TMAO: protecting proteins from feeling the heat

Mayank Boob[1], Shahar Sukenik[2,7], Martin Gruebele[*,1-3] and Taras V. Pogorelov[*,1,2,4,5,6]

[1]Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States.

[2]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States.

[3]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States.

[4]School of Chemical Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States.

[5]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States.

[6]National Center for Supercomputing Applications, University of Illinois at Urbana–Champaign, Urbana, IL 61801

[7]Present address: Department of Chemistry and Biochemistry, UC Merced, Merced, CA 95344, United States.

*To whom correspondence may be addressed: Email: mgruebel@illinois.edu or pogorelo@illinois.edu,

Abstract
Osmolytes are ubiquitous in the cell and play an important role in controlling protein stability under stress. The natural osmolyte trimethylamine N-oxide (TMAO) is used by marine animals to counteract the effect of pressure denaturation at large depths. The molecular mechanism of TMAO stabilization against pressure and urea denaturation has been extensively studied, but unlike the case of other osmolytes the ability of TMAO to protect proteins from high temperature has not been quantified. To reveal the effect of TMAO on folded and unfolded protein ensembles and the hydration shell at different temperatures, we study a mutant of the well-characterized, fast-folding model protein B (PRB). We carried out >190 μs in total all-atom simulations of thermal folding/unfolding of PRB at multiple temperatures and concentrations of TMAO. The simulations show increased thermal stability of PRB in presence of TMAO. Partly structured, compact ensembles are favored over the unfolded state. TMAO forms two shells near the protein: an outer shell away from the protein surface has altered hydrogen bond lifetimes of water molecules and increases hydration of the protein to help stabilize it; a less-populated inner shell with opposite TMAO orientation closer to the protein surface binds exclusively to basic side chains. The cooperative co-solute effect of the inner and outer shell TMAO has a small number of TMAO molecules 'herding' water molecules into two hydration shells at or near the protein surface. The stabilizing effect of TMAO on our protein saturates at 1 M despite higher TMAO solubility, so there may be little evolutionary pressure for extremophiles to produce higher intracellular TMAO concentrations, if true in general.

Significance Statement

Osmolytes are key adaptors of life in extreme conditions. The osmolyte tri-methyl amine oxide (TMAO) has been studied extensively as a protectant against high pressure, but not against high temperature. Our extensive all-atom molecular dynamics analysis of temperature stabilization of protein B by TMAO shows that the effect of TMAO on the water structure around the protein saturates at low concentration, offering a possible explanation why extremophile organisms do not utilize higher concentrations of TMAO although it has the necessary solubility. TMAO molecules bind very specific amino acids on the protein surface, and TMAO molecules in a shell further away from the protein ‘herd’ water molecules to enhance protein stability.

Introduction

Osmolytes are small organic molecules that help maintain homeostasis in the cell. Osmolytes can be broadly divided into their effects on water: chaotropes ('structure breakers') and kosmotropes ('structure makers') (1). The mechanism of stabilization or destabilization of protein structure by these osmolytes depends on their interaction with the hydrogen bond network of water and also their direct interaction with the protein (2). Some kosmotropes are important for protein stability *in vivo* (3). For example, the concentration of trimethylamine N-oxide (TMAO) increases almost linearly with oceanic depth in the bodies of fishes, up to about 0.8 M, indicating its role in stabilizing proteins against effects of hydrostatic pressure (4, 5).

TMAO is characterized by a partial charge separation with the nitrogen and oxygen at the polar end, and its methyl groups at the nonpolar end (Figure 1, top right). The oxygen typically forms three hydrogen bonds with water at normal pressure, and up to four such 'hydrophilic bonds' at high pressure (6), whereas the methyl groups interact with the oxygen of water by 'hydrophobic bonding' (7–9). While the hydrophilic bonds have a very distinct spectroscopic signature in the terahertz region, the hydrophobic bonds look very similar to bulk water (10). Amphipathic TMAO reduces self-interaction because aligning of two N-O dipoles requires intervening methyl groups, and increases solubility in water as well.

The mechanism used to explain the stabilization of proteins under pressure by TMAO is the 'osmophobic effect,' whereby kosmotropes such as TMAO are excluded from the protein surface, while chaotropes, such as urea, preferentially bind to the backbone and side chains (11–13). In this case, stabilization of protein structure is achieved indirectly through altered protein surface hydration. Experiments and simulations show that the extraordinarily high osmotic coefficients of TMAO compared to other solutes like urea and NaCl indeed lead to exclusion from the protein surface (14, 15). Residence times of TMAO are shorter near proteins than those of urea (16), and yet TMAO stabilizes proteins against urea denaturation in a roughly 1:2 ratio, i.e. 1 M TMAO is enough to protect the protein against denaturation by 2 M Urea (17, 18). Recent analysis of experiments and simulations with calibrated force fields nonetheless reveals that TMAO-protein interaction also occurs, in contrast to kosmotropes such as glycine that obey the 'osmophobic effect' almost perfectly (19–21).

