SUMMER PROJECT REPORT-2015

Hydrophobic interactions analysis of Urea and aqueous Guanidinium Chloride mixtures with amino acids analogs- Molecular Dynamics Studies.

NIT JALANDHAR

IISER BHOPAL

SUBMITTED BY:

Tanu Rana
Roll no. 14320001
M.Sc Chemistry
NIT Jalandhar.

SUBMITTED TO:

Dr. Rajesh Kumar Murarka
Assistant Professor
Department of Chemistry,
IISER Bhopal

Acknowledgements

I take this opportunity to express my sincere gratitude to Dr. R.K. Murarka for allowing me to work in his lab as a summer intern and for several insightful discussions and constant encouragement to pursue science.

My thanks and appreciation also goes out to Mr. Rasenjit Ghosh Dastidar, Mr. Navjeet Ahalawat and all the other members of RKM research group for their constant support and guidance through this project at IISER Bhopal.

July-2015

IISER Bhopal Tanu Rana

Certificate

The work embodied in this report is the result of studies carried out by **Ms. Tanu Rana** in the **Department of Chemistry, Indian Institute of Science Education and Research, Bhopal** under my supervision during June 1 to July 21, 2015. Ms. Tanu Rana was a Summer Fellow of the Indian Institute of Science Education and Research, Bhopal during this period.

Dr. Rajesh Kumar Murarka
Department of Chemistry
IISER, Bhopal

Dr Rajesh Murarka
Asstt. Professor, Chemistry
IISER, BHOPAL

Table of Contents

Abstract

Proteins are marginally stable, and this stability can be perturbed by the addition of small organic molecules known as cosolvents. Cosolvents that shift the equilibrium toward the unfolded ensemble are termed denaturants, whereas those that favor the folded ensemble are known as protecting osmolytes. Here, in view of understanding collapse of proteins in the mixture of denaturants we explored the effects on hydrophobic interactions between small methane molecules using extensive molecular dynamic (MD) simulations. As a measure of interactions of methane with other components of the modeled systems we computed the radial distribution functions (RDFs) and potential of mean force (PMFs) between varying solute-solvent pairs at varying concentrations of denaturants. Analysis of the free energies of interactions (calculated from PMFs) for various pairs revealed that the hydrophobic association of methane increased with increased concentration of GdmCl in the mixture relative to that in pure water and aq urea solutions (decreased with respect to aq. GdmCl solution). These results for the model systems faintly links the protein collapsing mechanism to the denaturant's ability (preferably GdmCl) to form hydrogen bonds readily with water and other denaturant (urea), thereby proposing that GdmCl interacts (predominant interactions) indirectly with methane whereas urea interacts directly with methane (feeble interactions), thereby affecting its hydrophobic interactions. For a robust understanding of this mechanism further studies are indispensable.

Theory of Molecular Dynamic Simulations

One of the principal tools in the theoretical study of biological molecules is the method of molecular dynamics simulations (MD). This computational method calculates the time dependent behavior of a molecular system. MD simulations have provided detailed information on the fluctuations and conformational changes of proteins and nucleic acids. These methods are now routinely used to investigate the structure, dynamics and thermodynamics of biological molecules and their complexes. They are also used in the determination of structures from x-ray crystallography and from NMR experiments.

What is molecular dynamics?

Molecular dynamics (MD) is a computer simulation technique that allows one to predict the time evolution of a system of interacting particles (atoms, molecules, granules, etc.). The basic idea is simple. First, for a system of interest, one has to specify: (a) a set of initial conditions (initial positions & velocities of all particles in the system); (b) interaction potential for deriving the forces among all the particles. Second, the evolution of the system in time can be followed by solving a set of classical equations of motion (Newton's Equation) for all particles in the system. Within the framework of classical mechanics, the equations that govern the motion of classical particles are the ones that correspond to the second law of classical mechanics formulated by Sir Isaac Newton over 300 years ago:

$$m_i\vec{a}_i = \vec{F}_i \quad \text{or} \quad m_i \frac{d\vec{v}_i}{dt} = \vec{F}_i \quad \text{or} \quad m_i \frac{d^2\vec{r}_i}{dt^2} = \vec{F}_i \quad \text{for the } i^{th} \text{ particle}$$

Solving a set of classical equations of motion for all particles in the system:

$$m_i\vec{a}_i = \vec{F}_i \quad \text{or} \quad m_i \frac{d\vec{v}_i}{dt} = \vec{F}_i \quad \text{or} \quad m_i \frac{d^2\vec{r}_i}{dt^2} = \vec{F}_i \qquad \text{for the } i^{th} \text{ particle}$$

If the particles of interest are atoms, and if there are a total of Nat of them in the system, the force acting on the ith atom at a given time can be obtained from the interatomic potential V(r1, r2, r3, …, rNat) that, in general, is a function of the positions of all the atoms:

$$\vec{F}_i = -\vec{\nabla}_i U(\vec{r}_1, \vec{r}_2, \vec{r}_3, \ldots, \vec{r}_{N_{at}})$$

Once the initial conditions and the interaction potential are defined, the equations of motion can be solved numerically. The result of the solution are the positions and velocities of all the atoms as a function of time.

