In silico studies of the human IAPP in the presence of osmolytes

Ashma Khan1 · Ishrat Jahan1 · Shahid M. Nayeem1

Received: 3 August 2021 / Accepted: 30 May 2022 / Published online: 14 June 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
The human islet amyloid polypeptide or amylin is secreted along with insulin by pancreatic islets. Under the drastic environmental conditions, amylin can aggregate to form amyloid fibrils. This amyloid plaque of hIAPP in the pancreatic cells is the cause of type II diabetes. Early stages of amylin aggregates are more cytotoxic than the matured fibrils. Here, we have used the all-atom molecular dynamic simulation to see the effect of water, TMAO, urea and urea/TMAO having ratio 2:1 of different concentrations on the amylin protein. Our study suggest that the amylin protein forms β-sheets in its monomeric form and may cause the aggregation of protein through the residue 13–17 and the C-terminal region. α-Helical content of protein increases with an increase in TMAO concentration by decreasing the SASA value of protein, increase in intramolecular hydrogen bonds and on making the short-range hydrophobic interactions. Electrostatic potential surfaces show that hydrophobic groups are buried and normalised configurational entropy of backbone, and side-chain atoms is lesser in the presence of TMAO, whereas opposite behaviour is observed in the case of urea. Counteraction effect of TMAO using Kast model towards urea is also observed in ternary solution of urea/TMAO.

Keywords Human IAPP · Aggregates · IDPs · α-Helix · Urea · TMAO

Introduction
Intrinsically disordered proteins (IDPs) are incapable to form stable three-dimensional structures. But IDPs are dynamically disordered and vary promptly in conformational space from extended random coil to compact protein [1]. Various types of disease mostly related with IDPs are Parkinson’s disease, Alzheimer’s disease, Huntington’s disease, amyotrophic lateral sclerosis (ALS), type 2 diabetes (T2D), cancer, etc. [2]. Human islet amyloid polypeptide (hIAPP), or amylin, is one of the IDPs which has propensity to aggregate into insoluble fibrils. Amylin, a 37-residue peptide, is neuroendocrine hormone co-secreted with insulin as a monomer found in the β-cells of the pancreas. Amylin controls glucose level in the blood and induces hormone for satiety which helps against overeating (Hay et al., 2015; Hinshaw et al., 2016). Amyloid accumulation in and around the β-cells in patients with T2D has hIAPP as the primary constituent [3]. There is experimental evidence that the aggregates of hIAPP lead to β-cell death [4]. IAPP is discharged in excess amount with the overproduction of insulin that leads to the pathological aggregation of amyloid fibrils on β-cell area [5]. Ninety-five percent of T2D patients have mature fibrillar of hIAPP (Höppener et al. 2000), and 100 million people have been affected worldwide by T2D [6]. Therefore, the study of molecular mechanism involved in protein misfolding and their associated aggregation is required.

Stability, folding and assembly of proteins are influenced in living organism under various physico-chemical conditions, involving the level of hydration and the presence of cosolvents and co-solutes at different concentrations [7]. Under physiological situations, proteins need to counteract towards any changes on the thermodynamic properties in order to avoid partial or full loss of their function. Accumulation of particular cosolvents is one of the natural mechanisms to counteract osmotic and other environmental disturbance. Cosolvents that do not interact with proteins and peptides in a specific manner are termed as compatible osmolytes. Osmolytes are small organic
molecules and accumulate at relatively high concentration in the intracellular surrounding that can stabilise the folded protein without disturbing any other cellular processes [8]. Osmolytes that stabilise the folded structure of protein such as carbohydrates (e.g., trehalose), methylamine (e.g., trimethylamine N-oxide (TMAO) and betaine) are often termed as chemical chaperones. TMAO counteracts the effects of water stress by allowing organisms to survive in extreme conditions [9]. For example, marine animals counteract the high pressure perturbed conditions by increasing the concentration of TMAO in their natural surroundings [10]. On the other hand, non-compatible osmolyte and waste product like urea destabilise the proteins [9, 11] and favour the unfolded conformations over the native state of proteins. Urea makes favourable interactions with the protein surface that leads to unfolding. This unfolded state is characterised by protein conformations with larger exposed surface area towards the solvent [12].

A counteracting effect of TMAO on urea was found with the most active ratio 2:1 of urea and TMAO, which is used by marine animals for stabilising the protein under various stress conditions [13, 14]. Osmoprotectant TMAO is used for protein found in human body to counteract the structural changes under deteriorated condition for inhibiting the protein aggregation. Urea is widely used in protein folding studies to denature the native state [15] while TMAO favours the native state of proteins [16]. TMAO stabilises the native structures of protein through exclusion mechanism, i.e., TMAO moves away from the vicinity of the protein first solvation shell (FSS). In order to exclude from the protein surface, TMAO is strongly attracted towards the water [17]. For the stabilisation of protein, the counteracting effects of TMAO from the denaturing effect of urea have been suggested to emerge from direct and water-mediated interactions [18]. A water-mediated interaction between TMAO and urea having hydrated and long-lived TMAO-3H2O complexes as TMAO provide shielding layers between TMAO and the urea’s donor site [19]. The opposing effects of urea and TMAO on protein stability have been investigated through experimental and theoretical methods over the last two decades. The effect of TMAO and urea on protein conformations and stability mainly depends on the concentrations of osmolytes present in the solution. It has been argued that reorganisation of the urea and water structure around the protein surface has decreased protein solubility in TMAO and helps in the compaction of protein [20]. The preferential solvation model [21] shows that water and urea choose to solvate TMAO over protein surface, and a decrease in the number of peptide-water hydrogen bonds have been observed [22]. From neutron scattering experiment, it has been proved that TMAO counteracts the urea’s deleterious effect by direct interaction of TMAO and urea without any interaction with protein [20]. Most of the researches have been done by using the TMAO Kast model based on the ab initio methods [23].

