 32468 water molecules – for 90% ww water dilution. In addition, 3566 molecules of water were used to make 100% water sample. Furthermore, 8M Urea sample was prepared by mixing 2040 urea and 10668 water molecules. As it was mentioned, in all cases under study one molecule of lysozyme was introduced to the simulated environment with extra 8 chloride ions to neutralize the positive charge of the protein. Depending on the dilution, the lysozyme concentration in the simulated boxes varied from 1.5 mM to 15.7 mM. The snapshot of an example of the box with protein and solvent is represented on Figure 3.1.

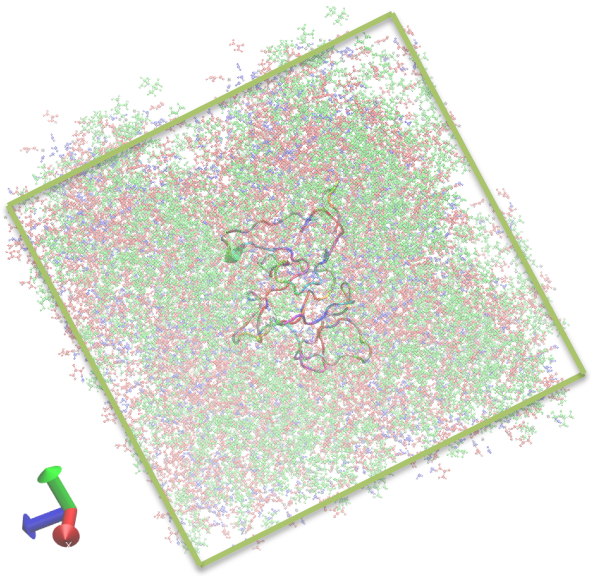

Figure 3.1 Thermally unfolded protein structure in reline

After filling the box with protein molecule, 250 reline molecules and specific number of water molecules, energy minimization was performed, with a constraint of maximum force to be less than 1000 kJ/mol/nm on any atom, in order to optimize the starting configuration of the simulation box. Following energy minimization, NVT and NPT equilibrations were conducted with temperature of 298 K and pressure of 1 bar, for 0.1 ns each. Finally, production run for each of the systems was performed for 50 ns (50000 ps). A large time was particularly chosen to analyze the dynamics of protein within the solvent.

During the simulations, all bonds were constraint with LINCS algorithm. A cut-off of 1.0 nm was used for Coulomb and LJ short-range interactions. Particle Mesh Ewald (PME) method was applied for long-range electrostatics with fourier spacing of 0.16 nm. In order to keep constant temperature at 298 K modified Berendsen pressure-coupling method was applied. In addition, Parrinello–Rahman type pressure coupling was used at reference pressure 1 bar. Furthermore, periodic boundary conditions in all directions were applied for all MD simulations.

In what follows, to visualize and study the simulations results Visual Molecular Dynamics (VMD) tool was used, in addition to the analysis accomplished via gromacs package. To be more detailed, Solvent Accessible Surface Area (SASA) analysis available on gromacs was used to study the deviations in the structure of lysozyme in different environments by evaluating all atoms of 129 amino acid residues of protein accessible for the interactions

with solvent. For SASA analysis, the solvent probe radius of 0.1 nm was used. Similarly, Radius of Gyration (RoG) analysis available on gromacs was performed to look at the gyration of protein radius in different solvents evaluating all atoms of protein. In addition, Radial Distribution Function (RDF) of all atoms distribution was used to study the interaction patterns between protein and solvent components via VMD.

A series of simulations were performed to analyze the structural stability of the protein in different solvents at room temperature. After that, the similar procedure of MD simulation was done to check the thermal stability of the protein via changing the temperature in equilibration steps from 298 K to 498 K keeping all other parameters unchanged, including the compositions of the simulated environment in the boxes, mentioned in Table 3.1.

Furthermore, the unfolded structure of lysozyme produced from its thermal denaturation in water at 498 K was taken after 10 ns of simulation to study the protein refolding patterns at 298 K in reline and its aqueous solutions. Similarly to the previous parts of the study, all of the other parameters of the simulation, as well as the solvent compositions mentioned in Table 3.1 were kept the same changing only the initially taken structure of the protein from initially folded to already unfolded one.

Chapter 4 - Results and Discussion

4.1 Structure of the reline

In the previous work of our group, the interactions between the components of reline were analyzed from the prospective of hydrogen bonds formation and radial distribution function (RDF) (34). The number of hydrogen bonds formed between urea molecules in reline was compared to the number of hydrogen bonds formed between the molecules in pure urea.

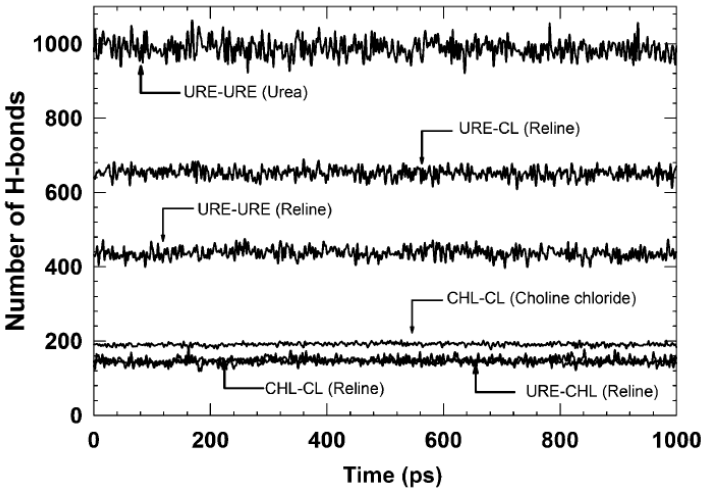

Figure 4.1 Number of Hydrogen - bonds in Reline and in pure components (34)

According to the results, represented on Figure 4.1, the number of hydrogen bonds between urea molecules decreased in reline solution. Furthermore, it was found that in reline, urea molecules interact more strongly with chloride ions than with urea and choline molecules. In addition, the number of H-bonds between choline and chloride ions is decreased in reline, in comparison to pure choline chloride solution. The rdf analysis, used to study the

probability of finding a molecule at a particular distance also confirmed the decrease of urea – urea interactions in reline and showed strong interactions between urea and chloride in this DES (34).

Furthermore, two different systems of reline and slightly diluted reline (10% ww water) was explored from the prospective of RDF analysis of the interactions between urea molecules (Figure 4.2), as well as the interactions between urea and chloride ions (Figure 4.3) present in each system.

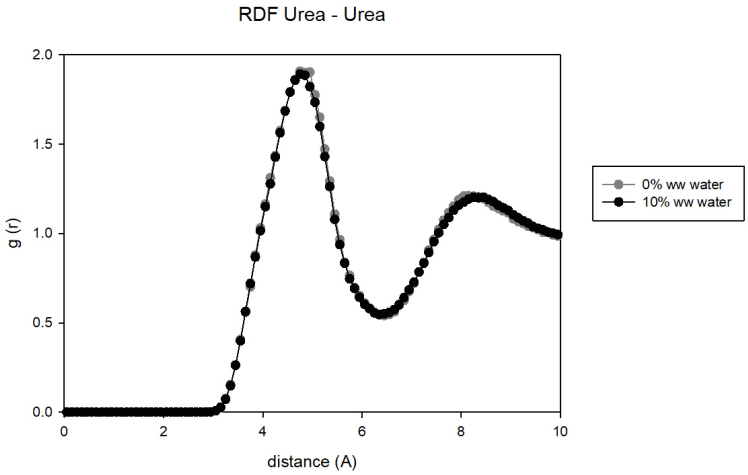

Figure 4.2 RDF analysis of urea – urea interactions

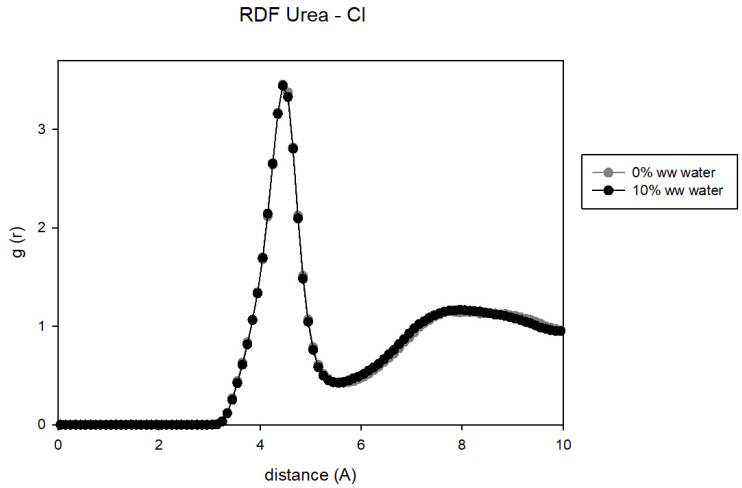

Figure 4.3. RDF analysis of urea – Cl interactions

According to the results of RDF analysis represented on Figure 4.2 and Figure 4.3, the interactions between urea – urea and urea - chloride did not change after the addition of water to reline solution.