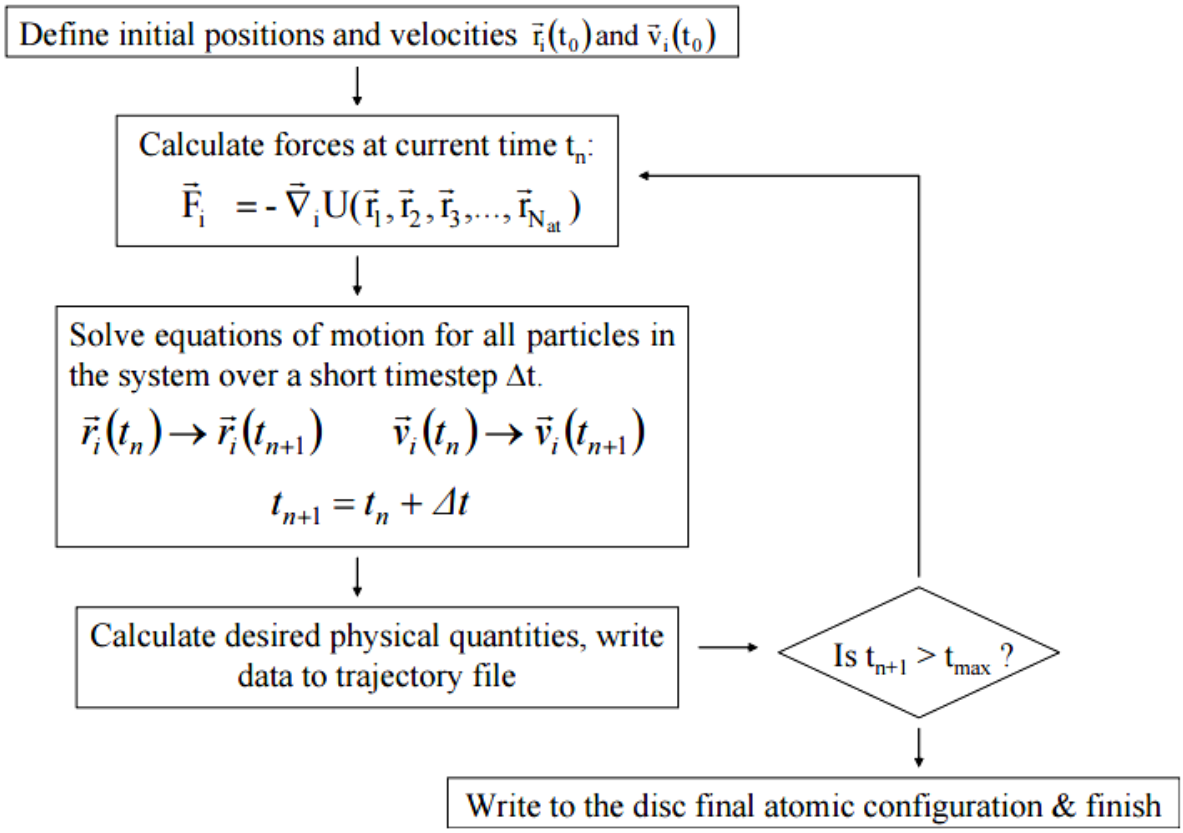
One can often treat a MD simulation like an experiment. Below is a common flowchart of an ordinary MD run:

[system setup]		[equilibration]		[simulation run]		[output]
sample selection (pot., *N*, IC, BC)	→	sample preparation (achieve *T*, *P*)	→	property average (run *L* steps)	→	data analysis (property calc.)

Advantages of MD:

- The only input in the model – description of interatomic/intermolecular interaction
- No assumptions are made about the processes/mechanism to be investigated
- Provides a detailed molecular/atomic-level information

Schematic diagram of a basic MD code



- MD is a deterministic technique: given initial positions and velocities, the evolution of the system in time is, in principle, completely determined (in practice, accumulation of integration and computational errors would introduce some uncertainty into the MD output).
- MD can be also used as a statistical mechanics method: it generates a set of configurations that are distributed according to statistical distribution functions. In many cases we are not interested in trajectories of individual atoms, we are interested in macroscopic properties of the material. MD information can be averaged over all the atoms in the system and over time to obtain thermodynamic parameters.
- The main strengths of the MD method is the ability to study fast non-equilibrium processes with atomic-level resolution (e.g. microscopic mechanisms of damage/plastic deformation due to a shock wave propagation, dynamic fracture and crack growth, ion bombardment, cluster impact, etc.). For many of these problems, MD method does not have an alternative

1. Introduction

Solvents such as guanidinium chloride (GdmCl) and urea, denature proteins. Protein denaturation is partial or complete disorganisation of a protein's characteristic three dimensional shape as a result of disruption of its secondary, tertiary, and quaternary structure, leading to loss of its biochemical activity. Despite extensive studies, the denaturation mechanism is not fully understood largely because of the perplexity of protein structures in the presence of denaturants. Furthermore, experimental data could not be used to draw probable denaturation mechanism since the interaction free energies between the denaturants and protein are very small[1] which makes it difficult to use experimental data to derive a denaturation mechanism. Recent studies have suggested two distinct mechanisms of denaturation (Direct and Indirect) based on different transfer experiments that measure solubilities of various peptide units and amino acid side chains or analogs in water, apolar solvents, urea, and GdmCl [1] [2] .