While there is developing consensus on how TMAO resists chemical and pressure denaturation, the interaction of TMAO and water at higher temperature has not been studied in detail. The unfolding of proteins due to temperature is distinct from chemical and pressure denaturation (22). Proteins undergo both cold and heat denaturation (23): due to loss of hydrophobic interactions at low temperature (24) and increased configurational entropy of the chain at high temperature (25). Marine organisms experience temperature gradients along with high pressure (26), and the concentration of TMAO varies with temperature in smelt fish plasma even at constant pressure, perhaps as protection against cold denaturation (27).

Here we use over 190 μs of unbiased all-atom molecular dynamics simulations to understand the molecular-level effects of TMAO on the temperature denaturation of a protein. We study Protein B (PRB), a small albumin-binding domain consisting of 47 amino acids (28), mutated such that its three alpha helices fold fast, on the microsecond time scale (29). This makes PRB ideal for probing its interactions with TMAO on the currently accessible simulation timescales (16, 30, 31). Previous studies have shown that the folding timescale of PRB is around 4 μs in both experiments and

simulations (29, 32). Performing all-atom microsecond simulations at different temperatures and TMAO concentrations (12 conditions modeled for ~16 μs each), enables us to see how TMAO affects water structure and dynamics, interacts with the protein at different temperatures and different TMAO concentrations, and how the stabilizing effect saturates between 1 and 2 M TMAO concentration.

Methods

Simulation setup. All-atom simulations of PRB were initially carried out using the GROMACS 5.1.2 simulation package and GROMACS 2020.3 along with GROMACS 2022.3 was used for carrying out additional runs to support our results (33). Four different temperatures (340 K, 350 K, 360 K and 370 K) and three different concentrations of TMAO (0 M, 1 M, and 2 M) were chosen for a total of twelve different simulation conditions. For each of the simulations, a structure of PRB (PDB ID 1PRB) was mutated to match the fast-folding sequence shown below using the Mutator Plugin in VMD (34). The simulations were conducted in the NVT ensemble, with a 2.5 fs time step, and a 1 nm cutoff for Lennard-Jones interactions. Electrostatic forces were calculated using the particle-mesh Ewald method (35). Temperature was held constant using a Nosé-Hoover thermostat (36, 37). Frames were saved every 100 ps and analyzed using the methods described below. The CHARMM36 force field/ TIP3P water combination (38) was chosen to carry out the simulations, so the simulations may be compared with equally extensive simulations of PrB in aqueous solution (29). The protein was solvated with TIP3P water (39) in a periodic box of size 57.36 × 57.36 × 57.36 $Å^3$. NaCl (2 Na^+ and 1 Cl^- ions) was used to balance out charges and standard protonation states were assumed for all the amino acid side chains. Additionally, we ran five 100 ps runs each to capture water dynamics at 360K and all concentrations of TMAO (total 15 simulations) using GROMACS 2020.3. We also ran eight 250 ns simulations at 370K to get additional folding and unfolding statistics in presence of TMAO (total of 24 simulations) using GROMACS 2022.3.

Protein Sequence (PDB ID: 1PRB): LKNAIEDCIA ELKKAGITSD FVFNAWNKAK TVEEVNALVN EILKAHA

TMAO Force field. We started simulations by choosing three different force fields from Kast *et al.* (40), Netz *et al.* (41) and one generated by us using *CGENff* (42). The difference between Kast and *CGENff* was minimal, so we simulated further only with the *CGENff* and Netz force fields. The major difference between *CGENff* and Netz is the distribution of partial charges, which could potentially change the interaction between the protein side chains and TMAO. From a simulation of 5 μs of PRB with 1 M TMAO with the two different force fields at 340 K, we could not find any major differences in direct protein-TMAO interactions (e.g. basic side chains at the protein surface) or effect on H-bonding of TMAO with water, so we chose our own force field based on the widely-used *CGENff* for the bulk of our simulations.

Analysis. *Python 3.7* (43) and the molecular dynamics analysis package *MDTraj* (44) were used for analysis. Volume calculations were performed in *ProteinVolume* (45). Solvent accessible surface area (SASA) calculations were done in *VMD* (34) with a probe radius of 1.4 Å. They were normalized according to equation 1 with the SASA of the PDB structure. The pairwise distance measurements between two atoms for calculating radial distribution function were done in *MDTraj*, along with assignment of the nearest atom in contact. The volume occupied by shells around the protein was done with an in-house implementation of Monte Carlo method in *VMD*.