4.2 Protein structure in reline and its aqueous solutions at 298 K

4.2.1 RoG and SASA

Moving to the first objective of this work, the structure of native lysozyme in pure water, reline and its aqueous solutions was studied by analyzing radius of gyration (RoG, Figure 4.2) and solvent accessible surface area (SASA, Figure 4.3) of protein over 50 ns. For reference, the values of RoG and SASA of the native structure of lysozyme are 1.4 nm and 73 nm^2, respectively, which changes quickly with first ns of the simulations.

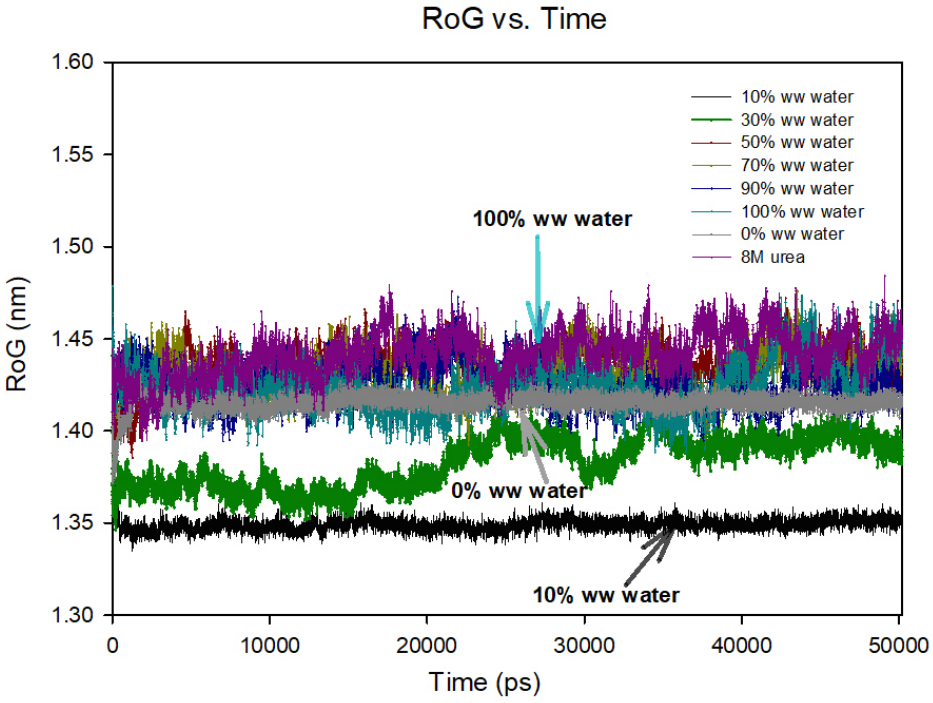

Figure 4.4. RoG of protein in pure water, pure reline and aqueous reline solutions

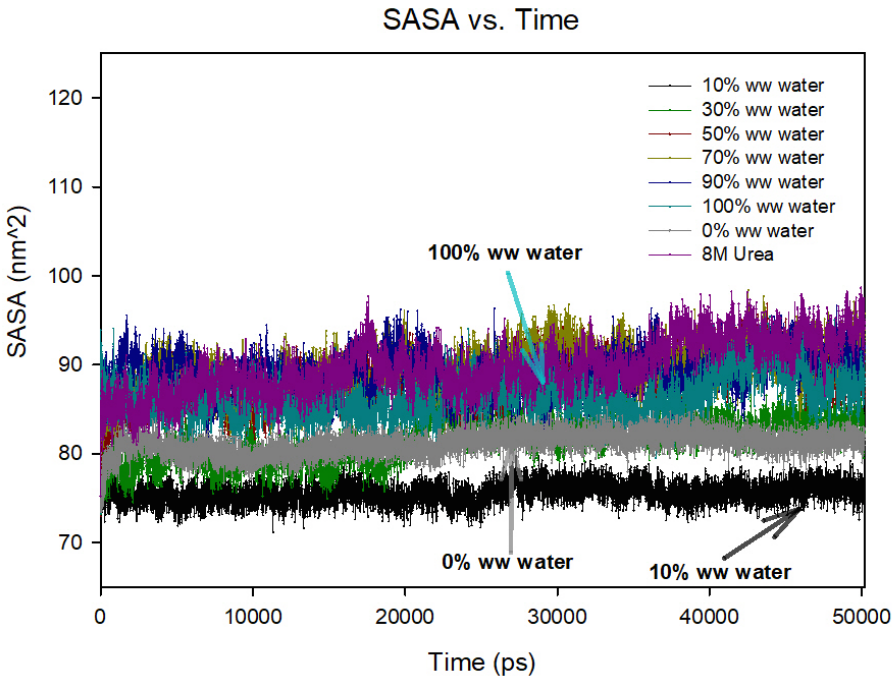

Figure 4.5. SASA of protein in pure water, pure reline and aqueous reline solutions

We begin with analysis of the simulation on the protein structure in water. The results from the simulations of protein in water, act as a control to analyze and discuss the structure of the protein in reline and its aqueous solutions. According to the results represented on Figure 4.4, the average value of RoG in pure water (100% ww water; cyan) is 1.44 nm. The presence of fluctuations on the plot indicates dynamics of protein structure during the simulation. Furthermore, in pure reline (0% ww water; grey), there are smaller fluctuations in RoG values around 1.41 nm, which stands for more stable structure. In addition, in slightly diluted reline (10% ww water; black), the protein is more compact with RoG values around 1.35 nm. With further addition of water to the reline solution, the stability of protein decreases and we observe high fluctuations, as well. In particular, maximum RoG was observed at

1.46 nm in the most diluted reline (90% ww water). Taken together, the results indicate that the structure of the protein remains compact in pure water, pure reline, and in aqueous reline solutions containing 10% water, whereas the protein expands slightly in aqueous reline solution containing high water amount.

The results of protein stability were also confirmed by SASA analysis, represented on Figure 4.5. A low value of SASA shows that protein has low surface area accessible for solvent due to its compact shape. According to the results, in pure water, SASA of protein fluctuates from 73 nm^2 to 88 nm^2, and becomes stable after 40 ns of the simulation. In addition, in pure reline and slightly diluted reline (10% ww water) there are smaller fluctuations in SASA values, with an average of 80 nm^2 and 75 nm^2, respectively. As noted above, more compact structure of lysozyme is observed in the aqueous reline solution containing 10% water by weight. Furthermore, with addition of water, SASA values show high fluctuations reaching maximum 95 nm^2 in more diluted reline solutions.

In addition to analyzing the structure of the protein in water, reline and its aqueous solutions, MD simulation was performed in the solution of 8 M urea, which commonly used to unfold the protein. In this case, the value of RoG increases from 1.4 nm to 1.46 nm in 50 ns (Figure 4.4) and likewise SASA also increases continuously indicating an unfolding pattern (Figure 4.5). The similar unfolding pattern of protein in urea was obtained by MD studies performed by

Rocco et al., which found that the unfolding of protein is more significant in the solution of 10 M urea than in water at any temperatures (35). In addition, in the study of Eleftheriou et al., it was indicated that 1000 ns simulation time is not enough for full denaturation of lysozyme in 8M urea (36).

The snapshots of protein structures in pure water, pure reline, slightly diluted reline (10% ww water) and 8M urea solutions obtained from our studies are represented on Figure 4.6.