In the direct mechanism, the denaturant molecules disrupt the protein structure by interacting with the amino acid residues and the peptide backbone of protein either through electrostatic interactions or by forming hydrogen bonds. The Indirect mechanism usually operates by disrupting the structure of water around the protein, thereby enhancing the solubility of hydrophobic groups of protein, altering the hydrophobic interactions associated with the protein. This present work also revolves around these hydrophobic interactions (interactions which leads to bringing in contact the hydrophobic molecules in a polar solution).

For both urea and GdmCl various theories have been proposed suggesting both the direct [3] and the indirect mechanisms [1][4] for different proteins, leading to a controversial debate.

Since proteins solvated in a mixture of solvents are considered to be a more realistic representation of cellular environments, these are of immense fundamental importance. In view of understanding protein behaviour in a mixture of denaturants, we performed set of transfer experiments to assess the behaviour of small hydrophobic amino acid analog (such as methane; analog of alanine) in different mixture of urea and GdmCl solutions through MD Simulations.

Proteins generally adopt extended or unfolded conformations in either aqueous urea or GdmCl, so they are expected to become even further extended or unfolded on the addition of second denaturant to the first denaturant solution. However, on the contrary, a collapse (contraction) of the denatured protein is revealed [5] when a mixture of denaturant is used. As both denaturants compete with each other in interacting with the protein, it leads to unexpected complex behavior of protein energies and dynamics. This present work deals with comprehending this unexpected non-linear effect.

Our primary focus deals with comprehending solvent induced changes in the hydrophobic interactions between solute molecules (methane-methane interactions) and elucidation of the factors that affect these interactions.

2. Methods

2.1 Models:

We simulated the interactions between methane molecules in aqueous urea and guanidinium (Gdm^+) chloride (Cl^-) solutions using all-atom representations of all the chemical species. The molecules interact with each other through pairwise potentials that are composed of electrostatic and van der Waals interactions between the atomic centers. The electrostatic interactions arise from partial charges on each atomic center, whereas van der Waals interactions are modeled using the Lennard-Jones potential,

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

where r and σ are the distances and diameter between the two atomic centers, respectively, and ε is the well depth, $1/r^{12}$ term represents a repulsive hard core interaction between the atoms and $1/r^6$ term represents an attractive van der Waals interaction between the non-polar atoms. The potential V_{LJ} has its minimum at $r = 2^{1/6}\ \sigma$, $V = -\varepsilon$.

2.2 <u>Simulation Details</u>:

All the MD simulations were carried out with GROMACS version 4.6.5. OPLSAA force field was adopted for the solute (methane) molecules. TIP3P model was used for water. Kirkwood–Buff force field (KBFF) was employed for urea and modified KBFF for GdmCl. The electrostatic interaction was handled by the particle-mesh Ewald (PME) method with a cutoff of 1.2 Å. LINCS constraint algorithm was employed. The system was contained in a cubic 40 A box.

Solvent systems modeled were i) 8M Urea; ii) 6M GdmCl; iii) 3M Urea + 3M GdmCl; iv) 2M Urea + 4M GdmCl ; v) 4M Urea + 4M GdmCl; vi) 4M Urea + 2M GdmCl; vii) Pure water (All these contained 20 solute molecules).

System	[Urea]	[GdmCl]	Nu	Ng	Nw	Nm
1	8	-	311	-	1349	20
2	-	6	-	232	1227	20
3	3	3	116	116	1389	20
4	2	4	77	154	1335	20
5	4	4	154	154	1140	20
6	4	2	154	78	1443	20
7	-	-	-	-	2132	20

<u>Table 1</u>: *Different modeled systems and their respective concentrations in molarities.(Nu, Ng, Nw, Nm being the no. of urea, GdmCl, water, and methane molecules respectively in different solutions).*

The systems were prepared by adding 20 methane molecules in a pre-equilibrated box (equilibrated for 10 ns) containing different amounts of urea, GdmCl and water; the inserted methane molecules are included such that they do not overlap with other molecules and solvent molecules within their 2 A distance were removed. These systems were then equilibrated for 30 ns each. Finally using the final box size from previous NPT equilibration

step, production runs (500 ns long simulations in the NPT ensemble) were carried out (at 300 K and 1 atm) with Berendsen thermostat and barostat for equilibration and Parrinello-Rahman barostat for production runs respectively. Coordinates were saved every 10 ps for further study.