$$SASA_{norm} = \frac{SASA(t) - SASA_{PDB}}{SASA_{PDB}} \quad (1)$$

Fraction of Native contacts (Q). The determination of folded and unfolded ensembles was done by assigning a value *Q* between 1 and 0 to each timeframe. Here 1 represents a completely folded structure, while 0 represents a completely unfolded structure. This value was calculated according to equation 2 as defined previously (46, 47) to indicate the number of native contacts preserved compared to a native state. The native state was determined by clustering the trajectory using RMSD by the *gmx_cluster* command in GROMACS (33) separated by an RMSD difference of at least 4 Å. The largest cluster in the 340K 0M TMAO trajectory was designated as the folded ensemble and native contacts maintained for at least eighty percent of the trajectory were used to calculate the *Q* value for all the trajectories. Timeframes with a *Q* value between 1 to 0.8 were defined as the folded structures, while values between 0.2 to 0 were defined as the unfolded structures, to determine folded and unfolded ensembles. The value of α and β were set to be 30 and 1.4 respectively to optimize for the difference between the folded and unfolded state.

$$Q(t) = \frac{1}{N}\sum_{i,j}\frac{1}{1+\exp\left\{\alpha\left[d_{ij}(t)-\beta d_{ij}^{0}\right]\right\}}, \quad (2)$$

where d_{ij}^{0} is the averaged native contact distance, α and β are smoothing factors.

Hydrogen bond lifetimes and diffusion rates. The hydrogen bond lifetimes of water around the protein were calculated using an in-house code written with MDTraj (44) package. The half-lives of the bonds were calculated by fitting the autocorrelation curves to either sum of two exponentials or three exponentials depending on whether it was continuous or intermittent hydrogen bonds. This is discussed in detail in SI. The diffusion rate was calculated using the mean square displacement module in *MDAnalysis* package (48, 49) and fitting the result with the equation 3.

$$MSD = 6Dt, \quad (3)$$

where MSD is the mean square displacement, D is the diffusion coefficient and t is time.

Results

A low concentration threshold of TMAO stabilizes protein against thermal denaturation. We simulated the PRB-TMAO-water system using a CGenFF-derived force field for TMAO (41, 42) after benchmarking three force fields for 5 μs each (see Methods). The fraction of native contacts (*Q*) is a well-established measure to quantify protein folding/unfolding dynamics as a function of temperature *T* and TMAO concentration, (46). To this end we plot the timeseries for all temperatures and concentrations of TMAO used in our simulations (Fig. 1A). The *Q* values for folded structures are defined to be the range from 1 to 0.8, indicating close resemblance to the native state as defined by clustering of the 340 K, 0 M TMAO simulation. Similarly, the range of 0.8 to 0.2 is defined as structurally intermediate ensemble(s), while 0.2 to 0 is classified as the unfolded ensemble. On average, *Q* decreases as the temperature increases in all simulations, most prominently in the absence of TMAO, where there is almost no folded structures present at 370 K. This highlights that the protein is almost completely unfolded at this temperature. In the presence of TMAO, long-lived folded states are observed even at higher temperatures. However, the stabilization due to TMAO is similar both at 1 M and 2 M TMAO, with similar dwell times in folded and unfolded states observed. This trend is also consistent with other structural metrics for unfolding such as solvent accessible surface area (SASA discussed next), root mean square deviation (RMSD), and radius of gyration (R_g) (Figs. S1-S3). By these particular measures, the effect of TMAO has already saturated at 1 M TMAO, similar to the highest concentrations observed in marine fish (4).