IAPP pathological aggregation is enhanced by peptide concentration, pH and temperature variation above and below critical thresholds [24]. The aggregation of insulin has taken place due to the direct interaction and copolymerisation between insulin and hIAPP [25]. Both IAPP and insulin have enough opportunity to interact due to sharing the same secretory pathway in the β-cells. Many studies have shown that insulin has the ability to inhibit the hIAPP fibrillation in vitro [26]. The fragment hIAPP20–29 easily makes amyloid fibrils as the smaller fragments hIAPP23–27 and hIAPP21–27 form. The hIAPP21–27 has been described with X-ray crystallography for showing its tendency of amyloid formation [27]. Pathological aggregation is activated by the disordered NFGAIL region of hIAPP, which makes link between β-stranded regions [5]. Some researchers have suggested that distal regions of N-terminus is nucleation for IAPP [24]. Human IAPP has the ability to form highly ordered β-sheet-rich amyloid fibrils, but rat IAPP does not form. Rat IAPP varies from the hIAPP by only 6 amino acids in the residue 18–29 “the core mutation region” [28]. Atomic force microscopy, cryo-electron microscopy, solid-state nuclear magnetic resonance (NMR) and two-dimensional infrared (2D IR) have been used for the hIAPP fibril structures [29–37]. The hIAPP fibril structures have the numerous U-shaped monomers of hIAPP longitudinally mount on top of each other to form a parallel in-register β-sheet [38]. The pathway of the fibril formation of hIAPP has not yet explained. A β-hairpin monomer conformation has been proposed to favour fibrilisation. The full-length hIAPP monomers were studied to exhibit wide variety of conformations, including a compact helix-coil structure, an unstructured coil and β-sheet structure [28]. The effect of osmolyte (TMAO) and denaturant (urea) cosolvents on the aggregation and amyloid fibril formation of hIAPP has been studied experimentally [7]. Investigating structural properties of IAPP in solution is usually challenging in wet lab, and still much of it is unknown specially the aggregation pathways of amylin. Thus, it is necessary to study the structure of hIAPP at monomeric level by using the molecular dynamic (MD) simulations.

The objective of this paper is to study the structural changes of hIAPP protein in the presence of TMAO-water, urea-water and urea-TMAO-water ternary solutions using MD simulations at different concentration of osmolytes. The study gives valuable information about the interaction of hIAPP monomer with two different osmolytes and the fragment which involve in aggregation from this protein.
Methods

Initial coordinates of the hIAPP_{1-37} protein are taken from a NMR-based structure of human amylin bound to SDS micelles [39] due to the absence of crystal structure of hIAPP. It consists of α-helix structure in between the residues 7–29. The monomeric hIAPP peptide was taken from the RCSB-Protein Data Bank (model 1 of PDB ID-2kb8). The reason of selection of this α-helical form is to see the counteraction behaviour of TMAO towards urea such that TMAO can inhibit the aggregation of hIAPP in the presence of denaturant urea. Atomistic MD simulations were employed using the GROMACS 5.1.4 and 5.1.2 software [40]. Here, we performed MD simulations to study the behaviour of amylin protein in two different osmolytes TMAO and urea as well as a ternary solution of TMAO and urea in the ratio of 1:2. Simulations of dimeric form of protein were also performed in water at 300 and 400 K temperature. Cubic simulation boxes of side ~7.8 nm were used for all the systems. Simulations were performed at pH 7.0 as each system was neutralised with 2 chloride ions. The AMBER99sb-ILDN force fields [41] for bonded and non-bonded parameters and TIP3P water model was applied to study the structural changes during the simulation in amylin protein (Fig. 1). Kast et al. [23] and Netz [42] forcefield parameters for TMAO were used. Initially, steepest descent algorithm was used for energy minimisation for 3000 steps followed by conjugate gradient method, for removing bad contacts between the solute and solvent, and then, a position restrained was performed for 1 ns for all the systems. The resultant systems were then carried for simulations under NPT ensemble for the production run.

The LINCS algorithm [43] was employed to constrain all bonds within protein, TMAO and urea. Long-range electrostatic interactions were dealt with particle-mesh Ewald (PME) method [44], and a typical 14 Å cut-off was used for van der Waals (vdW) interactions. The details of the simulated systems are given in Table 1.

The simulation temperature was controlled at 300 K using V-rescale algorithm [45] with a coupling constant 0.1 ps, and constant pressure with coupling constant 0.5 ps was maintained with Berendsen methods [46]. Newton’s equation of motion was integrated over a 2-fs time step by applying a leapfrog algorithm. Cartesian periodic boundary condition was applied in all directions. A total of 10 different MD simulations had been performed with each simulation time of 500 ns, and data was collected for every 10 ps. Only the last 200 ns of the trajectories were used for all stable analysis except for RMSD and configurational entropy calculation. The total simulation time for this study is 5 µs. Estimation of conformational sampling in space is given in the supporting information of Fig. S9 to S11.

Simulation analysis

GROMACS built-in tools were used to analyse the MD trajectories for measuring quantities such as RMSD, RMSF and spatial distribution function. The conformational changes of the protein were analysed during the MD simulations by root mean-square deviations (RMSDs) of Cα atoms. The root mean square fluctuations (RMSFs) of the residues in the protein were measured for the flexibility of the specific residue. Radius of gyration (Rg) of protein Cα atoms were analysed for the compactness of structure. The radial distribution function (g(r)) of osmolytes around the protein was calculated for the density of probability to find a particle
at distance $r$ from the reference protein. Solvent accessible surface area (SASA) of protein was determined by gmx sasa command for the global conformational sampling. The secondary structure content of protein was recognised using the DSSP [47] method by do_dssp command. Fraction of native contact is calculated if the two Cα atoms are within 7 Å of each other in the conformation at particular frame with respect to the native state. For the calculation of protein–protein hydrogen bonds, h_bond command of gromacs was used to compute the reduction in the number of hydrogen bonds during unfolding process. The hydrophobic residue contact map was used to measure the short- and long-range hydrophobic interaction. Spatial distribution functions (SDFs) of osmolytes around the protein in different system were calculated for the clearer visualisation of the 3D distribution by gmx spatial. The electrostatic surface potential of protein was calculated using the Delphi [48] program for last frame or at 500 ns of MD simulation in different environment. Configurational entropy was calculated to see protein unfolding during the simulation.