2.3 Radial Distribution Function [g(r)]:

It is a pair correlation function, which describes how, on average atoms in a system are packed radially around each other, it is especially useful for determining the average structure of disordered molecular systems such as liquids.

In a system g(r) describes how density varies as a function of distance from a reference particle. It's plot shows various peaks starting from the one giving atomic distribution in the first coordination shell leading to the subsequent distributions, at very low range every RDF plot tends to one as at this stage local density becomes equal to bulk density. Figure 1(a) (on the next page) shows a usual g(r) plot. In our present work we have computed g(r) for the following combinations:

- Methane-urea, methane-GdmCl, and methane-water pairs in different solutions for comparing the affinities of urea, GdmCl and water for the solute (CH_4) respectively, all these plots were also compared with each other for different denaturant concentrations.
- Methane-methane pairs: these determine the degree of solute aggregation due to effective hydrophobic interactions.
- Urea-GdmCl, urea-water, urea-urea, GdmCl-water, and GdmCl-GdmCl pairs: these determine how urea and GdmCl interacts with the rest of the system and with each other and how these molecules are distributed throughout the solution.

All these RDF plots (given in the results section on page 9) give useful insights about the average structure of the molecular systems we have adopted.

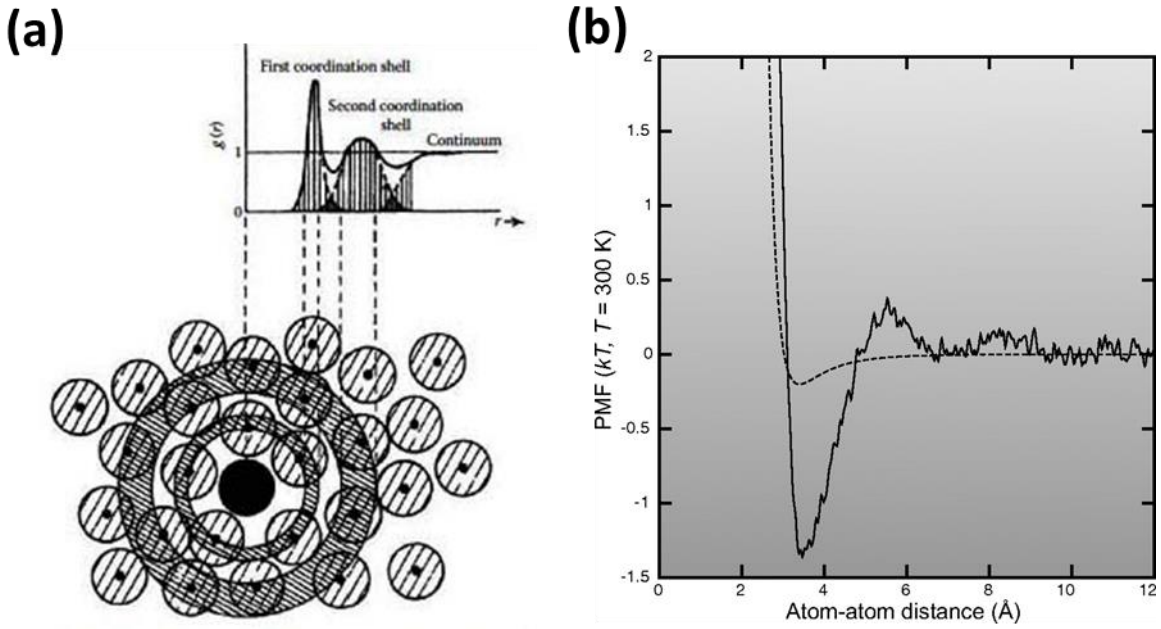


Figure 1: ***(a)*** *A usual g(r) plot, showing average distribution around the central molecule in solution,* ***(b)*** *A usual g(r) plot, showing average distribution around the central molecule in solution*

2.4 Potential Of Mean Force (Free Energy Profile):

Hydrophobic interaction free energy changes (amount of work done in bringing two non-polar groups towards each other from infinite separation to a distance r apart) in terms of reaction coordinates can be described in terms of PMF. In a system of N molecules, it is the potential that gives an average force over all configurations of the system.

PMF $W(r) = u(r) + \Delta w(r)$

It is a combination of direct energy interactions [u(r)] (interactions between the isolated solutes in absence of solvent) plus solvent mediated interactions [$\Delta w(r)$] (due to additional influence of solvent in either suppressing or promoting aggregation). By drawing a comparison between plots of Lennard-Jones potential of the solutes in water with the ones present in different concentrations of urea and GdmCl, it can be determined whether the solvents suppress or promote cluster formation of hydrophobic solutes (by comparing the well depths of the two plots) .

The simplest method of calculating for calculating PMF is to use separation r between two particles as the reaction coordinate. We have calculated PMF between two particles using this method (where g(r) is the radial distribution function, K_B is the Boltzmann distribution constant, T is the simulation temperature).