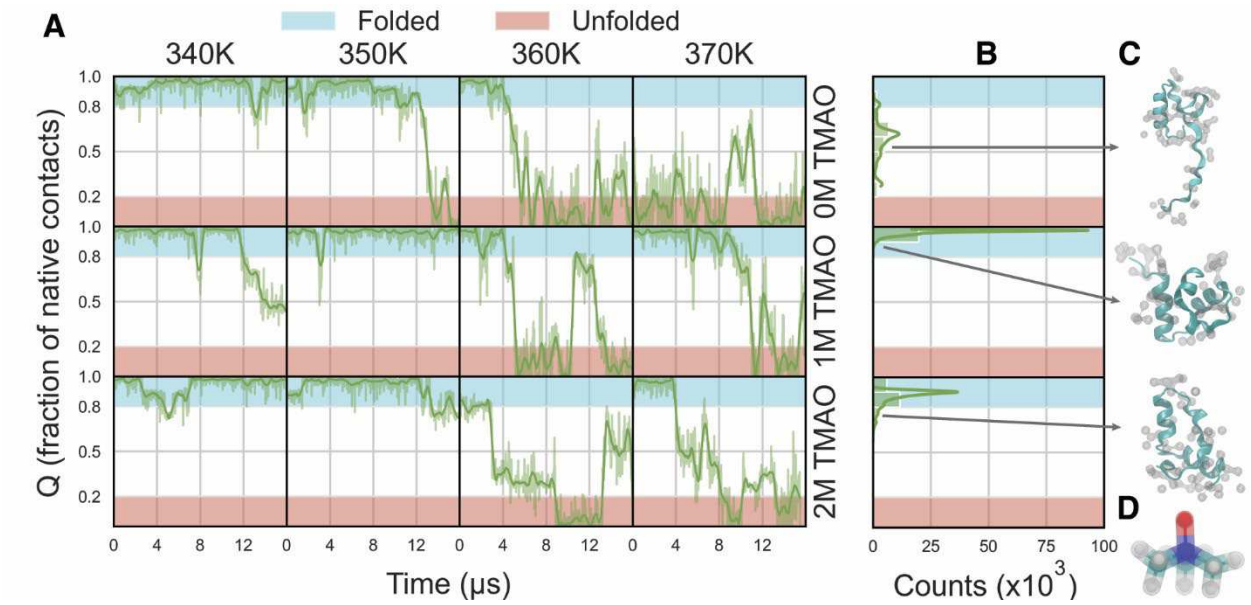

Fig. 1. (A) Fraction of native contacts (Q) for all trajectories with different concentrations of TMAO (0 M, 1 M, and 2 M) and at different temperatures (340 K, 350 K, 360 K and 370 K) are shown as moving average (solid green line) over 101 frames (101 ns) with a 3rd degree polynomial and using a Savitzky-Golay filter. The light blue band indicates the Q values for the folded state, while the light red band indicates the values for the unfolded state defined according to *Q*. (B) The histograms (right) are binned at 370K temperature across eight 250 ns runs (seeded across the entire length of the 16μs trajectories) for a particular concentration of TMAO. (C) Representative structure snapshots show the arrangement of water (grey) within 3 Å around the protein (teal) at different Q values binned in the histogram. (D) The stick figure shows a TMAO molecule structure - oxygen (red), nitrogen (blue), carbon (light blue), and hydrogen (grey).

Increased TMAO concentration leads to non-native states with lower solvent exposure. For all concentrations of TMAO, increased thermal unfolding of PRB leads to an increase in the population of unfolded states with larger SASA, as expected (Fig. 2). There is a large spread of SASA in the unfolded ensemble, indicating extensive solvation and desolvation during structural fluctuations of unfolded PRB. As TMAO concentration increases from 1 to 2 M, the fluctuations give rise to ensembles with smaller Q (unfolded-like) and SASA (folded-like), indicating the presence of non-native, collapsed structures with SASA similar to that of the folded protein. Presence of these compact non-native dried conformations has also been reported in our study of a fragment of λ-repressor (50). At high temperature, the protein's potential of mean force $-k_BT\ln(P)$, where P is the probability distribution in Fig. 2 has at least one additional minimum leading to metastable states with an ~2 μs lifetime (Fig. 1A) TMAO promotes population of multiple compact intermediate states over the unfolded state.