Results and discussion

Conformational changes of hIAPP in the presence of osmolytes

The RMSD values for protein with respect to time were calculated over the entire 500-ns trajectory. The RMSDs provide insight into the structural dynamics of the protein during the MD simulation. To examine the quantitative change in the simulated and the initial structures quantitatively, we have calculated the RMSD of Cα atoms of protein between them for all the systems of water, TMAO, urea and urea/TMAO at various concentrations of osmolytes as shown in Fig. 2A–C. In pure water, the RMSD of protein fluctuates at around 1.4 nm and starts showing much structural changes only at initial of simulation. Figure 2 shows that mostly all the system except three (at 4.0 and 6.0 M urea for binary system and 4.0 M urea in the case of ternary system) reaches an equilibrium after 25 ns of the simulation. A large increase in the RMSD indicates the fast divergence away from the initial conformation. In binary TMAO solution, the RMSD of hIAPP undergoes smaller fluctuation at all concentrations in comparison to urea that indicates the lesser conformational change of hIAPP and shows the protecting nature of TMAO (Fig. 2A and B). In urea solution, the RMSD of hIAPP shows largest changes initially up to 300 ns at 4.0 M, and after that, it shows similar structural changes as that of RMSD of hIAPP in water.

Human IAPP shows least conformational changes at 6.0 M urea up to 200 ns as compared to all other concentrations, and then, increase in RMSD value shows the denaturation behaviour of urea [49] (Fig. 2B). In ternary urea/TMAO solution, the RMSD at 4 M urea shows least changes as compared to hIAPP in water. While at 2.0 and 6.0 M concentration of urea in ternary system, hIAPP has equivalent RMSD values (Fig. 2C) and these RMSD values are somewhat higher than at 4.0 M urea.

We have also calculated the RMSD for binary solution of TMAO and ternary systems using the Netz model of TMAO as shown in Fig. S1. Conformational changes obtained using Netz model of TMAO shows that hIAPP shows lesser dynamical fluctuation in the presence of binary solution of TMAO. In ternary system, hIAPP shows maximum changes at 4.0 M urea same as Kast model. Dimer of hIAPP peptide shows the stability in structure as compared to monomeric form at 300 K as shown in Fig. S2. At higher temperature of 400 K, dimer shows highest RMSD as compared to monomer and dimer at 300 K.

Thus, the hIAPP protein shows maximum fluctuations in binary urea solution and least in ternary mixture at 4.0 M, and this lower divergence at 4.0 M urea in ternary system shows the protecting nature of TMAO in contrast to urea.
Structural flexibility at residue level

Since hIAPP is an IDP, it is important to capture the dynamic fluctuations of each residue of protein in different environments. Time-average RMSF of all Cα atoms of each residue of protein is plotted and is shown in Fig. 3A–C.

The RMSF plot shows that loop region residue 1–6 shows highest structural fluctuations of 1.0 to 0.4 nm in water as compared to the protein in TMAO (Fig. 3A). The average residual mobility is confined around 0.2 nm for all the residues of protein in the presence of TMAO. Among binary urea solution, protein shows highest fluctuation at 6.0 M urea in comparison to water for all the residues as shown in Fig. 3B, and lowest RMSF value at 2.0 M except for residue (31–37) which shows little increase in the residual mobility at 2.0 M urea. Maximum residual flexibility of the protein in the presence of urea at higher concentration may cause the protein to lose its secondary structure content. In ternary solution, hIAPP shows the highest residual fluctuation at 4.0 M urea. Six molar concentration of urea in ternary solution shows the lower residual mobility in comparison to binary solution of urea at 6.0 M indicating the counteraction of urea in presence of TMAO.

Thus, the lower RMSF value in the presence of TMAO indicates the lesser disruption of secondary conformation, and the higher RMSF value indicates the structural unfolding in the presence of urea. Systems having the maximum flexibility show the highest RMSF value around the N-terminal region.

Secondary structure analysis

To investigate the secondary structure change of the hIAPP for different solution, the DSSP algorithm of Kabsch and Saunder was used. For all the systems, we have shown the DSSP plot in Fig. 4.

In pure water, at the end of the simulation, there is significant conformational change in the α-helical content and deviation is observed from the initial structure having 65% α-helical content. At the end of the simulation in water, the percentage of α-helical content remaining is 9%, and on the addition of TMAO initially 2.0 M, the percentage of α-helical content remains the same as that of water at the end of the simulation (Table 2). On increasing the concentration of TMAO at 4.0 M, the α-helical content of protein increases up to 42%, and on further increasing the concentration of TMAO up to 6 M, the α-helical content decreases to 27% (Table 2). On the other hand, the percentages of the α-helical content of hIAPP confirm the loss of initial conformation and resulting denaturation in binary urea solution. At 4.0 M urea, there is formation of small β-sheets near residue (13–17) and C-terminal end (Fig. 4F).

These observations substantiate the facts that hIAPP in monomeric form can adopt a α-helical, a β-sheet or an unstructured coil conformation [50]. It also suggests that the β-hairpin structure may be an amyloidogenic precursor of human amylin protein [28]. Urea mostly shows non-monotonic behaviour at 4.0 M concentration due to the formation of β-sheets [51]. It is observed that the conformation of the hIAPP changes from helix to coil in urea as the concentration of urea increases from 2.0 to 6.0 M (Fig. 4E–G). The hIAPP is unfolded into a disordered conformation containing a bulk population of coil, bend and turn, in the binary solution of urea. TMAO can offset the denaturing effect of urea in ternary TMAO/urea solution, and some helical content reappears in these solutions. Coil content of the hIAPP is increasing with an increase in urea concentration. Netz model of TMAO also counteracts the effect of urea but less at 6.0 M urea due to the formation of β-sheets as shown in Fig. S3. Dimer at 400 K also shows the formation of β-sheets around the residues 13–17 and the C-terminal of protein as shown in Fig. S4.