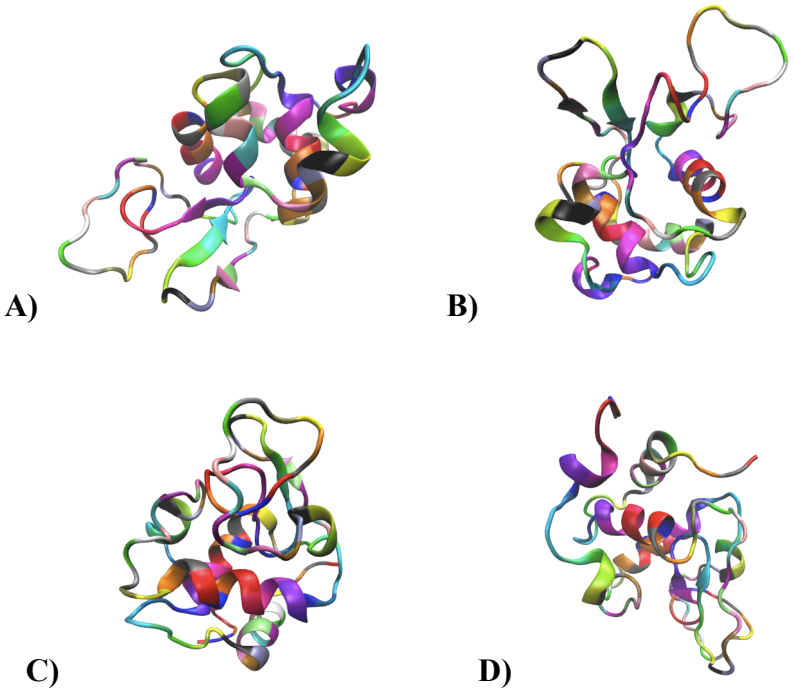

Figure 4.6. Native structure of lysozyme after 50 ns simulation at room temperature **in: (A) (A) pure water, (B) pure reline, (C) 10% ww water system and (D) 8M urea**

4.2.2 RMSF of protein active site

Next, RMSF (Root Mean Square Fluctuations) of lysozyme active site (Glu 35 - Asp 52) were studied to determine the fluctuations in the activity of the protein. The results are represented on Figure 4.7.

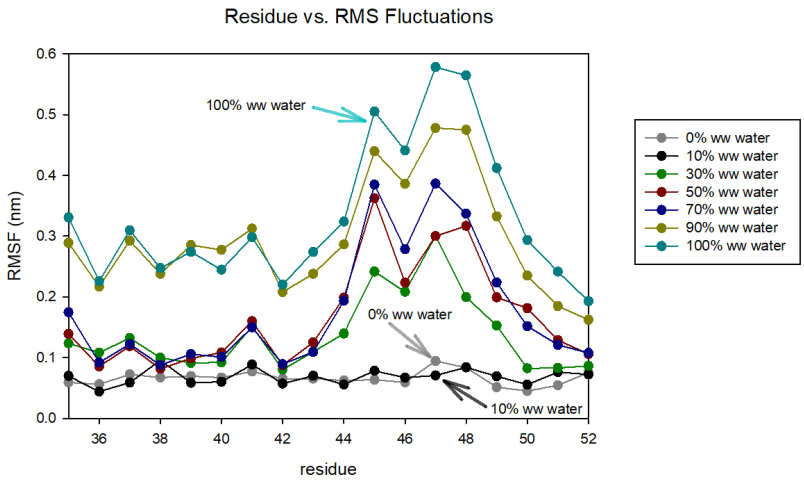

Figure 4.7. RMSF of protein active site in pure water, pure reline and aqueous reline solutions

On the basis of the results of Figure 4.7, protein active site has negligible fluctuations with the values of less than 0.1 nm in pure reline and slightly diluted reline (10% ww water), indicating the stability of protein active site. In comparison, in more diluted reline solutions and in pure water, the fluctuations, particularly from 45th to 48th residue increase more significantly reaching RMSF in the range of 0.3 – 0.6 nm.

4.2.3 Hydrogen Bonds analysis

The secondary and tertiary structures of a protein strongly depend on hydrogen bonds existing between amino acids. Following this, the protein structure stability was studied in terms of H-bonds formed between the components. In addition to the hydrogen bonds within the protein, the number of H-bonds between protein and solvent were also computed. Table 4.1

represents, the number of hydrogen bonds present in pure water, pure reline, 10% aqueous reline and 8M urea.

Table 4.1: Hydrogen bonds formation

Description	Protein - Protein	Protein – Non-protein	Protein - Urea
Pure reline	71±9	30±4	30±4
10% ww water	65±9	77±10	11±3
Pure water	51±10	275±62	0
8 M Urea	52±7	307±58	26±9

According to Table 4.1, the number of hydrogen bonds between protein – protein are slightly higher in pure reline (71±9) and in 10% water system (65±9), in comparison to pure water (51±10) and 8 M urea (52±7). The results are in-line with the compact structure of the protein, as we observed above in terms of RoG and SASA.

In pure water, the number of hydrogen bonds between protein and nonprotein (water) molecules is high (275±62), indicating strong interactions between protein and water. Whereas, in pure reline, there are a few hydrogen bonds between protein and nonprotein molecules (30±4). Interestingly, all of protein-nonprotein hydrogen bonds are formed with urea and none with other molecules like Cl or choline ions. Moreover, the number of protein – urea hydrogen bonds in pure reline is higher than in 8M-urea environment (26±9). However, we propose that due to lack of solvent molecules, the protein tends to

remain in the native state in pure reline, as compared to 8 M Urea solution. With addition of water, for example, in 10% diluted reline the number of hydrogen bonds between protein and urea molecules came down to 11±3, indicating lesser tendency to unfold in 10% aqueous reline solutions.

4.2.4. RDF analysis

Furthermore, the interactions between protein and solvent molecules were studied from the prospective of radial distribution function (rdf) analysis. The rdfs shows the probability of finding a particular component at a distance (we show up to maximum 15Å from the protein) (Figure 4.8).

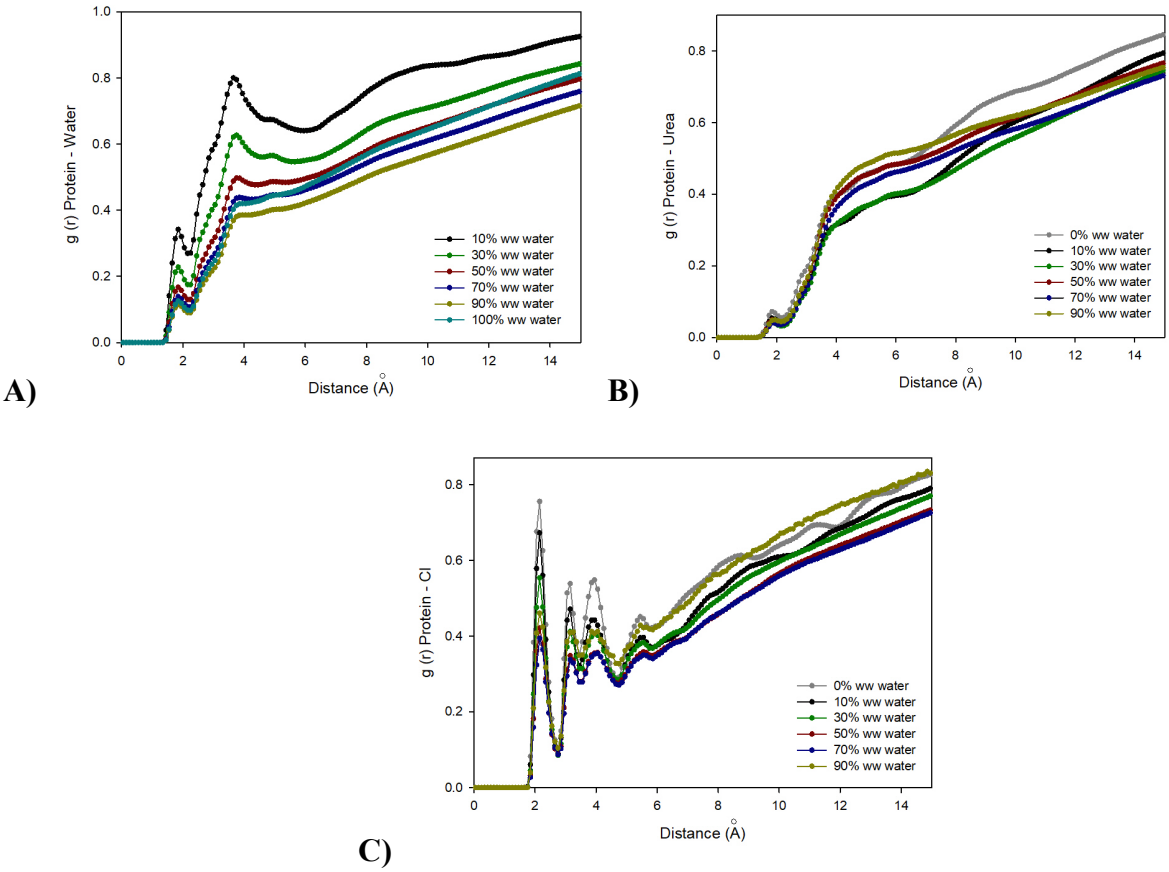

Figure 4.8. rdf between protein – solvent components in different systems: (A) protein-water, (B) protein-urea and (C) protein-Cl

The rdfs between protein and water are shown in Figure 4.8A for all different systems. Interestingly, although the numbers of water molecules are less in aqueous reline solutions, we observe stronger protein-water interactions in them, as compared to pure water. The results indicate that in the presence of reline, protein-water interactions are strong.