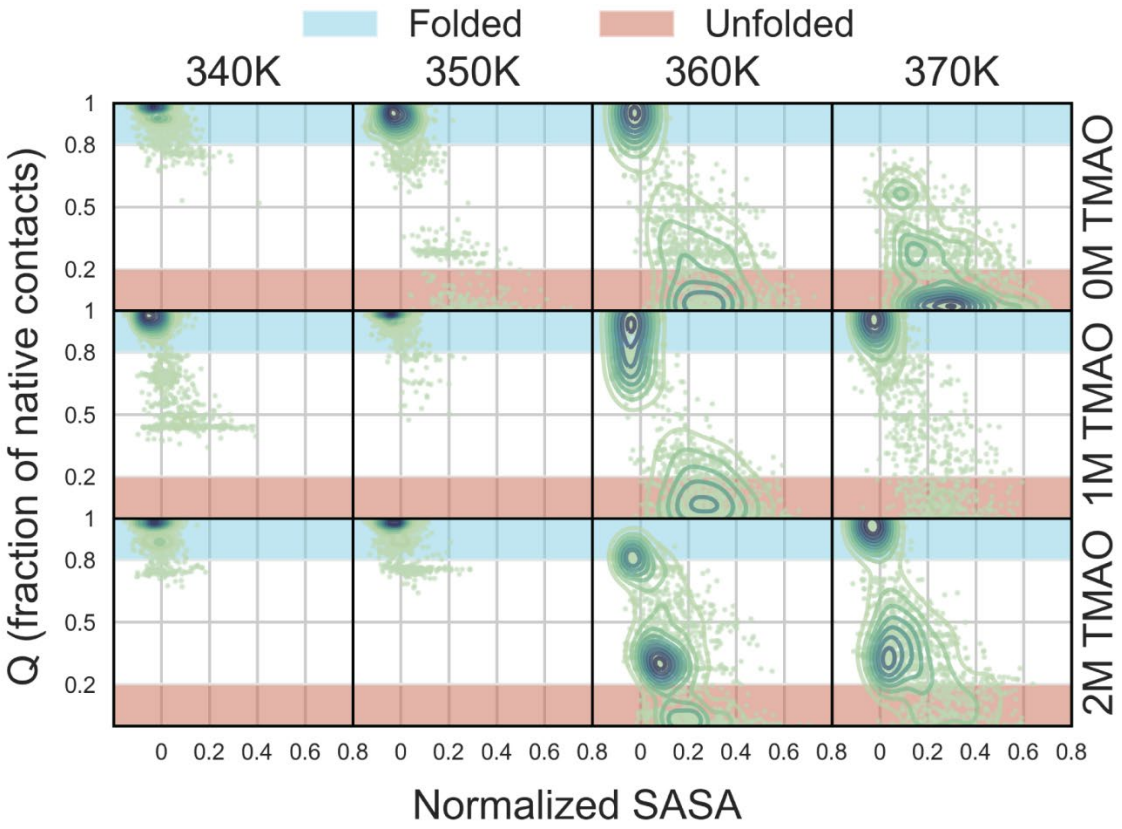

Fig. 2. Fraction of native contacts (*Q*) vs. Normalized solvent-accessible surface area (SASA) plotted as a scatter plot with kernel density estimator iso-probability contours to indicate the number of distinct populations for protein B between folded to unfolded state at different concentrations of TMAO and different temperatures. SASA is normalized compared to the SASA of the PDB structure by equation (1) used for the simulations. The light blue band indicates when the protein is in the folded state, while the light red band indicates protein in the unfolded state defined according to *Q*. In TMAO, states with residual native structure are favored over the unfolded state even at very high temperature.

TMAO is excluded from the protein surface unless it interacts with basic side chains. The radial distribution function of 1 M TMAO at 360 K, measured from the nearest protein residue, shows how TMAO organizes around the protein and influences the organization of water molecules around the protein near the thermal denaturation midpoint (Fig. 3). TMAO distributes over an 'inner' and 'outer' shell. The majority of TMAO lies in the outer shell between 4 to 6 Å (nitrogen) or 5 to 7 Å (oxygen), as revealed by the area under the $g(r)$ curve, which is over 4.5 times larger for the outer than the inner shell (Fig. 3A). The local concentration of TMAO in 2 M solution doubles in each of the shells (Fig. S6), so there is no preferential concentration-dependent population of shells.

As PRB becomes more unfolded and solvent exposed, the outer shell is more highly populated relative to the inner shell (Fig. S5 and Fig. S7). The TMAO interactions with the three protein conformational ensembles (folded, intermediate, unfolded) indicate that the osmolyte does not interact preferentially with any of the protein ensembles because the increased amount of TMAO in the unfolded and intermediate ensembles compared with the folded ensemble (Fig. S5) can be accounted for simply by the increase in the solvent accessible surface area of unfolded and intermediate structures.

In the outer shell, the TMAO oxygen atom is more likely to point away from the protein (Fig. 3A), and the methyl groups are more likely to point towards the protein surface. To quantify whether this shell of TMAO molecules is evenly distributed over the protein surface or prefers certain side chains, we plotted a histogram of normalized count of TMAO nearest to each side chain (Fig. 3C and 3D). TMAO molecules in the outer shell have no preferential water-mediated interactions with any particular type of residue on the protein surface (Fig. 3D).