Thus, the hIAPP in the presence of TMAO is capable to form α-helix, and the decrease is observed at higher concentration of urea. Human IAPP having α-helix structure during the start of the simulation has a tendency to
aggregate into β-sheet conformation at residue 13–17 and around the C-terminal at 4.0 M urea. Similar type of results were also observed by Kajava et al. [52]. They suggested a structure of hIAPP with β-strands around the residues 12–17, 22–27, and 1–37. Luca et al. [53] also proposed that the hIAPP monomer have the β-sheets with residues 8–17 and 28–37.

![Fig. 4](image)

**Table 2** Percentage of the secondary structure contents of hIAPP in water, TMAO, urea and urea/TMAO solutions

| S. no | System       | α-Helix | β-Sheet | β-Bridge | Coil | Turn |
|-------|--------------|---------|---------|----------|------|------|
| 1     | Water        | 9       | -       | 1        | 38   | 25   |
| 2     | 2M_tm        | 9       | -       | -        | 43   | 24   |
| 3     | 4M_tm        | 42      | -       | -        | 21   | 21   |
| 4     | 6M_tm        | 27      | -       | -        | 34   | 26   |
| 5     | 2M_u         | 24      | -       | -        | 26   | 32   |
| 6     | 4M_u         | -       | 5       | 1        | 54   | 14   |
| 7     | 6M_u         | -       | -       | -        | 66   | 12   |
| 8     | 2M_u+1M_tm   | 9       | -       | -        | 36   | 23   |
| 9     | 4M_u+2M_tm   | 9       | -       | 1        | 45   | 24   |
| 10    | 6M_u+3M_tm   | 1       | -       | -        | 48   | 16   |

**Solvent accessible surface area of hIAPP during the simulation**

SASA is defined as the area above which the interaction between the protein and the solvent molecules takes place. Here, we have plotted the probability distributions of SASA for different systems as shown in Fig. 5A–C. It is observed...
that TMAO-containing systems lead to a narrower distribution of SASA values in comparison to water. And with increase in concentration of TMAO, probability of slightly large value of SASA is increased at 6.0 M as the number of TMAO molecules is increased. On the other hand, probability distribution curve of SASA shifts towards the larger value of SASA as the concentration of urea increases. This increased SASA value shows the unfolding behaviours due to denaturing effect of urea. For the ternary solution, the value of SASA is decreased in comparison to the binary solution at respective concentration of urea in the presence of TMAO. This shows the counteraction behaviour of TMAO towards urea. Lower the SASA value means that the smaller number of amino acid residues is exposed to the solvent.

Effect of osmolyte concentration on the initial conformation of hIAPP

The initial structure of protein has the α-helical content, and this α-helical content is affected in the presence of denaturant urea and protectant TMAO. So, we have calculated the fraction of native contacts with respect to time as shown in Fig. 6A–C. The fraction of native contacts is explained as a list of native contact pairs (a,b), and its value varies from 0 to 1 [54]. For all-atom model heavy atom pairs a and b belong to residue \( \theta_a \) and \( \theta_b \) are in native contact when \( |\theta_a - \theta_b| > 3 \) and the distance between a and b is less than 7 Å.

More number of native contacts is formed in the presence of TMAO as compared to urea as the TMAO is a protecting osmolyte. Fraction of native contact increases with an increase in TMAO concentration up to 4.0 M, but at 6.0 M, it shows slight decrease from that at 4.0 M (Fig. 6A). On the other hand, protein in binary solution of urea shows the greater loss of initial structure as the concentration of urea increases (Fig. 6B). And in ternary solution, the fraction of native contacts of protein decreases with an increase in urea concentration, but this decrease is smaller than the binary solution of urea (Fig. 6B and C) except at 2.0 M. Protein at 2.0 M urea does not show much loss in secondary structure, as previously reported result also shows the stability of α-helical content at lower urea concentration [55]. And for higher concentration of urea, TMAO indicates the protecting nature for protein. Netz model of TMAO also shows the increase in fraction of native contacts in the presence of TMAO as shown in Fig. S5. Overall human IAPP does not lose much of its initial structure at higher concentration of TMAO as compared to water and urea using the Kast model. In most researches, the Kast model of TMAO has been used as an osmoprotectant [21, 56]. TMAO is also able to counteract the protein against urea denaturation in 2:1 ratio of urea and TMAO respectively [57].

Radial distribution function of osmolytes around the protein

To obtain the hIAPP solvation in different environment, we have calculated the radial distribution function (rdf) of oxygen atoms in TMAO (O_T) and carbonyl oxygen atoms of urea (O_U) around the backbone amide hydrogen atoms as shown in Fig. 7A–D. In binary urea solution and in ternary
urea/TMAO solution at all concentrations, we observed highest $g(r)$ value around 0.5 nm from the protein surface (Fig. 7B, D), and for TMAO, the first peak height appears at 0.55 nm (Fig. 7A) except for TMAO at 6.0 M in ternary system (Fig. 7C). The first peak height in ternary system of TMAO at 6.0 M urea is observed at 0.46 nm. The peak height for the urea is always higher than the TMAO at all concentrations. This shows the accumulation of urea in
the first solvation shell (FSS) of the protein. The increase in number of urea molecules around the protein is due to unfolding and is maintained by the increase in its surface area. The distribution of $O_T$ atoms around the backbone amide hydrogens in binary solution of TMAO at different concentrations is approximately similar. While in the case of urea, first peak height decreases with an increase in urea concentration, and this decrease around the protein indicates that density of $O_U$ is lower in FSS than in the bulk for higher concentration of urea [49, 58]. In ternary solutions, the rdf of TMAO atoms also does not show any differences in the peak height except a slight increment at 6.0 M urea, but it is moved towards the protein surface. The rdf of urea for ternary systems also shows similar behaviour as that is observed in the case of binary solutions of urea, the peak height decreases with increase in urea concentration.