Figure 4.8B shows rdf between protein and urea for all the systems. Although the peak height is small, first of all, the results indicate that protein and urea molecules are close to each other (<2Å). Secondly, the interactions become stronger from slightly diluted reline (10% ww water) to highly diluted reline (90% ww water).

Furthermore, rdf between protein and chloride ions (Figure 4.8C) shows that in general the interactions between protein and Cl are relatively strong (in the range from 0.4 to 0.78 at 2Å) in comparison to the interactions between protein and other nonprotein molecules (<0.4 and <0.1 at 2Å for protein – water and protein – urea, respectively). The main reasons for this are: (i) the overall positive charge (+8) of protein, that attract chloride anions and (ii) interactions between protein and urea molecules, which in turn, because of high interactions between urea and Cl in reline, brings Cl ions closer to the protein.

Moreover, as the number of water molecules increases, the probability of finding chloride ion at the distance of 2Å from the protein decreases from 0.78 (pure reline) to 0.4 (very diluted reline), suggesting that the interaction between protein and chloride ions decreases. The observed trend can be attributed to the

fact that, with increasing water amount, the chloride ions are solvated by water, which leads to a decrease in protein-Cl interactions.

Overall, taking in the consideration the results of the analysis performed in this study, it can be concluded that protein structure is more stable and compact in pure reline (0% ww water) and slightly diluted reline (10% ww water) in comparison to pure water environment. The reason is that in pure water, water molecules allow protein to move more freely, in comparison to pure reline, which is more viscous and keep the structure of protein more compact and stable.

Furthermore, the protein is more compact in 10% ww water DES system, as in this case; firstly, water present in the aqueous solution surrounds the protein and decreases protein – urea interactions. Reline that finally surrounds the water in this system still helps protein to keep its structure more compact due to more viscous environment of DES.

4.3 Protein Structure at high temperature

Moving to the next objective of this work, the native structure of lysozyme in pure water, pure reline and aqueous reline solutions at high temperature, when protein usually unfolds, was analyzed in terms of RoG (Figure 4.9) and SASA (Figure 4.10).

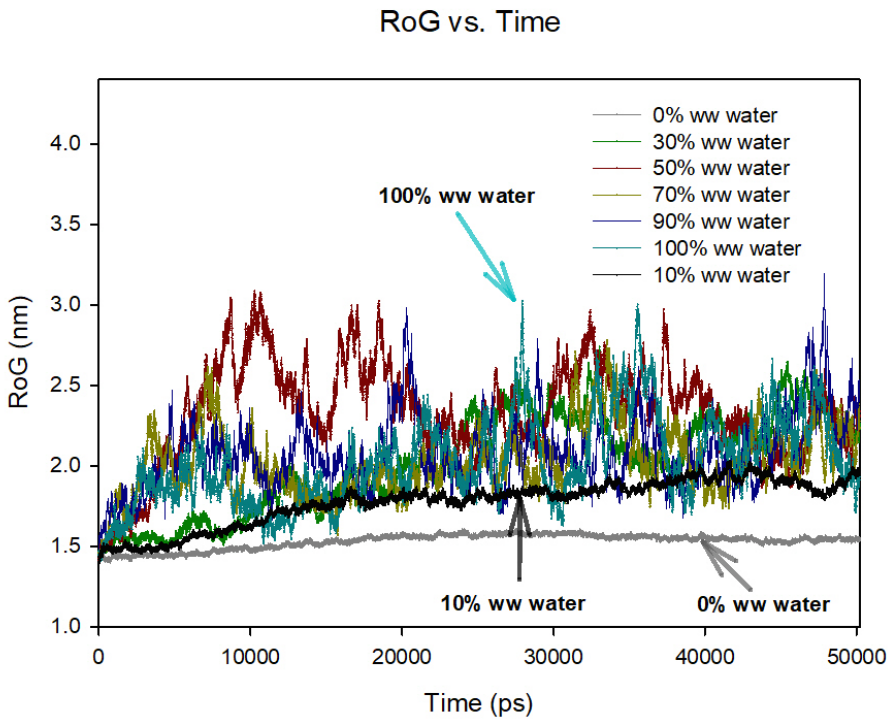

Figure 4.9. RoG of protein in pure water, reline and aqueous reline solutions at 498K

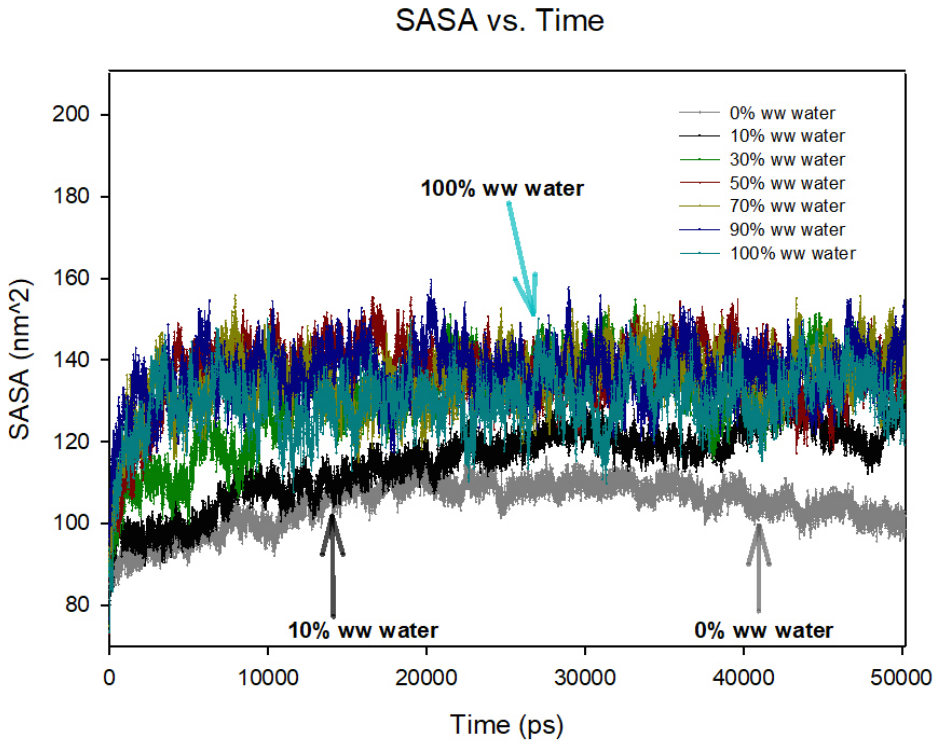

Figure 4.10. SASA of protein in pure water, reline and aqueous reline solutions at 498K

According to results, the RoG increase from its initial value of 1.4 nm to 1.9 nm in pure water and likewise SASA changes from 73 nm^2 to 130 nm^2 (RoG_{max} = 3.0 nm, $SASA_{max}$ = 150 nm^2). While the trend shows protein unfolding, as expected at such high temperature, the fluctuations are high. In comparison to pure water, the protein remains relatively stable in pure reline (RoG_{ave} = 1.53 nm, $SASA_{max}$ = 110 nm^2) and 10% ww water system (RoG_{max} = 1.9 nm, $SASA_{max}$ = 116 nm^2). With increasing water fraction in the solution, the fluctuations in RoG and SASA values increase, indicating higher changes in the native structure of lysozyme.

Similarly, the RDF analysis was used to study the interactions between protein and other components in the systems under study. The results are shown in Appendix A. According to the results of rdf analysis, the interactions between protein and urea are lower in 0% ww and 10% ww water systems in comparison to more diluted systems. Similarly to the previous analysis of protein structure at room temperature (Figure 4.8), at high temperature the interactions between protein and water decrease from slightly diluted reline (10% ww water) to more diluted reline solution (90% ww water).