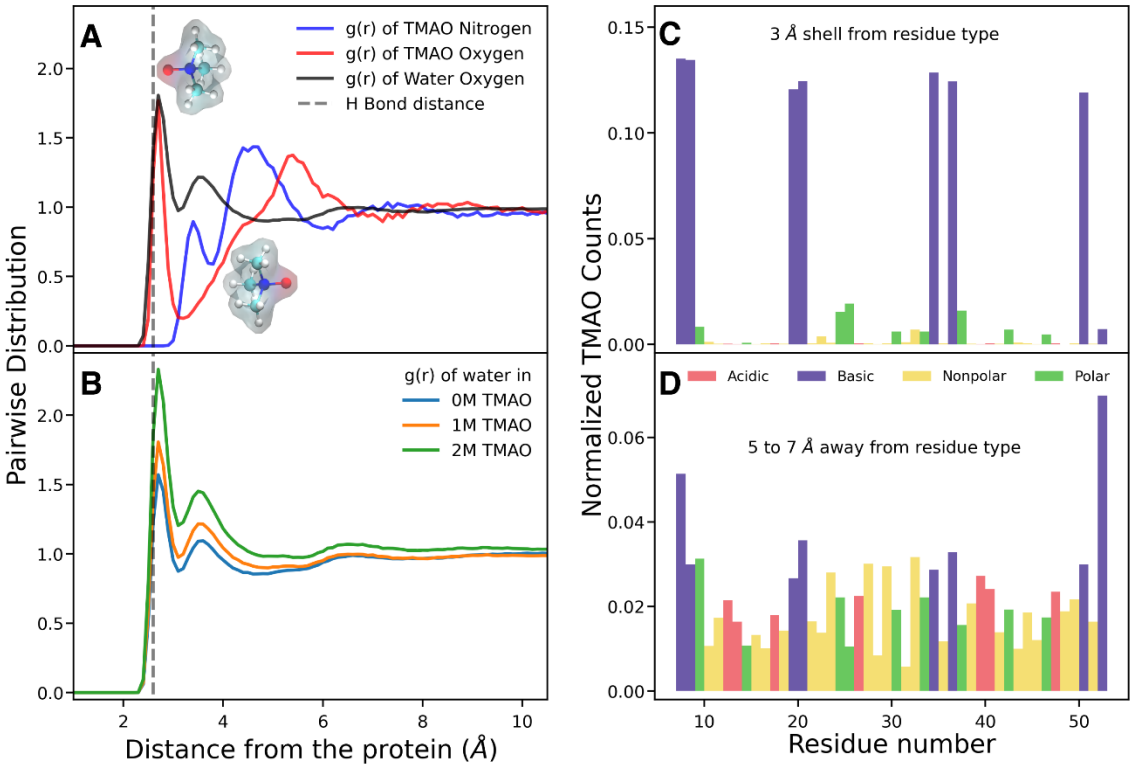

Fig. 3. Water and TMAO interaction with PRB for the simulation at 360K and 1 M TMAO. (A) The radial distribution function $g(r)$ is plotted versus distance from the protein surface (excluding hydrogens) to TMAO and water. To highlight the orientation of the TMAO towards the protein, two different distances are shown, from the nitrogen (blue) and the oxygen (red) atoms. The dashed line indicates a typical hydrogen bond distance from the protein surface. The images of TMAO molecules (inserts) are displaced and oriented towards the deduced mode of interaction that accounts for the visible peak. The area under the curve for first hydration (1 Å to 3 Å) is 4.3 while the area under the second hydration (5 Å to 7 Å) is 20.1 for the red curve (TMAO Oxygen). (B) Pairwise distance distribution of water from the PRB protein surface for the simulations at 360 K and 0 M (light blue), 1 M (orange) and 2 M (green) TMAO. The presence of TMAO pushes more water in the hydration shell of protein. (C) The first TMAO peak at 2.7 Å in (A) shows a normalized contact count that favor basic residues strongly over all others. (D) The normalized count of residues nearest TMAO in the second peak at 5 to 6 Å in (A) shows that TMAO evenly distributed over the protein surface at that distance.

In contrast, the inner shell of TMAO at 2.7 Å (Fig. 3A) reveals an additional and very specific interaction of TMAO with the protein surface side chains. To quantify this interaction, we again plotted a normalized count of TMAO nearest to each side chain, this time for TMAO in the inner shell (Fig. 3C). The partial negative charge of the oxygen atom of TMAO avoids hydrophobic, polar, and acidic side chains, interacting instead almost exclusively with the basic side chains of the protein. Indeed, the experimental transfer free energies of side chains into 1 M TMAO from water are the lowest (most negative) for positively charged residues (51). In an agreement with a spectroscopic study

(52), we observe that the methyl group of TMAO molecules in this inner shell tend to point away from the surface of the protein—opposite to the orientation of the molecules in the outer shell (Fig. 3A). TMAO is excluded from the protein surface mainly near hydrophobic side chains, in agreement with the exclusion measured by X-ray scattering (53) and spectroscopy (54).

Table 1. Area under the curve for Fig. 3A showing the approximate number of TMAO molecules in shell in absolute units

Shell	1 M TMAO	2 M TMAO
1 to 3 A (Oxygen) First peak	4.31	9.00
3 to 4 A (Nitrogen)	5.18	10.45
5 to 6 A (Oxygen)	10.59	19.32
5 to 7 A (Oxygen) Second peak	20.06	35.80