Considering the hydration of protein, the Rdfs of water oxygen around the backbone amide hydrogen are calculated as shown in Fig. 8A–C. In pure water, the first peak appears at 0.46 nm, and on addition of TMAO, no difference in the peak height is observed (Fig. 8A). Whereas, on addition of urea, the peak height of $O_w$ decreases with an increase in urea concentration (Fig. 8B) which indicates the removal of water molecules from the FSS of protein. In ternary solution, the peak height of $O_w$ of water also decreases with increasing osmolyte concentration, but this decrease is smaller than the binary solution of urea.

Thus, TMAO shows the protecting nature for the hIAPP protein and counteracts the denaturation of urea by replacing the urea molecules with water.

**Preferred interaction coefficient of osmolytes around the protein**

The interaction of osmolytes with protein can be estimated by measuring the preferential interaction coefficient, described as

$$
\Gamma = -\left( \frac{\delta \mu_2}{\delta \mu_3} \right)_{c_2,T,P} = \left( \frac{\delta c_3}{\delta c_2} \right)_{\mu_3,T,P}
$$

(1)

where $\mu$ is the chemical potential, $c$ is the concentration and the subscripts 1, 2 and 3 represent water, solute and the osmolytes respectively [59]. $\Gamma$ indicates the change in the chemical potential of a solute due to the presence of osmolytes and can also be defined as the change in concentration of osmolyte for balancing the constant chemical potential when solute is added to the solution. This has been calculated using a two-domain model due to the difference in the osmolyte concentration in the local domain of the solute and the bulk solution [60, 61], i.e.,

$$
\Gamma = \langle N_{3\text{local}} \rangle - \left( \frac{N_{3\text{local}}}{N_{1\text{local}}} \right) N_{1\text{local}}
$$

(2)

where $N$ denotes the number of molecules. Negative value of $\Gamma$ indicates that the TMAO molecules are excluded from the...
FSS of protein. Positive value $\Gamma$ shows the accumulation of urea molecules with a net favourable interaction.

Here, we have calculated the preferential interaction coefficient of TMAO and urea around the protein as shown in Fig. 9A and B. In binary system, the exclusion of TMAO increases with increase in TMAO concentration and the accumulation of urea increases with increase in urea concentration. While, in ternary system, TMAO counteracts the effect of urea at 2:1 and shows less preferential interaction of urea around the protein in the presence of TMAO. This was also shown in rdf of oxygen atoms of water (Fig. 7), where peak height increases in the presence of TMAO.

**Probability distribution analysis for protein**

The probability distribution graph obtained as a function of both the end-to-end distance ($R_{ee}$) and the radius of gyration ($R_g$) as the collective variables for the hIAPP protein in water, TMAO, urea and urea/TMAO solutions using the following equation:

$$\Delta G_x = -k_B T \left[ \ln P_x - \ln P_m \right]$$

where $k_B$ is the Boltzmann constant, $T$ is the absolute temperature, $\Delta G_x$ is the free energy of coordinate $x$ for the probability distribution $P_x$ and the maximum probability is represented by $P_m$ [57]. This graph is plotted to determine the dominant native and metastable states of hIAPP and its variation during the simulation as shown in Fig. 10A–J. Structural coordinates are obtained from the lowest energy states to understand the structural changes during the simulation. Dark blue colour represents the lowest Gibbs energy states.

The occurrence time of these coordinates is then tracked and marked for obtaining the most probable configuration during the simulation. From the plot, it is observed that in pure water, the most probable state of lowest energy is losing its initial structure to some extent and its end-to-end distance and $R_g$ is the radius of gyration of protein Cα atoms. Structures in the inset are lowest energy structure obtained during the simulation of protein with different probability.

![Fig. 10 Population densities of hIAPP conformations for protein in presence of (A) water, (B–D) TMAO, (E–G) urea and (H–J) urea/TMAO solutions, where $R_{ee}$ is a measure of the peptide end-to-end distance (from centre of mass of C terminus to N terminus groups) and $R_g$ is the radius of gyration of protein Cα atoms. Structures in the inset are lowest energy structure obtained during the simulation of protein with different probability.](image-url)
also increases indicating formation of extended structure. While on the addition of TMAO, the most probable structure is moving towards the structure having high α-helical content up to 4.0 M, but at 6.0 M TMAO, the configuration obtained is slightly extended with less probability as shown in Fig. 10D. On the contrary, protein is losing some α-helical content at low 2.0 M urea concentration, and the extended conformations are obtained with increase in urea concentration. The conformations obtained at 6.0 M urea lose all its α-helical content and have probability of random coil structure at lower end-to-end distance.

On the other hand in ternary system, at 2.0 M urea, TMAO counteracts the effect of urea as the one conformation is obtained with low end-to-end distance with almost similar Rg value at all the concentration. And with increase in urea concentration for ternary system, the conformations obtained are showing higher end-to-end distance. The probability density is also increasing with two basins at 6.0 M urea with most probable conformation having some α-helical content. Hence, it is concluded that TMAO offset the effect of urea.

**Intramolecular hydrogen bond formation by the hIAPP**

Protein stability is the consequence of an adjustment between intramolecular interactions of protein and their interactions with the solvent. The formation of intramolecular hydrogen bonds (H-bonds) in protein is important for its stability. Here, we have calculated the probability distribution of intramolecular H-bonds for different systems as shown in Fig. 11A–C. The H-bond is expected to be formed if the distance between a donor (D) and an acceptor (A) is less than 3.5 Å and the D-H-A angle is greater than 150°.

In TMAO, the protein intramolecular H-bonds increase with increase in TMAO concentration as the probability distribution curve peak height shifts to a larger number of H-bonds. While in binary solution of urea, protein only shows the maximum probability for higher number of H-bonds at 2.0 M in comparison to water and the curve shifts to the lower number of intramolecular H-bonds with increase in urea concentration. On the other hand in ternary systems, probability distribution curve shifts towards the lower number of H-bonds with increase in urea concentration. This decrease in number of H-bonds is lower in comparison to binary solution of urea. Thus, this indicates that TMAO is preserving the intramolecular H-bonds of the protein.