PMF [W(r)] = $-K_BT \ln g(r)$ + constant

The constant is chosen so that the most probable distribution corresponds to a free energy of zero. A relatively small change in the free energy (i.e. a small multiple of K_BT) may correspond to g(r) changing by an order of magnitude from its most likely value. When the same two particles were brought together in the gas phase, the free energy would simply be the pair potential u(r), which has only a single minimum. But the PMF between two particles in liquid oscillates with maximum and minimum. The lowest minima corresponds to the contact separated minima (CM) (hydrophobic association) of the solutes which is separated from the second lowest minima (solvent separated minima (SSM)) (hydrophobic solvation) through an energy barrier (desolvation barrier) (Figure 1(b)).The energy difference in going from the SSM to the CM gives the free energy of association of the corresponding particles. We have plotted the PMF curves (given in the results section) for various combinations of methane, urea, GdmCl, water for determining the stabilities of various interactions amongst them and computed their respective free energies (ΔG) of interaction.

Association thermodynamics of the solute pairs (in water + in solutions) can also be studied through the PMF curves (more is the well depth of the CM, more is the hydrophobic association and higher is the energy barrier between the CM and SSM lower is the hydrophobic solvation).

3. Results and Discussion:

3.1 Resemblance of radial distribution curves corresponding to different mixtures:

Radial distribution curves of different mixtures for the same combinations of solute-solvent, solvent-solvent and solute-solute pairs did not show any significant variations in terms of their respective first and subsequent coordinate shells, proposing that the change in urea and GdmCl concentration has little effect on the molecular distribution. The radial distribution functions between different solute-solvent or solvent-solvent pairs are shown in Figure 2.

RDF results indicate that methane molecules are surrounded by water as their nearest neighbours followed by other methane molecules (suggesting cluster formation), which are further followed by urea and GdmCl. Similarly molecular distributions around urea, GdmCl, water in different systems (calculated from RDF plots of different systems) is listed in Table 2.

System	Methane(M)	Urea (U)	GdmCl (G)	Water (W)
1	W-M-U (M>W>U)	W-M (U>W>M)	----------	W-M-U (W>M>U)
2	W-M-G (M>W>G)	----------	W-M-G (W>M>G)	G-M (G>M)
3,4,5,6	W-M-U-G (M>W>U>G)	W-M-G-U (G>U>M>W)	W-U-M-G (U>W>M>G)	G-M-U (W>G>M>U)
7	M-W (M>W)	---------	----------	W-M (W>M)

Table 2: *Molecular distributions around methane, urea, GdmCl, water in different systems starting from their nearest neighbors on the left to their farthest neighbors on the right.(the bracketed terms in each column represents the probability of occurrence in decreasing order).*

3.2 Effect on methane clusters in mixture compared to that in aqueous solutions:

Methane-methane association in mixtures is raised compared to that in pure water, as the g(r) peak of methane-methane interactions in mixtures shifted to a higher value (~3.8 kcal/mol) from ~3.2 kcal/mol in pure water. Similarly on analyzing the various PMFs curves of system 1 and 2 it was observed that methane-methane association is maximum in the aqueous GdmCl solution (System 2), and nearly minimum in aq. urea solution (System 1) among all the systems adopted, so methane clusters decreases in mixtures as compared to that in aq. GdmCl whereas increases as compared to that in aq. urea. This suggests maximum hydrophobic interactions in system 2 and minimum in 1.