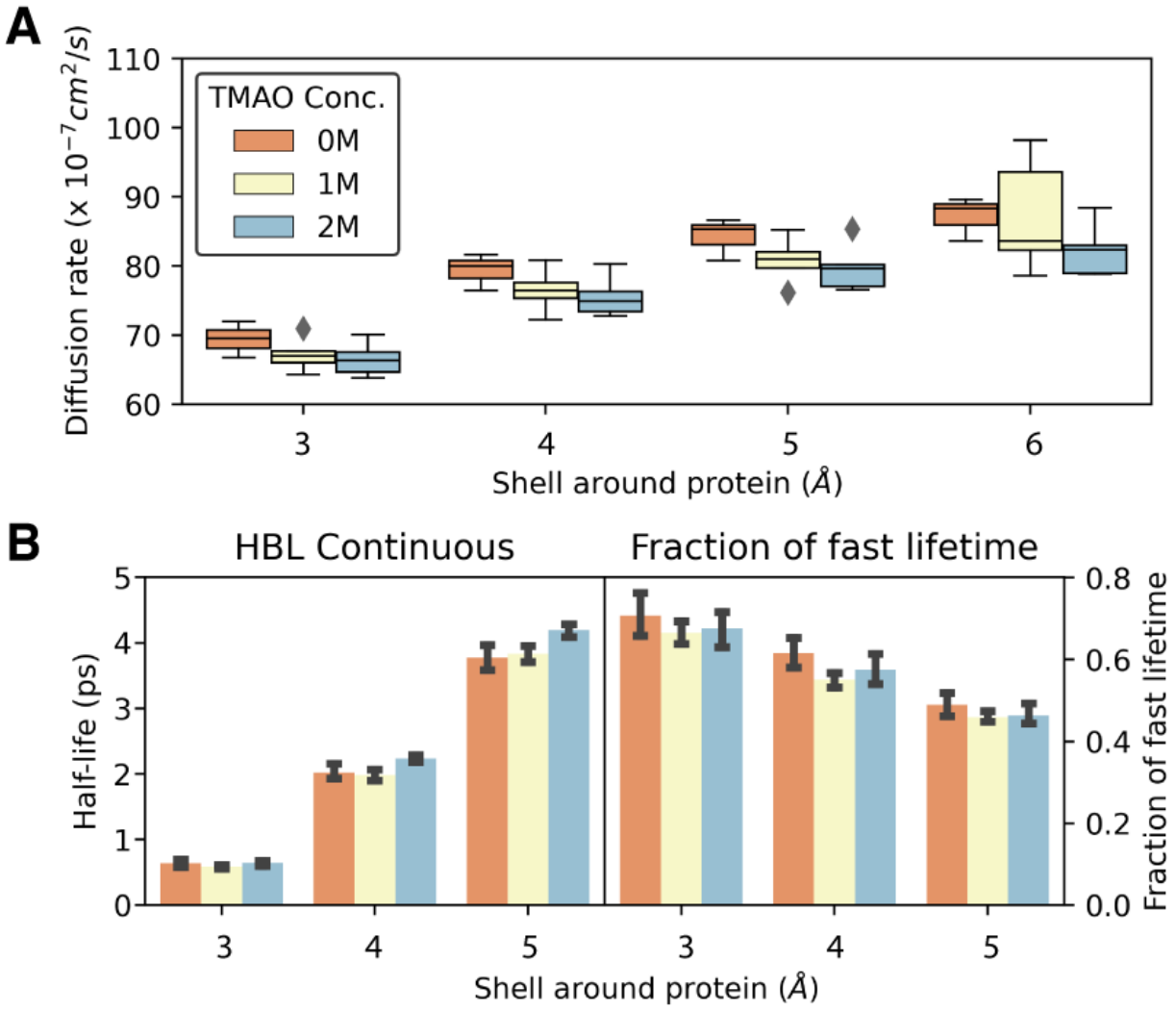

Fig. 4 Water diffusion and water-water hydrogen bonds. (A) The diffusion coefficient of water within shells from the protein surface at 3, 4, 5 and 6 Å for the 0 M (orange), 1 M (yellow), and 2 M (blue) TMAO concentrations

based on fifteen 100 ps (5 at each concentration of TMAO) simulations at 360K from a folded structure. The box in the boxplots shows the interquartile range, the whiskers show the minimum and maximum, the line within the box show the median, while the diamonds are the outliers. (B) The continuous water-water hydrogen bond lifetimes (HBL, left) and the fraction of the fast lifetime (A_1 as described in equation S5 in the SI) contributing to it (right) for the three shells around the protein for the fifteen 100 ps simulations at 360K - 5 each with 0 M, 1 M, and 2 M TMAO.

Water is herded into the vicinity of protein B in the presence of two TMAO shells. The amount of water in two hydration shells peaking at 2.7 and 3.5 Å from the protein surface increases in the presence of 1 M TMAO at 360 K (Fig. 3B). The first hydration shell is due mostly to direct hydrogen bonding between water molecules and protein side chains/backbone. The probability of water being present in the first hydration shell is increased in 1 M TMAO, despite the observation that TMAO binds to basic side chains with an oxygen radial distance also ≈ 2.7 Å.

The inner and outer shells of TMAO cooperate to increase water occupancy near the protein. TMAO, a kosmotrope, locally orders water in our simulations—both the oxide and the methyl moieties of TMAO are hydrated (Fig. 3A)—in agreement with experimental density data (55). Water occupancy (area under curve in absolute units is 0.9 at 0M vs 1.02 at 1M and 1.16 at 2M) is also enhanced significantly between the methyl groups of the inner and outer TMAO shells (3.5 Å peak in Fig. 3B). The amount of water near the protein continues to increase in 2 M TMAO (Fig. 3B), but this does not correlate with a further increase of protein stability, which seems to saturate at 1 M as discussed earlier.

Water diffusion slows down but hydrogen bond lifetimes remain nearly unchanged as TMAO is added. To understand how the presence of TMAO affects the diffusion and hydrogen bonding of water near the protein surface, we selected five protein-water structures at 4 μs intervals from simulations at 360 K for each TMAO concentration (0, 1, and 2 M) and performed fifteen short 100 ps MD simulations, with 2.5 fs temporal sampling of the trajectories, to study the residence times and diffusion rates of water molecules around the protein in the presence and absence of TMAO.