**Hydrophobic contact map**

Not only the H-bonds in protein, but also the hydrophobic contacts can play important role in stabilisation of protein. Ability of hydrophobic residue to be buried away from water is accepted as the native state stabilising force for protein [64]. The inter-residue interactions determine the initiation of folding in the unfolded protein under folding conditions. Non-polar groups are not soluble in water, and the hydrophobic interactions between them are important to protect the protein from aqueous environment. Long-range (LR) and short-range interaction (SR) is produced if $x > 10$ and $x \leq 10$ respectively [65].

![Probability distribution curve of intramolecular H-bonds of hIAPP in (A) TMAO, (B) urea and (C) urea/TMAO solutions](image)
where \( a \) and \( b \) are amino acids in the primary structure. LR could play role in finding the tertiary structure, and SR interactions could largely contribute to the secondary structure formations [62, 63].

Here, we have monitored the effect of TMAO, urea and urea/TMAO solutions on the interaction between the hydrophobic residues; the hydrophobic contact map for all the systems was plotted as shown in Fig. 12A–K. Here, dark blue colour represents the LR hydrophobic contacts. In the presence of pure water, there is a formation of LR hydrophobic contacts around the regions 4–8/1, 12–14/1, 1/4–8 and 1/12–14. While in TMAO, there is no significant change in the case of LR contacts but few new SR interactions appeared at 2.0 M TMAO. On increasing the TMAO concentration, there is a continuous increase of LR contacts. At 6.0 M TMAO, there is formation of LR hydrophobic contacts around the regions 5–14/1–4 and 1–4/5–14. On the other hand, there is a large increase of LR hydrophobic contacts with the loss of some SR interactions at higher concentration of binary solution of urea. Urea mostly shows the SR interactions around the residue 23–27 of hIAPP. In ternary solution, there is a decrease of LR hydrophobic interactions on the addition of TMAO in comparison to binary solution of urea. This decrease of LR hydrophobic interactions shows the appearance of \( \alpha \)-helical content in protein and the counteraction of urea by TMAO.

Fig. 12 Hydrophobic contact maps of hIAPP in (A) water at 0 ns, (B) water, (C) 2 M TMAO, (D) 4 M TMAO, (E) 6 M TMAO, (F) 2 M urea, (G) 4 M urea, (H) 6 M urea, (I) 2 M urea and 1 M TMAO, (J) 4 M urea and 2 M TMAO and (K) 6 M urea and 3 M TMAO solutions.
Spatial organisation of solvent molecules around the protein

We have also calculated the spatial distribution functions (SDFs), g(r, θ, φ) which contains both the radial and angular spherical coordinates, rendering a full 3D packing arrangement of the solvent molecules around solute. According to the work by Kusalik and Svishtchey [66], we have computed the 3D local density from the outcomes of simulations. The space was binned with width of 1 Å, and its occupation was captured for each frame of the dynamics. Here, we have analysed the SDF of TMAO and urea molecules in FSS around the hIAPP protein as shown in Fig. 13A–K.

We observed that TMAO spreads more with increase in TMAO concentration from the FSS of protein and urea shows compact arrangement as compared to TMAO. And this compactness or distribution of urea decreases with increase in urea concentration which was also observed in rdf of urea (Fig. 7B), as rdf of urea has shown the decrease of urea distribution in FSS on increasing the urea concentration. In ternary system, there is also a sparseness of TMAO around the protein with increase in urea concentration and there is complete absence of TMAO molecule in FSS at 6.0 M urea. For urea in ternary system, highest distribution of urea around the protein at 2.0 M and least at 4.0 M are observed. SDF for water in FSS for all the systems has been given in Fig. S6. Water shows some more distribution in the presence of TMAO as compared to urea.

Hence, we conclude that urea shows more distribution in the FSS of protein as compared to TMAO, and the density decreases with increase in osmolyte concentration.

Electrostatic potential surface map of protein

Electrostatic interactions have important role in biological systems because biomolecules consist of atoms carrying partial charges [67]. Electrostatic potential surface (ESP) maps of hIAPP for different systems of water, TMAO, urea

![Fig. 13 Spatial distribution function of osmolyte in the FSS of protein (pink colour represent TMAO and cyan colour represent the urea molecules around the FSS of protein)](image)
and urea/TMAO solutions were generated by the Delphi software to see the hydrophobic surfaces [68] as shown in Fig. 14A–J. The Delphi program executes its internal Poisson-Boltzmann (P-B) solver for linear P-B equation to calculate surface potential which is mapped on protein surface. The parameters used for the calculation are the inner (2.0) and outer dielectrics (80.0), the probe radius which is used for how close an ion may come to the surface of a molecule is 1.4 Å, salt concentration (0) and filling percentage of cube is 80.

The hIAPP is an IDP which may undergo aggregation under unfavourable environment conditions. Aggregation process of proteins or peptides can be triggered by hydrophobic-hydrophilic interfaces [69]. Protein–protein interactions in the aggregation process can be electrostatic or hydrophobic. Protein in water has exposed mostly the hydrophobic groups (A5, A8, L12, A13, F15, A25, I26, L27, V32). While on the addition of TMAO, most of these hydrophobic groups of protein are relatively less exposed with increasing concentration of TMAO. However, all the hydrophobic groups get more exposed on the addition of urea except at low concentration of urea. At 2.0 M urea of binary solution, hIAPP does not show much change in the structure of protein as compared to initial state as shown in Fig. 6B. On the other hand in ternary system, hydrophobic groups of protein are less exposed in comparison to the binary solution of urea. This was also shown in Fig. S7A–C by calculating the difference in SASA of hydrophobic residues with respect to the residues in water for different systems. This decrease in exposure of hydrophobic groups shows the counteraction of urea by TMAO and results in burying of hydrophobic residue.