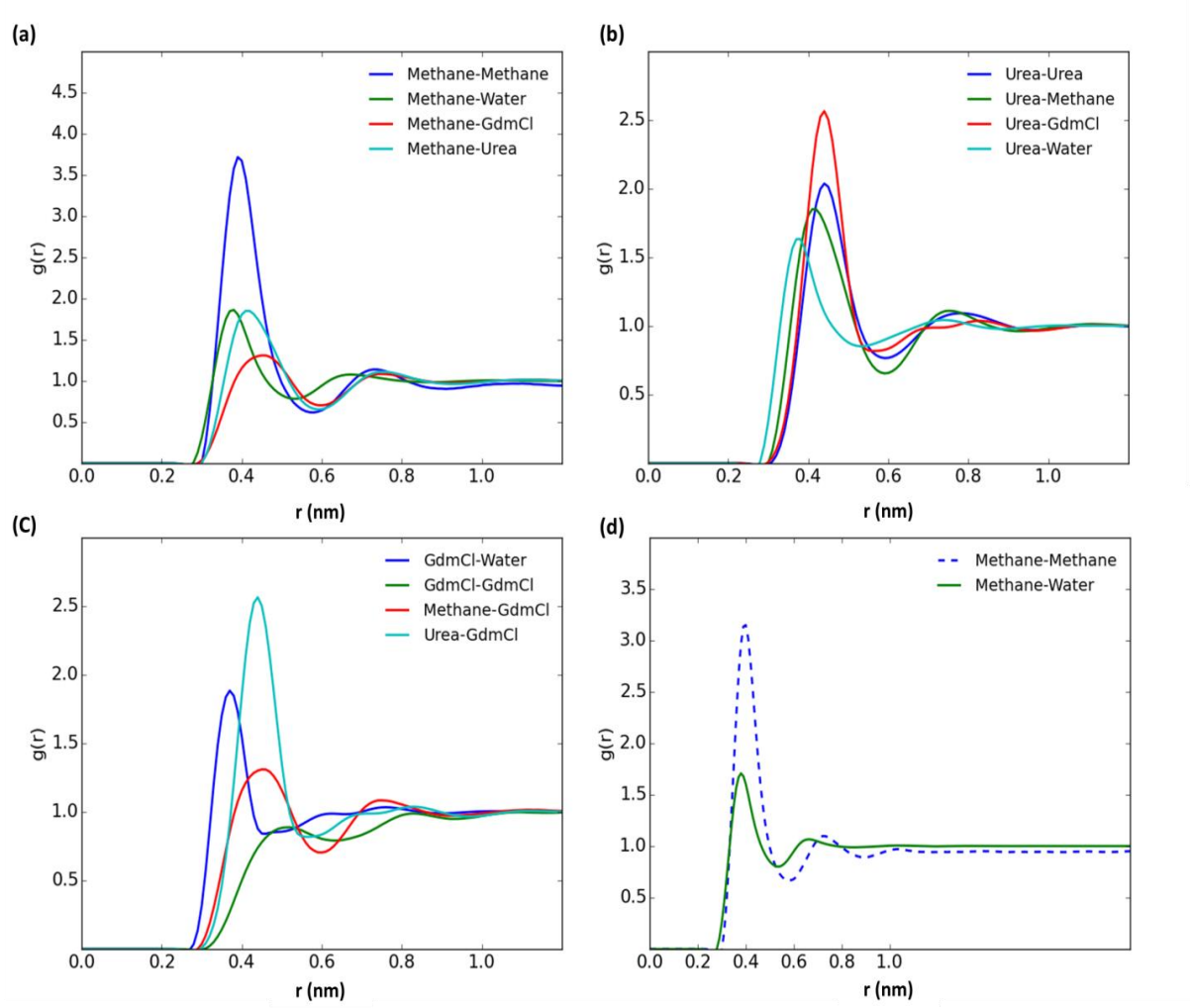


Figure 2: *RDFs of* ***(a)*** *methane,* ***(b)*** *urea,* ***(c)*** *GdmCl, and* ***(d)*** *water respectively with other components of the system showing distributions around them separately.(These curves show distribution in system 3 , distributions for other mixtures also show similar arrangements).*

3.3 PMFs calculated for different mixtures display minute changes:

As the concentration of GdmCl was raised in the systems the interaction free energies for different atom pairs were computed through their respective PMFs (Table 3)(difference between the contact minima (CM) and the solvent separated minima (SSM) was computed as the free energy of interactions).

ΔG	**System 1**	**System 2**	**System 3**	**System 4**	**System 5**	**System 6**	**System 7**
ΔG_{mm}	-0.7497	-0.8315	-0.725	-0.7559	-0.7717	-0.7869	-0.647
ΔG_{mg}	----	-0.1955	-0.113	-0.1367	-0.156	-0.1216	----
ΔG_{mu}	-0.1258	----	-0.292	-0.3257	-0.345	-0.3401	----
ΔG_{mw}	-0.3743	-0.3958	-0.349	-0.3655	-0.3758	-0.3757	-0.3039
ΔG_{gg}	----	+ve	+ve	+ve	+ve	+ve	+ve
ΔG_{uu}	-0.4724	----	-0.385	-0.3895	-0.372	-0.3610	----
ΔG_{gu}	----	----	-0.582	-0.5895	-0.5881	-0.5831	----
ΔG_{wg}	----	-0.5377	-0.484	-0.5035	-0.518	-0.4865	----
ΔG_{wu}	-0.2727	----	-0.294	-0.2859	-0.273	-0.2842	----

Table 3: *Free energies of interaction (ΔG) between different atom pairs in different systems.(Calculated from their respective PMFs), m = methane, g = GdmCl, w = water, u = urea.*

Inferences made from Table 3:

- For every mixture (system 3, 4, 5, 6) the ΔG_{mm} is more negative as compared to that in water, leading to the notion that methane-methane association is more favoured in mixtures that in pure water solution.
- ΔG_{mm} is most negative in aqueous GdmCl solution, leading to maximum methane-methane association among all systems.
- The interaction free energies of methane in each mixture follows the order $\Delta G_{mm} > \Delta G_{mw} > \Delta G_{mu} > \Delta G_{mg}$ (Figure 3a) , showing that methane has highest tendency to form aggregates and least tendency to interact with GdmCl. This means that GdmCl does not interact directly with methane.
- Similarly for water it was observed that in every mixture water has maximum tendency to interact with GdmCl and least with urea. This matches with the idea that GdmCl interacts directly with water and urea interacts directly with methane.