4.4 Refolding of thermally unfolded protein at room temperature

Proper refolding protein is an important issue, which has been studied experimentally and computationally. Herein, as mentioned above, we further analyze the refolding of thermally unfolded protein structure in water, pure

reline and in diluted reline solutions. The thermally unfolded lysozyme structure from the unfolding simulation at T = 498 K in water after 10 ns of the simulation was used to analyze refolding patterns. The unfolded protein was placed in water, reline and its aqueous solutions at 298 K. For reference, the initial values of RoG and SASA of the unfolded protein structure are 1.9 nm and 133 nm^2, respectively. Refolding patterns of lysozyme in different environments were studied by analyzing radius of gyration (RoG, Figure 4.11) and solvent accessible surface area (SASA, Figure 4.12) of the protein. In addition, RoG (Figure 4.13) and SASA (Figure 4.14) of protein refolding in the first 100 ps (0.1 ns, nvt equilibration) were represented separately.

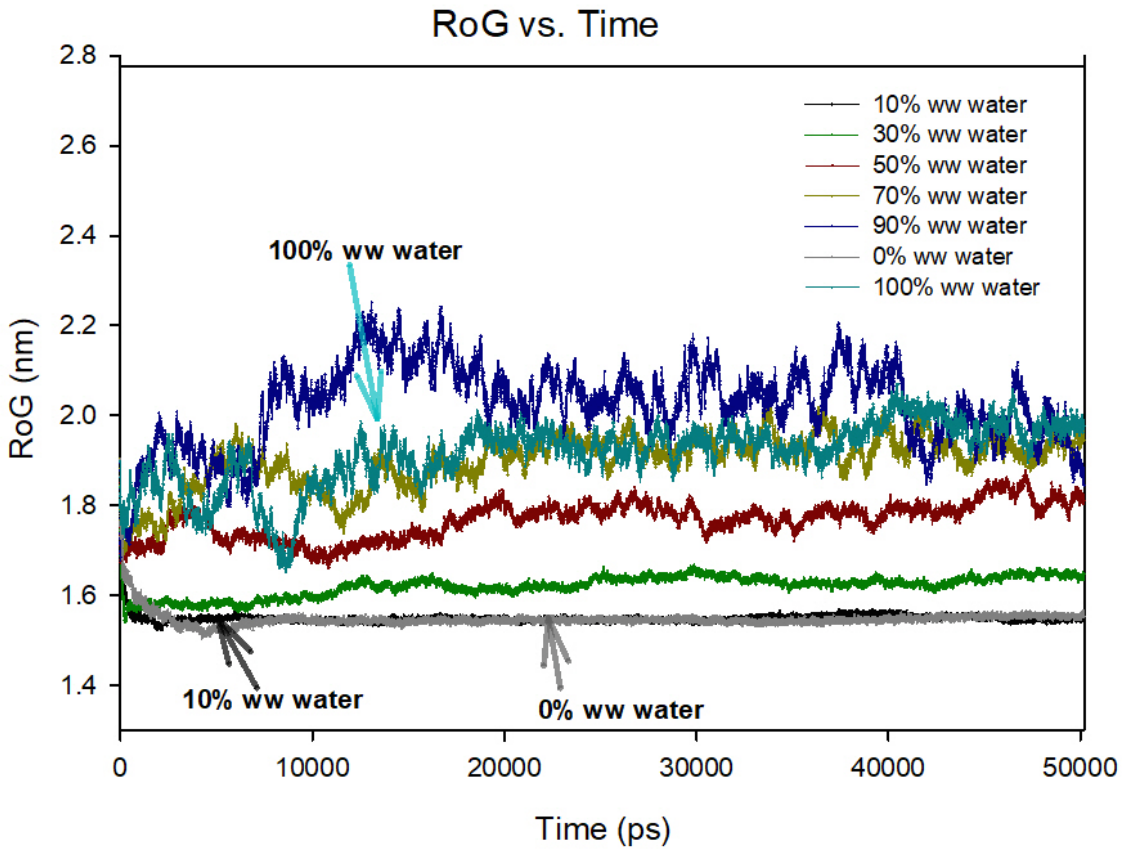

Figure 4.11. RoG of protein in pure water, pure reline and aqueous reline solutions

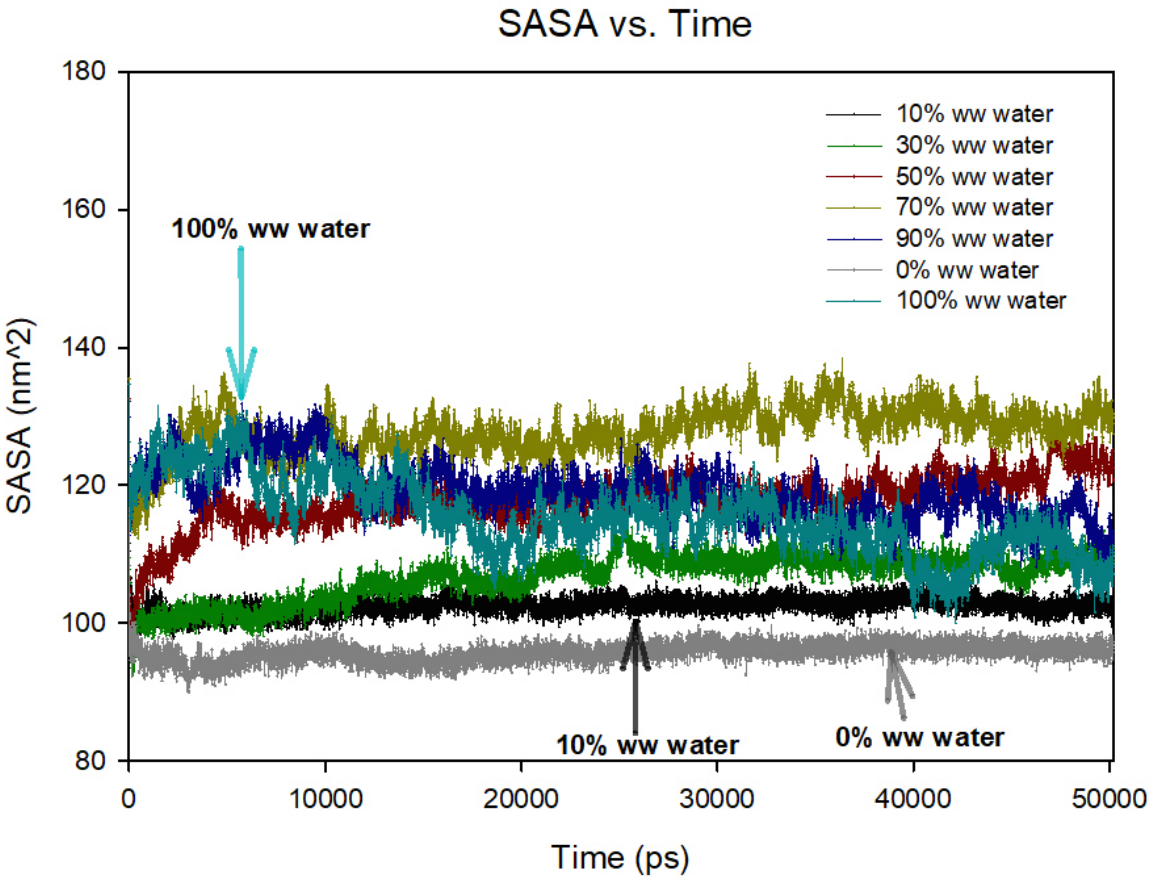

Figure 4.12. SASA of protein in pure water, pure reline and its aqueous solutions