**Entropy of protein backbone and side-chain atoms during the simulation**

Schäfer et al. shows the importance of entropic term in protein folding and ligand binding. Schlitter’s [70] formula was used for the calculation of configurational entropy, which suggested an ad hoc approximation for evaluation using covariance matrix of cartesian positional coordinates.
obtained by simulation and requires only estimation of a determinant, following the equation

$$S_{Sch} = \frac{1}{2} R \ln \det \left[ 1 + \frac{kT e^2}{\hbar^2 M} \sigma \right] = \frac{1}{2} R \ln \det \left[ 1 + \frac{kT e^2}{\hbar^2 M} \sigma \right]$$

(5)

where $R$ is the molar gas constant, $k$ is the Boltzmann’s constant, $T$ is the absolute temperature, $e$ is Euler’s number, $\hbar$ is the Planck’s constant divided by $2\pi$ and $M$ is the mass matrix having masses of the atoms on the diagonal and the entire off-diagonal elements equivalent to 0. $\sigma$ is the covariance matrix of 3 N coordinates, and $N$ is the number of considered protein atoms. The covariance matrix elements $\sigma$ were given by,

Fig. 15 Normalised configurational entropies of the backbone atoms of hIAPP in (A) TMAO, (B) urea and (C) urea/TMAO solutions at different concentration.

Fig. 16 Normalised configurational entropies of the side-chain atoms of hIAPP in (A) TMAO, (B) urea and (C) urea/TMAO solutions at different concentration.
\[ \sigma_{ij} = \langle (x_i - \bar{x}_i)(x_j - \bar{x}_j) \rangle \]  

(6)

where \( x_i \) are the cartesian coordinates obtained after the least squares superposition of the trajectory configuration with respect to a specific subset of atoms.

Here, we have calculated the normalised configurational entropy of backbone and side-chain atoms of protein during the whole simulation time for different systems of water, TMAO, urea and urea/TMAO mixtures at various concentrations of osmolytes as shown in Figs. 15A–C and 16A–C. As the backbone and side-chain atoms have different numbers of atoms, here we have calculated the normalised entropy by dividing the entropy values with the number of atoms. The side-chain atom entropy is greater than that of backbone atom entropy. For all the systems, entropies have become constant after 200 ns of simulation. In TMAO solutions, both backbone and side-chain entropies are decreasing with an increase in TMAO concentration. While in binary urea solution, these increases with an increasing urea concentration and its value is high in comparison to TMAO. At 2.0 M urea, it shows only slight increase from that observed in 2.0 M TMAO and shows less structural changes, as observed from the RMSD (Fig. 2B). For the ternary system, lowest backbone entropy for 2.0 M and highest at 4.0 M urea have been observed. Highest backbone entropy value at 4.0 M shows the maximum structural changes during the simulation as seen from RMSD (Fig. 2C). And in ternary system, these backbone entropies are lower in comparison to the binary solution at respective concentrations of urea and show the counteraction properties of TMAO for urea at 2:1 ratio. There is a decrease in backbone entropies of ternary system in the presence of TMAO, whereas the change is not observed in the side chain entropies from that of binary solution of urea. Netz model of TMAO also shows the decrease in absolute configurational entropy of backbone atoms of protein in the presence of TMAO as compared to urea as shown in Fig. S8A–B.

Configurational entropy of backbone and side-chain atoms of protein is lower in the presence of TMAO as compared to urea, suggesting the protein folding in TMAO. Folded protein has lower entropy as compared to the unfolded one, and this loss in entropy upon folding is compensated by the solvent [71].

**Conclusion**

In this study, the effect of water, TMAO, urea and urea/TMAO (2:1) ratio of different concentrations on the human IAPP protein through all-atom MD simulation have been investigated. Through the different analysis, we observed that TMAO stabilises the protein in its initial state and also shows the counteraction property towards the denaturing effect of urea. Urea gives the unfolded structure of protein except at 4.0 M urea due to the formation of \( \beta \)-sheets. Human IAPP in its monomeric may have \( \beta \)-sheets around the residue 13–17 and at the C-terminal region, which may cause the aggregation of peptides through these residues. This study gives the evidences of the formation of \( \beta \)-sheets in IDP amylin protein in the worst conditions of environment and may enhance the aggregation.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00894-022-05180-1.

**Acknowledgements** The authors acknowledge the facility from the NPSF at CDAC, Pune, India, for providing the supercomputer facility to carry out the MD simulations. They also acknowledge UGC for providing financial support as JRF and the Department of Chemistry, AMU, through UGC-assisted DRS-SAP, DST-FIST, and DST PURSE Programme for carrying out the above research work.

**Author contribution** Ashma Khan (design, data acquisition, analysis and draft of manuscript), Israt Jahan (critical revision) and Shahid M. Nayeem (supervision, design and finalisation of MS).

**Code availability** Not applicable.

**Declarations**

**Ethics approval** The submitted work is original and have not published elsewhere in any form.

**Consent to participate** Consent taken from all authors.

**Consent for publication** Consent taken from all authors.

**Conflict of interests** The authors declare no competing interests.

**References**

1. Wright PE, Dyson HJ (2015) Intrinsically disordered proteins in cellular signalling and regulation. Nat Rev Mol Cell Biol 16:18–29
2. Uversky VN, Oldfield CJ, Dunker AK (2008) Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu Rev Biophys 37:215–246
3. Westermark P, Wernstedt C, Wilander E, Hayden DW, O’Brien TD et al (1987) Amyloid fibrils in human insulinoma and islets of Langerhans of the diabetic cat are derived from a neuropeptide-like protein also present in normal islet cells. Proc Natl Acad Sci USA 84:3881–3885
4. Wu C, Shea JE (2013) Structural similarities and differences between amyloidogenic and non-amyloidogenic islet amyloid polypeptide (IAPP) sequences and implications for the dual physiological and pathological activities of these peptides. PLoS Comput Biol 9:e1003211
5. Levine ZA, Teranishi K, Okada AK, Langen R, Shea JE (2019) The mitochondrial peptide humanin targets but does not
11. Li S, Micic M, Orbulescu J, Whyte JD, Leblanc RM (2012) Human islet amyloid polypeptide at the air-aqueous interface: a Langmuir monolayer approach. J R Soc Interface 9:3118–3128

26. Susa AC, Wu C, Bernstein SL, Dupuis NF, Wang H et al (2014) Defining the molecular basis of amyloid inhibitors: human islet amyloid polypeptide-insulin interactions. J Am Chem Soc 136:12912–12919