- For GdmCl (in mixtures)(Figure 3b) it was also seen that it has maximum interactions with urea, followed by water and methane ($\Delta G_{ug} > \Delta G_{gw} > \Delta G_{mg} > \Delta G_{gg}$).
- As the concentration of GdmCl increased in going from system 3 to 6, methane-methane, methane-urea and GdmCl-water interactions have raised slightly showing that methane-methane aggregation increases with increased GdmCl concentration.

Accounting for all the above observations, it could be said that in mixtures, both the denaturants play part in methane-methane aggregation (although GdmCl has a bigger part to play than urea), GdmCl has shown more possibilities of interacting indirectly with methane by interacting with water and urea while urea interacts directly with methane.

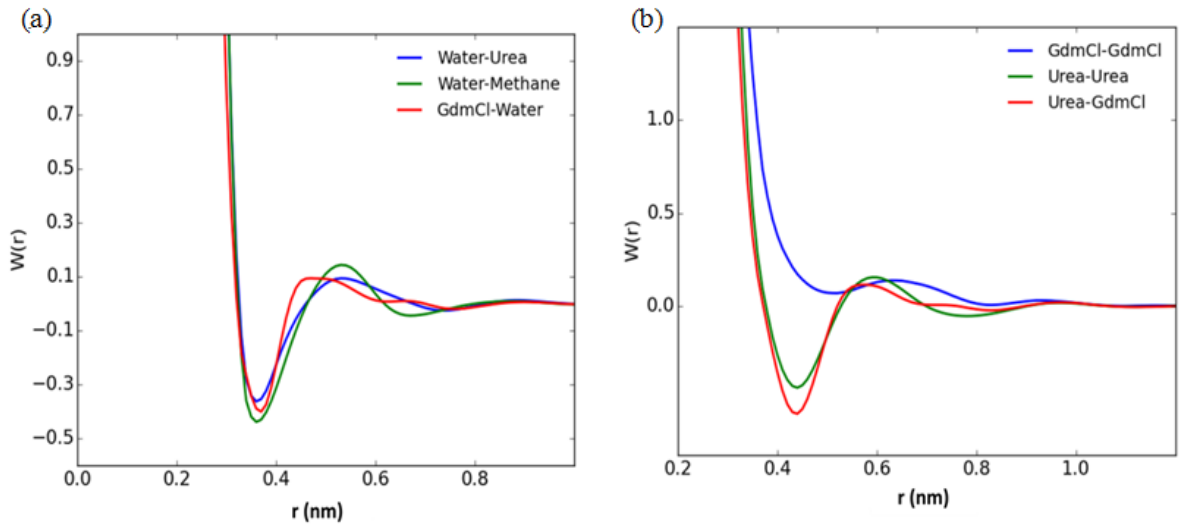


Figure 3: PMFs of **(a)** water with different components of system 3 displaying strength of interactions between them **(b)** PMFs between urea and GdmCl of system 3 displaying strength of interactions between them.

3.4 Comparison between aqueous solutions and mixtures:

Interactions among different components in the aq. solutions (system 1, 2) and mixtures (system 4; for simplicity only one mixture is taken) were compared (Figure 4). For methane-methane interactions similar observations were reported as suggested in section 3.2. Methane-urea and urea-water interactions in the mixture were found to be stronger than in aqueous system 1, as predicted above on the basis of ΔG, whereas opposite behavior in case

of GdmCl was seen as system 2 had stronger GdmCl-methane and GdmCl-water interactions as compared to the mixtures.