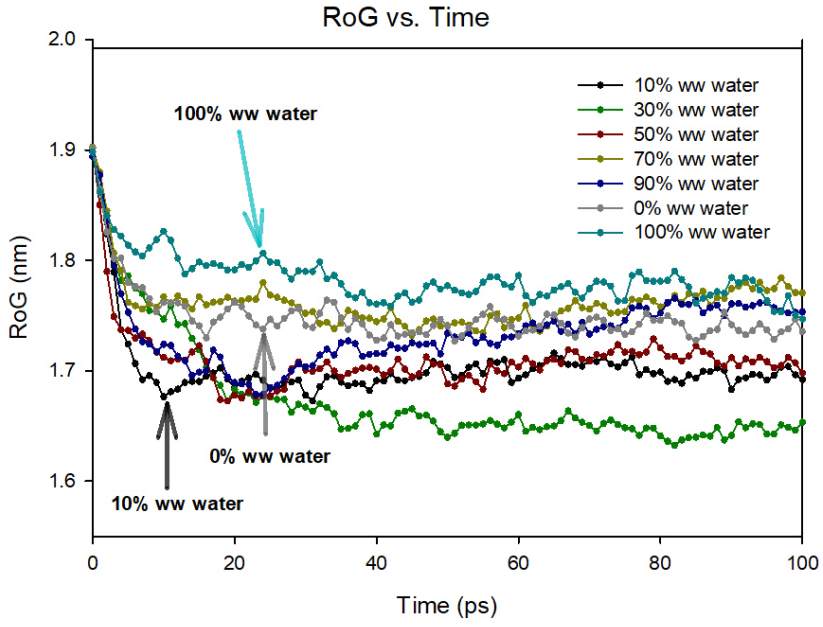

Figure 4.13. RoG of protein in pure water, pure reline and its aqueous solutions in the first 100 ps of simulations

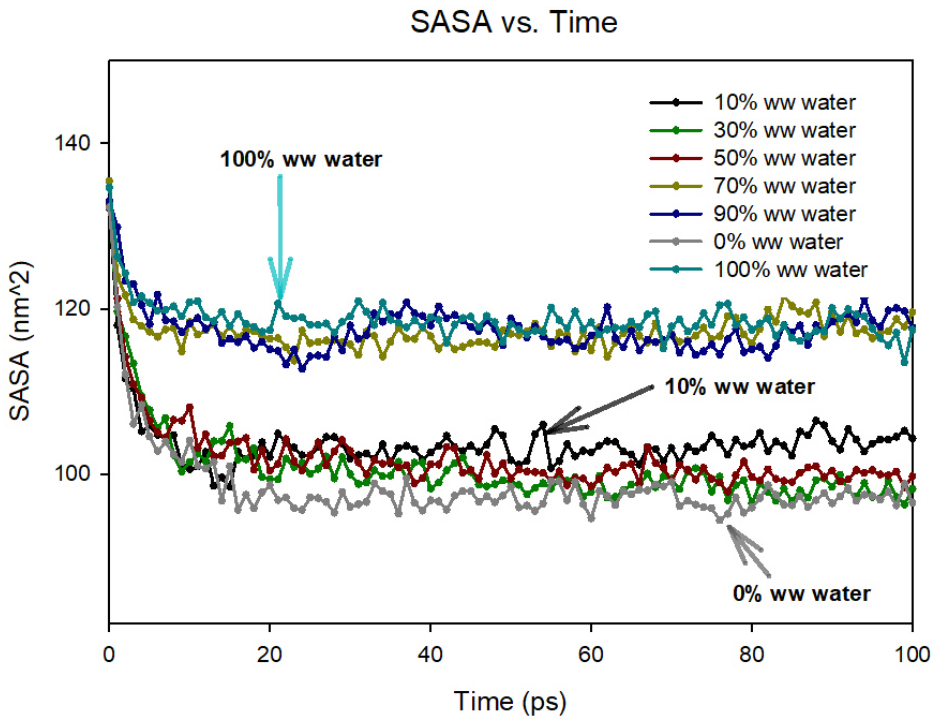

Figure 4.14. SASA of protein in pure water, pure reline and its aqueous solutions in the first 100 ps of simulations

On the basis of the results of RoG (Figure 4.11) and SASA (Figure 4.12) analysis, in 50 ns of simulation more effective protein refolding is obtained in pure reline, 10% and 30% diluted reline solutions. According to Figure 4.11, in pure reline, 10% and 30% diluted reline solutions; in general, RoG value was significantly decreased from 1.9 nm to 1.55 – 1.65 nm in 50 ns. Furthermore, according to Figure 4.12, SASA values decrease from 133 nm^2 to 95 nm^2 (63% completion of refolding, 0% ww water), 103 nm^2 (50% of refolding, 10% ww water) and 110 nm^2 (38% of refolding, 30% ww water and 100% ww water). Refolding is slower in more aqueous reline solutions and in pure water, as SASA and RoG values deviate more vigorously.

Furthermore, the RDF analysis was used to see the interactions between protein and components of each solvent (Appendix B). According to the results of RDF analysis, the interactions between protein and solvent components have the trends similar to the interaction patterns shown previously in section 4.2.4 (Figure 4.8). The snapshots of thermally unfolded protein structures obtained after 50 ns simulations in pure water, pure reline and slightly diluted reline (10% ww water) are represented on Figure 4.15. Nevertheless, even the protein was not refolded back completely to its native structure in 50 ns, the initial refolding patterns were indicated. In order to achieve a complete refolding, it is suggested to increase the simulation time or to use another methods such as REMD (Replica – Exchange Molecular Dynamics) (37).

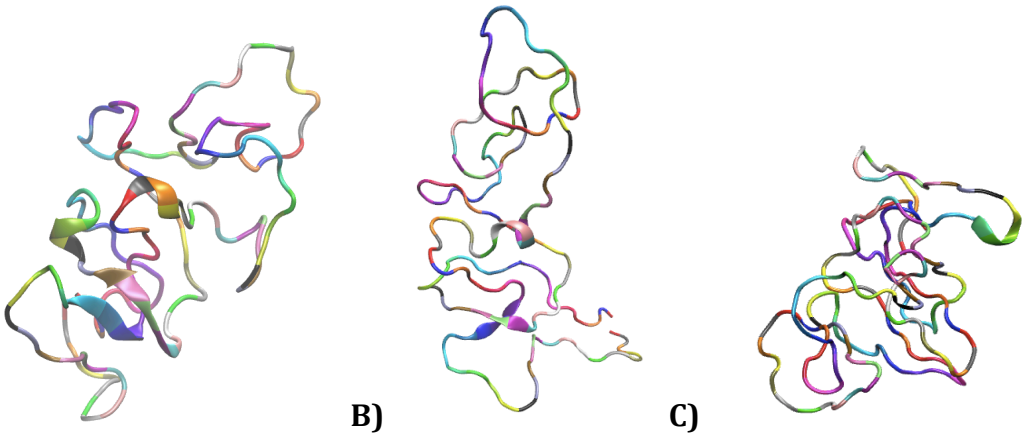

Figure 4.15. Structures of thermally unfloded lysozyme after 50 ns simulation at room temperature in: (A) pure water, (B) pure reline, (C) 10% ww water system

Chapter 5 – Conclusion and future work

In summary, Deep Eutectic Solvents (DESs) are new type of solvents, which are environmentally friendly, less expensive and can be easily synthesized. There are many different applications of DESs in catalysis, desulfurization process, gas adsorption, extraction, etc. This work was dedicated to explore one of the applications of DESs in protein structural stability by using Molecular Dynamics simulations on gromacs software.

To sum up, this study was based on three main objectives in analyzing: (i) protein structure at room temperature, (ii) protein structure at high temperature and (iii) protein structure during refolding at room temperature. In addition, the simulations were performed using the native structure of lysozyme to study its stability in cases (i) and (ii), while in case (iii) thermally unfolded protein structure was used to study its refolding pattern. In all cases under study, protein was put for 50 ns into (i) pure DES of choline chloride and urea combined in 1: 2 molar ratio (reline solution), (ii) aqueous solutions of reline and (iii) water for reference. The analysis of the results was performed by looking at the changes in protein structure via SASA, RoG and RMSF. Moreover, number of Hydrogen bonds and RDF plots were used to see the interaction patterns between protein and components of each particular solvent.

Moving to the conclusions based on the results of studying (i) protein structure at 298 K, it was seen that protein keeps its native structure in more

compact shape in reline diluted via 10% ww water. From the interactions patterns it was suggested that the reason is that protein surrounded by water keeps its compactness due to viscous DES. In pure reline, the protein structure shows stability and less compactness in comparison to slightly diluted reline mentioned previously, as there are more interactions between protein and urea. In both cases there are no changes in active site of the lysozyme. For reference, in less viscous environment of pure water, the native structure of protein is stable, but less compact. In more aqueous reline solutions, the native structure of lysozyme is less stable and less compact.

In addition, similar results were obtained from studying (ii) protein structure at high temperature, when usually unfolding occurs. In this case, in 50 ns of the simulation, the native structure of lysozyme showed more vigorous unfolding pattern in pure water and in more aqueous reline solutions in comparison to pure reline and least diluted reline solution (10% ww water), where protein was more thermally stable.