27. Buchanan LE, Dunkelberger EB, Tran HQ, Cheng PN, Chiu CC et al (2013) Mechanism of IAPP amyloid fibril formation involves an intermediate with a transient beta-sheet. Proc Natl Acad Sci U S A 110:19285–19290

28. Dupuis NF, Wu C, Shea JE, Bowers MT (2009) Human islet amyloid polypeptide monomers form ordered beta-hairpins: a possible direct amyloidogenic precursor. J Am Chem Soc 131:18283–18292

29. Soong R, Brendler JR, Macdonald PM, Ramamoorthy A (2009) Association of highly compact type II diabetes related islet amyloid polypeptide intermediate species at physiological temperature revealed by diffusion NMR spectroscopy. J Am Chem Soc 131:7079–7085

30. Bedrood S, Li Y, Isas JM, Hegde BG, Baxa U et al (2012) Fibril structure of human islet amyloid polypeptide. J Biol Chem 287:5235–5241

31. Goldsbury C, Goldie K, Pellaud J, Seelig J, Frey P et al (2000) Amyloid fibril formation from full-length and fragments of amylin. J Struct Biol 130:35–42

32. Hoppener JWM, Hendriksen LJ, Zeegers MP, van Noorden S, Koper LH (1999) Modeling islet amyloid polypeptide interactions with beta-sheet structure. J Am Chem Soc 121:3633–3637

33. Katabuse H, Obara K, Matsumoto T, Hori H, Oda Y (1994) The relationship between the fibrillation of islet amyloid polypeptide and the development of type II diabetes mellitus. J Bacteriol 176:2633–2638

34. Kimura K, Hori H, Katabuse H, Obara K, Matsumoto T (1995) The relationship between the fibrillation of islet amyloid polypeptide and the development of type II diabetes mellitus. J Bacteriol 177:2633–2638
47. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637

48. Li L, Li C, Sarkar S, Zhang J, Witham S et al (2012) DelPhi: a comprehensive suite for DelPhi software and associated resources. BMC Biophys 5:9

49. Paul S, Paul S (2015) Investigating the counteracting effect of trehalose on urea-induced protein denaturation using molecular dynamics simulation. J Phys Chem B 119:10975–10988

50. Singh S, Chiu CC, Reddy AS, de Pablo JJ (2013) alpha-helix to beta-hairpin transition of human amylin monomer. J Chem Phys 138:155101

51. Khan A, Nayeem SM (2020) Effect of TMAO and urea on dimers and tetramers of amyloidogenic heptapeptides ((23) FGAILSS(29)). ACS Omega 5:26986–26998

52. Kajava AV, Aebi U, Steven AC (2005) The parallel superpleated beta-structure as a model for amyloid fibrils of human amylin. J Mol Biol 348:247–252

53. Luca S, Yau WM, Leaman R, Tycko R (2007) Peptide conformation and supramolecular organization in amylin fibrils: constraints from solid-state NMR. Biochemistry 46:13505–13522

54. Best RB, Hummer G, Eaton WA (2013) Native contacts determine protein folding mechanisms in atomistic simulations. Proc Natl Acad Sci U S A 110:17874–17879

55. Gull N, Sen P, Kabir UdD, Khan RH (2007) Effect of physiologically concentrated urea on the conformation of human serum albumin. J Biochem 141:261–268

56. Paul S, Patey GN (2007) The influence of urea and trimethylamine-N-oxide on hydrophobic interactions. J Phys Chem B 111:7932–7933

57. Borgohain G, Paul S (2016) Model Dependency of TMAO’s counteracting effect against action of urea: Kast model versus osmotic model of TMAO. J Phys Chem B 120:2352–2361

58. Canchí DR, Paschek D, Garcia AE (2010) Equilibrium study of protein denaturation by urea. J Am Chem Soc 132:2338–2344

59. Timasheff SN (2002) Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. Proc Natl Acad Sci U S A 99:9721–9726

60. Parsegian VA, Rand RP, Rau DC (2000) Osmotic stress, crowding, preferential hydration, and binding: a comparison of perspectives. Proc Natl Acad Sci U S A 97:3987–3992

61. Record MT Jr, Anderson CF (1995) Interpretation of preferential interaction coefficients of nonelectrolytes and of electrolyte ions in terms of a two-domain model. Biophys J 68:786–794

62. Gromiha MM, Selvaraj S (1999) Influence of medium and long range interactions in protein folding. Prep Biochem Biotechnol 29:339–351

63. Anderson JS, Scheraga HA (1982) Effect of short- and long-range interactions on protein folding. J Protein Chem 1:281–304

64. Malleshappa Gowder S, Chatterjee J, Chaudhuri T, Paul K (2014) Prediction and analysis of surface hydrophobic residues in tertiary structure of proteins. ScientificWorldJournal 2014:971258

65. Sengupta D, Kundu S (2012) Role of long- and short-range hydrophobic, hydrophilic and charged residues contact network in protein’s structural organization. BMC Bioinformatics 13:142

66. Kusalik PG, Svirshchev IM (1994) The spatial structure in liquid water. Science 265:1219–1221

67. Gilson MK, Rashin A, Fine R, Honig B (1985) On the calculation of electrostatic interactions in proteins. J Mol Biol 184:503–516

68. Li L, Li C, Sarkar S, Zhang J, Witham S, Zhang Z, Wang L, Smith N, Petukh M, Alexov E (2012) DelPhi: a comprehensive suite for DelPhi software and associated resources. BMC Biophys 5:9

69. Lu Q, Tang Q, Xiong Y, Qing G, Sun T (2016) Protein/peptide aggregation and amyloidosis on biointerfaces. Materials (Basel) 9:740

70. Schüttler J (1993) Estimation of absolute and relative entropies of macromolecules using the covariance matrix. Chem Phys Lett 215:617–621

71. Schafer H, Daura X, Mark AE, van Gunsteren WF (2001) Entropy calculations on a reversibly folding peptide: changes in solute free energy cannot explain folding behavior. Proteins 43:45–56

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.