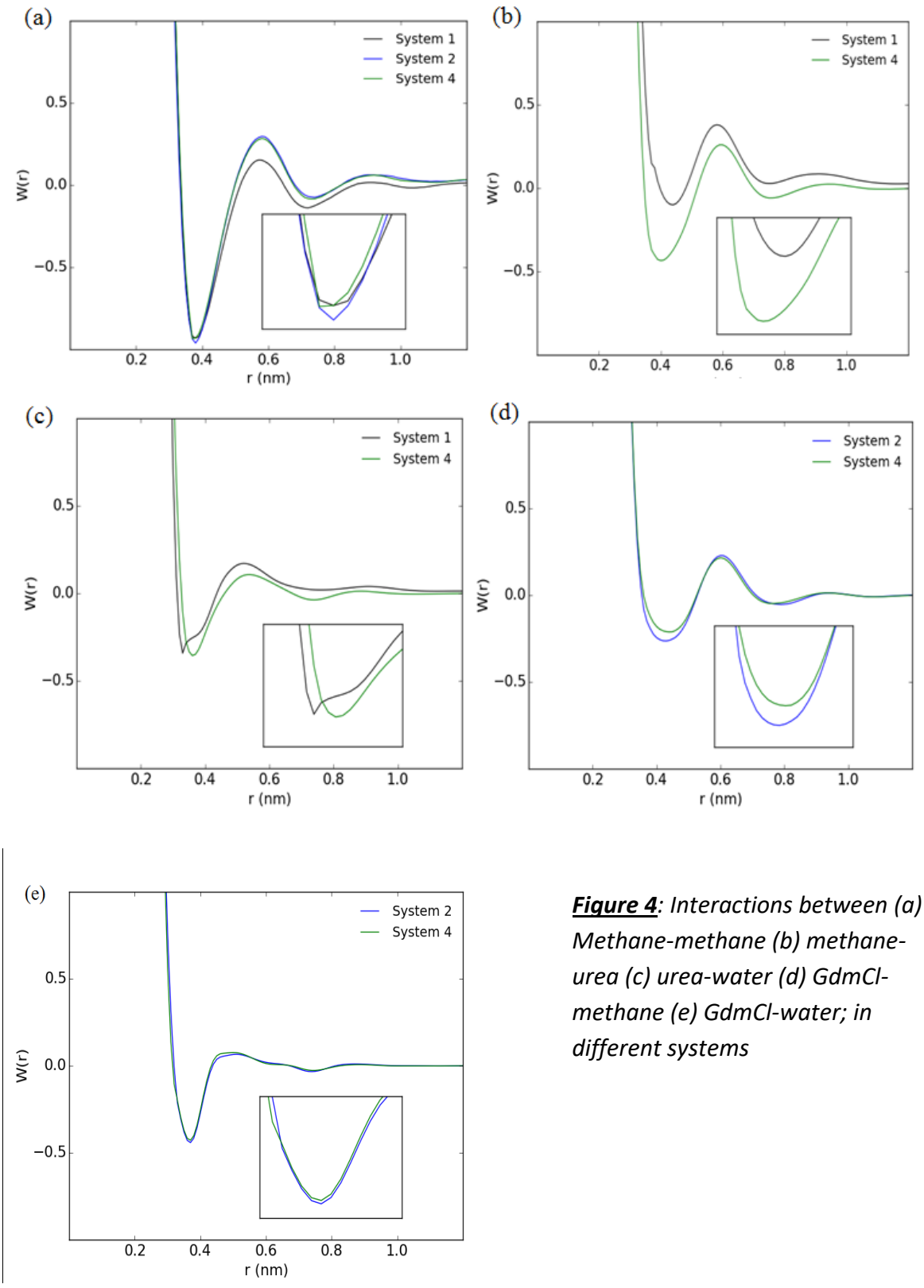


***Figure 4**: Interactions between (a) Methane-methane (b) methane-urea (c) urea-water (d) GdmCl-methane (e) GdmCl-water; in different systems*

4. Conclusion

Motivated by the need to understand the structural basis of mixture of denaturants induced collapse of proteins we have investigated the alterations in the hydrophobic interactions between small hydrophobic methane molecules in aqueous urea and GdmCl mixtures. The PMFs between methane molecules show that GdmCl has maximum stabilizing effect on methane-methane hydrophobic interactions whereas urea has nearly the least among all the mixtures.

Moreover, the strength of hydrophobic interactions show a slight increase in the mixture of urea and GdmCl as the amount of GdmCl is raised. These observations are consistent with molecular dynamics studies on small solutes[7] hydrophobic polymer [6] and dipeptides[8] . But the small free energy changes makes it difficult to infer a denaturation mechanism from energetic considerations alone, further investigations are required for deeper understandings.

These results faintly suggests that in mixtures of GdmCl and urea, urea interacts with methane directly (although weakly) without much interaction with water whereas GdmCl interacts with water and urea directly (preferably through hydrogen bonding, thereby masking urea and preventing it from interacting with methane), which indirectly affects the hydrophobic association of methane and methane-urea interactions leading to increased association of methane molecules . For a meticulous understanding of these systems further studies correlated with the above observations are requisite.

5. REFERENCES

(1) Nozaki, Y.; Tanford, C. J. Biol. Chem. 1963, 238, 4074-4081.

(2) Greene, R. F.; Pace, C. N. J. Biol. Chem. 1974, 249, 5388-5393.

(3) Robinson, D. R.; Jencks, W. P. J. Am. Chem. Soc. 1965, 87, 2462-2469.

(4) Tanford, C. AdV. Protein Chem. 1970, 24, 1-95.

(5) Xia, Das, Shakhnovich, Zhou. J. Am. Chem. Soc. 2012, 134, 18266-18274.

(6) Das, Xia, Zhou. Collapse. Langmuir. 2013, 29, 4877-4882.

(7) Wallqvist, A.; Covell, D. G.; Thirumalai, D. J. Am. Chem. Soc. 1998, 120, 427-428.

(8) Tobi, D.; Elber, R.; Thirumalai, D. Biopolymers 2003, 68, 359-369.

(9) Frenkel D. and Smith B, Understanding molecular simulations; Allen M. P. and Tildesley D. J, Computer simulation of liquids.