Finally, from the study of (iii) refolding protein of thermally unfolded protein at room temperature, it was also concluded that more effective refolding of unfolded lysozyme occurs in pure reline (63% of refolding) and least diluted reline solutions (50% of refolding, 10% ww water) in comparison to pure water and more diluted reline, where protein structure deviates more vigorously during 50 ns of simulation.

Nevertheless, the limited simulation time of 50 ns shows only the initial trends of protein stability, unfolding and refolding patterns. Consequently, in order to understand the whole picture of the interactions involved in the study of protein structure in DESs, for the future work it is suggested to:

- increase the simulation time to see the patterns of protein stability in larger period of time
- use other methodologies like REMD to accelerate MD simulations
- study and explore the application of other DESs with different compositions in protein stability at room and high temperatures, as well as in protein refolding.

Reference List:

1. Choi YH, van Spronsen J, Dai Y, Verberne M, Hollmann F, Arends IW, et al. Are natural Deep Eutectic Solvents the missing link in understanding cellular metabolism and physiology? Plant Physiol. 2011;156(4):1701-5.
2. Sanchez-Fernandez A, Edler KJ, Arnold T, Alba Venero D, Jackson AJ. Protein conformation in pure and hydrated Deep Eutectic Solvents. Phys Chem Chem Phys. 2017;19(13):8667-70.
3. Esquembre R, Sanz JM, Wall JG, del Monte F, Mateo CR, Ferrer ML. Thermal unfolding and refolding of lysozyme in Deep Eutectic Solvents and their aqueous dilutions. Phys Chem Chem Phys. 2013;15(27):11248-56.
4. Welton T. Room-Temperature Ionic Liquids. Solvents for Synthesis and Catalysis. Chem Rev. 1999;99(8):2071-84.
5. Smith EL, Abbott AP, Ryder KS. Deep Eutectic Solvents (DESs) and their applications. Chem Rev. 2014;114(21):11060-82.
6. Abbott AP, Capper G, Davies DL, Munro HL, Rasheed RK, Tambyrajah V. Preparation of novel, moisture-stable, Lewis-acidic ionic liquids containing quaternary ammonium salts with functional side chains. Chem Commun (Camb). 2001(19):2010-1.
7. Li X, Row KH. Development of Deep Eutectic Solvents applied in extraction and separation. J Sep Sci. 2016;39(18):3505-20.
8. Elledge SJ, Harper JW. The role of protein stability in the cell cycle and cancer. Biochim Biophys Acta. 1998;1377(2):M61-70.
9. Pace CN, Shirley BA, McNutt M, Gajiwala K. Forces contributing to the conformational stability of proteins. FASEB J. 1996;10(1):75-83.
10. Dill KA, MacCallum JL. The protein-folding problem, 50 years on. Science. 2012;338(6110):1042-6.
11. Sedov IA, Magsumov TI. Molecular dynamics study of unfolding of lysozyme in water and its mixtures with dimethyl sulfoxide. J Mol Graph Model. 2017;76:466-74.
12. Yon JM. Protein folding: a perspective for biology, medicine and biotechnology. Braz J Med Biol Res. 2001;34(4):419-35.
13. Pace CN, Treviño S, Prabhakaran E, Scholtz JM. Protein structure, stability and solubility in water and other solvents. Philos Trans R Soc Lond B Biol Sci. 2004;359(1448):1225-34; discussion 34-5.
14. Manning MC, Chou DK, Murphy BM, Payne RW, Katayama DS. Stability of protein pharmaceuticals: an update. Pharm Res. 2010;27(4):544-75.
15. Frokjaer S, Otzen DE. Protein drug stability: a formulation challenge. Nat Rev Drug Discov. 2005;4(4):298-306.
16. Kumari M, Dohare N, Maurya N, Dohare R, Patel R. Effect of 1-methyl-3-octyleimmidazolium chloride on the stability and activity of lysozyme: a spectroscopic and molecular dynamics studies. J Biomol Struct Dyn. 2017;35(9):2016-30.
17. Beck DA, Daggett V. Methods for molecular dynamics simulations of protein folding/unfolding in solution. Methods. 2004;34(1):112-20.
18. Abbott AP, Barron JC, Ryder KS, Wilson D. Eutectic-based ionic liquids with metal-containing anions and cations. Chemistry. 2007;13(22):6495-501.
19. Abbott AP, Capper G, Davies DL, Rasheed RK, Tambyrajah V. Novel solvent properties of choline chloride/urea mixtures. Chem Commun (Camb). 2003(1):70-1.
20. AlOmar MK, Hayyan M, Alsaadi MA, Akib S, Hayyan A, Hashim MA. Glycerol-based Deep Eutectic Solvents: Physical properties. J Mol Liq. 2016:98 - 103.

21. Zhang Q, De Oliveira Vigier K, Royer S, Jérôme F. Deep Eutectic Solvents: syntheses, properties and applications. Chem Soc Rev. 2012;41(21):7108-46.
22. Greaver TL, Sullivan TJ, Herrick JD. Ecological effects of nitrogen and sulfur air pollution in the US: what do we know? Front Ecol. 2012:365 - 72.
23. Li C, Li D, Zou SS, Li Z, Yin JM, Wang AL, et al. Extraction desulfurization process of fuels with ammonium-based Deep Eutectic Solvents. Green Chem. 2013.
24. Li C, Zhang J, Li Z, Yin J, Cui Y, Liu Y, et al. Extraction desulfurization of fuels with 'metal ions' based Deep Eutectic Solvents (MDESs). Green Chem. 2016.
25. D'Alessandro DM, Smit B, Long JR. Carbon dioxide capture: prospects for new materials. Angew Chem Int Ed Engl. 2010;49(35):6058-82.
26. Garcia G, Aparicio S, Ullah R, Atilhan M. Deep Eutectic Solvents: Physicochemical Properties and Gas Separation Applications. Energy Fuels. 2015.
27. Shishov A, Bulatov A, Locatelli M, Carradori S, Andruch V. Application of Deep Eutectic Solvents in analytical chemistry. A review. Microchem J. 2017.
28. Shahbaz K, Mjalli FS, Hashim MA, AlNashef M. Using Deep Eutectic Solvents for the removal of glycerol from palm oil - based biodiesel. J Appl Sci. 2010.
29. Durand E, Lecomte J, Villeneuve P. Deep Eutectic Solvents: Synthesis, application, and focus on lipase-catalyzed reactions. Eur J Lipid Sci Technol. 2013:379 - 85.
30. Durand, E., Lecomte J, Barea B, Piombo G. Evaluation of Deep Eutectic Solvents as new media for Candida antarctica B lipase catalyzed reactions. Process Biochem. 2012:2081 - 9.
31. Monhemi H, Housaindokht MR, Moosavi-Movahedi AA, Bozorgmehr MR. How a protein can remain stable in a solvent with high content of urea: insights from molecular dynamics simulation of Candida antarctica lipase B in urea : choline chloride Deep Eutectic Solvent. Phys Chem Chem Phys. 2014;16(28):14882-93.
32. Apol E, Apostolov R, Berendsen HJC, et al. GROMACS Reference Manual. 2018 ed.
33. 1AKI: THE STRUCTURE OF THE ORTHORHOMBIC FORM OF HEN EGG-WHITE LYSOZYME AT 1.5 ANGSTROMS RESOLUTION [Internet]. Available from: https://www.rcsb.org/structure/1AKI.
34. Shah D, Mjalli FS. Effect of water on the thermo-physical properties of Reline: An experimental and molecular simulation based approach. Phys Chem Chem Phys. 2014;16(43):23900-7.
35. Rocco AG, Mollica L, Ricchiuto P, Baptista AM, Gianazza E, Eberini I. Characterization of the protein unfolding processes induced by urea and temperature. Biophys J. 2008;94(6):2241-51.
36. Eleftheriou M, Germain RS, Royyuru AK, Zhou R. Thermal denaturing of mutant lysozyme with both the OPLSAA and the CHARMM force fields. J Am Chem Soc. 2006;128(41):13388-95.
37. Gromacs. Replica-Exchange Molecular Dynamics [Available from: http://www.gromacs.org/Documentation/How-tos/REMD.

Appendices

Appendix A

The results of RDF analysis between protein and components of each solvent obtained from 50 ns simulations on protein structure at high temperature.

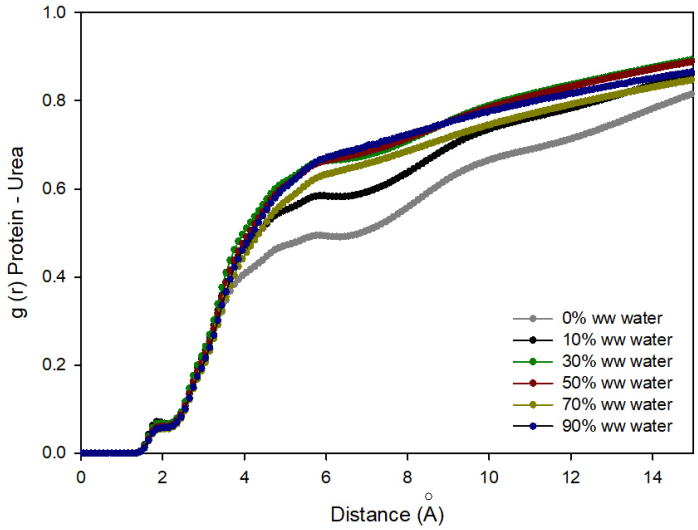

Figure A.1. RDF analysis of protein – urea interactions

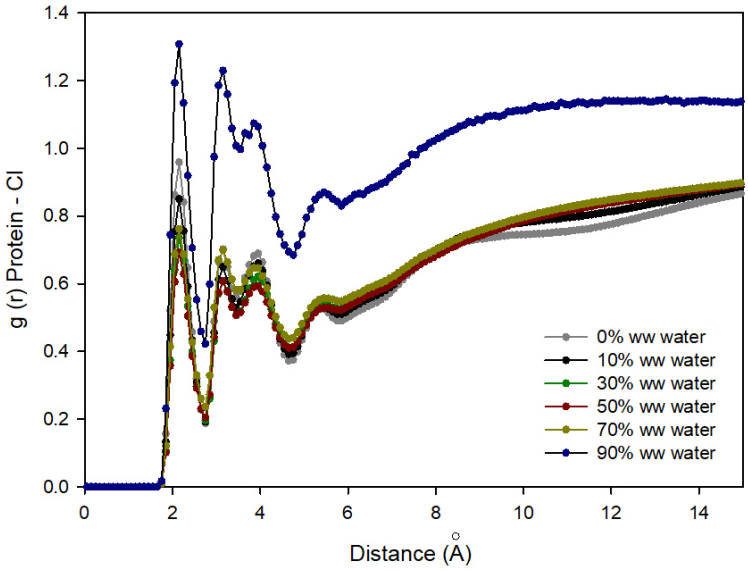

Figure A.2 RDF analysis of protein – Cl interactions

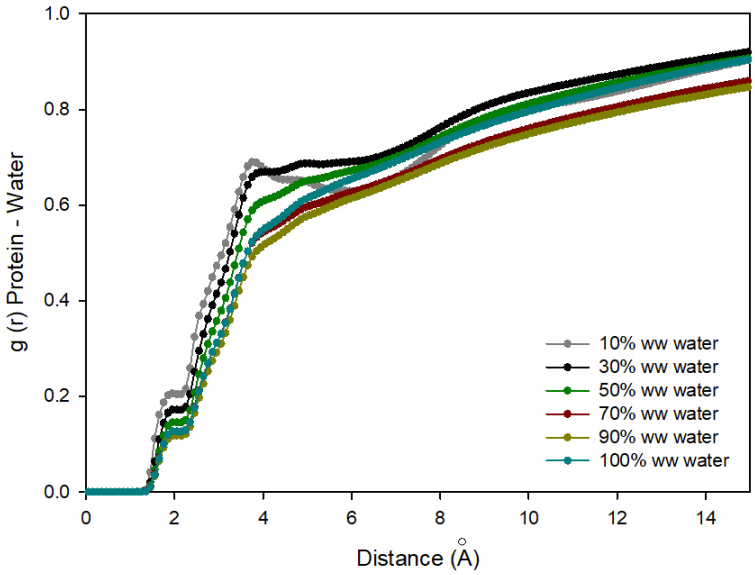

Figure A.3. RDF analysis of protein – water interactions

Appendix B

The results of RDF analysis between protein and components of each solvent obtained from 50 ns study on refolding patterns of thermally unfolded protein at room temperature.

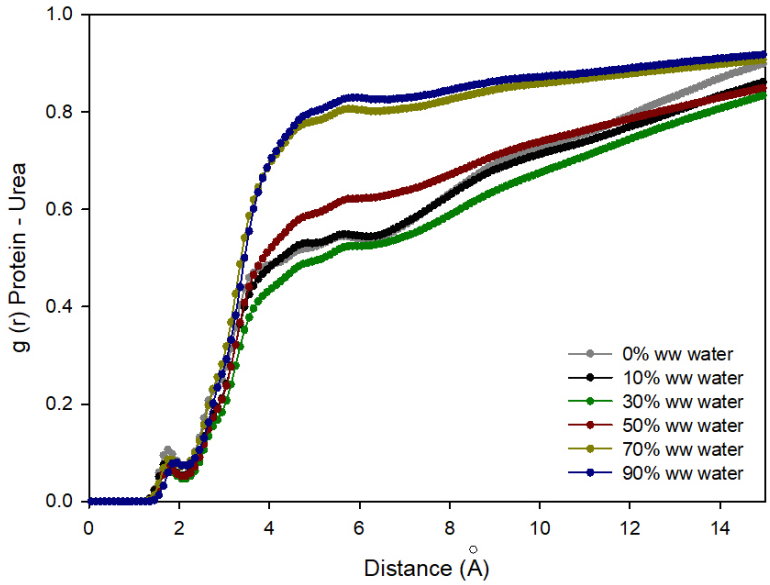

Figure B.1. RDF analysis of protein – urea interactions

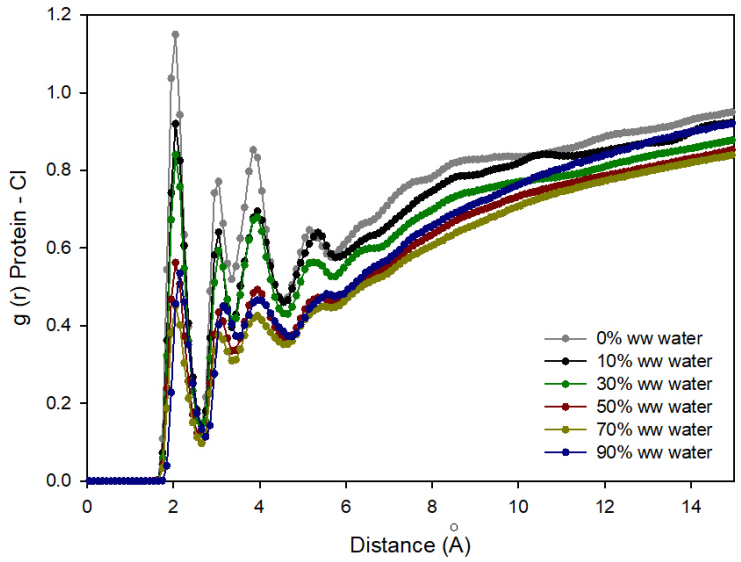

Figure B.2. RDF analysis of protein – Cl interactions

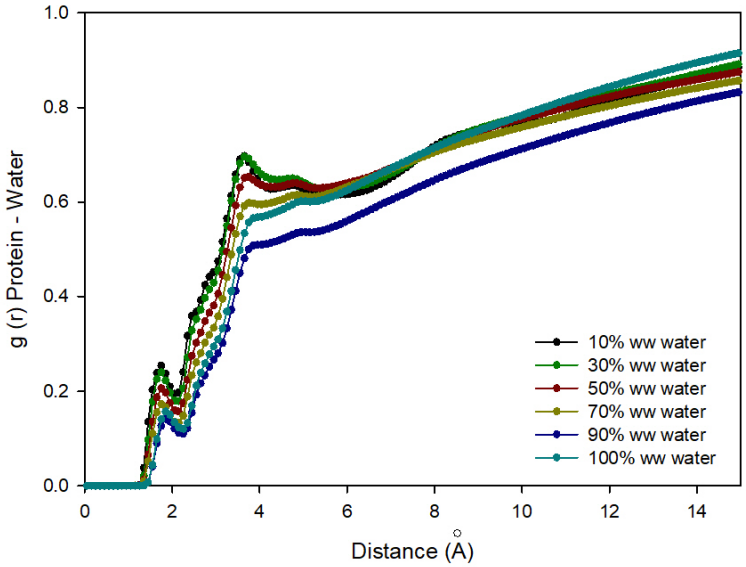

Figure B.3. RDF analysis of protein – water interactions