Near the protein surface, we find that the diffusion of water molecules is steadily reduced by about 10% progressing from 0 to 2 M TMAO, diffusion being dependent on distance from the protein surface (Fig. 4A). TMAO narrows the range of diffusion coefficients observed for the ensemble of water molecules at each of the protein-water distances in Fig. 4A, including near the outer TMAO shell (~6 Å), where the majority of TMAO molecules near the protein reside. Reduced water diffusion all the way to the outer TMAO shell is in keeping with the herding effect described above. A recent study similarly shows a decrease in the mobility (diffusion coefficient) of both bulk water, and water hydrogen bonded with TMAO, explained by an increase in effective viscosity of water in the presence of TMAO (56, 57).

In contrast, we observe that TMAO has a modest effect on water-water hydrogen bonding lifetimes possibly due to long lived TMAO-water hydrogen bonds (57), slightly increasing H-bond persistence (Fig. 4B and S8). The effect of the protein surface is much larger, with the H-bond half-life decreasing towards the faster component at a smaller distance from the protein surface. We used two definitions of lifetime to characterize hydrogen bonds: the ‘continuous’ lifetimes report the half-life of a hydrogen bond that remains continuously formed while ‘intermittent’ lifetimes allow for the bond to be broken briefly and formed again (see definition in SI). The continuous water-water hydrogen bond lifetimes can be fitted accurately by a double exponential accounting for shorter- and longer-lived H-bond populations, both of which slow down further from the protein surface (Fig. S8 and S9).

Discussion

In our simulations TMAO increases protein stability in the presence of thermal stress, consistent with its behavior in simulations at high pressure (6, 14) or in the presence of denaturants (14, 16, 21, 31). Mechanistically, most TMAO resides in an outer methyl-towards-protein shell at ~5.5 Å, but an oxygen-towards-protein inner shell of TMAO molecules interacts directly with basic side chains (~2.5 Å) at the protein surface, cooperating in a push-pull effect to herd water molecules towards the protein even at high temperature. The interaction of TMAO with basic residues prevalent at all temperatures in our simulations (Fig. 3) might also play a role in the stabilization against pressure. One way to further probe this would be to study frequency of basic surface side chains in deep sea organisms.

The stabilizing effect of 1 M vs. 0 M TMAO at high temperature is evident (Fig. 1A). Osmolytes generally induce a near-linear increase of the folding free energy as a function of osmolyte concentration (11). For example, the number of TMAO molecules increases roughly linearly with concentration in both TMAO shells (Table 1), crowding (58) changes roughly linearly up to 2 M TMAO, and the diffusion coefficient of water in Fig. 4A decreases roughly linearly with TMAO concentration.

Nonetheless, other indicators of protein-TMAO interaction saturate at 1 M TMAO, or even turn over between 1 and 2 M TMAO, especially at high temperature: For our simulation length at least, when integrating the probability distribution of native (Q>0.8) occupancy (Fig. 2), the protein occupies the folded region consistently less frequently in 2 M than in 1 M TMAO at high temperature (Fig. 5A and Fig. S14). If one looks at the area of in $g(r)$ of the first hydration shell (Fig. 5B), addition of TMAO increases the population of water molecules near the surface of the protein (herding), but it is evident that saturation (or even a plateau at the highest temperature) occurs between 1 and 2 M TMAO. The diffusion of water is damped by TMAO at high temperature as part of the herding, but not significantly more in 2 M than in 1 M TMAO (Fig. 5C). In this case, more shepherds do not herd better. This is also evident in Fig. 5D, which shows how the second shell of TMAO molecules behaves when the temperature is raised: 1 M TMAO concentration (light to dark green) at 5.5 Å from the protein surface is impervious to increasing temperature, whereas in 2 M TMAO (light to dark blue), there is a loss of osmolyte in the second shell to below the 1 M level.

The maximum amount of TMAO that has been found in marine organisms is about 0.9 M (4), and solubility is sometimes cited as the cause (4, 5). However, TMAO is soluble at concentrations above 4M in water at 20 °C, (20, 59) and remains soluble above 1.5 M even at 12 kbar (1, 2). We propose instead that the protein stabilization plateau (Fig. 5) is the cause: the cost of additional TMAO biosynthesis does not pay off with increased protein stability above 1 M. Indeed, a study previously reported that TMAO can have a destabilizing effect on proteins at high concentrations (60), so the stability plateau we observe above 1 M may be a sign of a threshold beyond which TMAO eventually destabilizes the protein, particularly at higher temperature. Why 1 M? 1 M TMAO corresponds to a mean distance of 4Å between TMAO molecules, and this about the distance over which TMAO perturbs its own hydration shell; when different TMAO hydration shells begin to overlap, adding more TMAO has a diminishing effect on water molecules. Analogous behavior has been seen in the THz spectrum of water as a function of protein concentration, where the effect of protein on water dynamics saturates at a protein-protein separation of 1 to 2 nm (61).

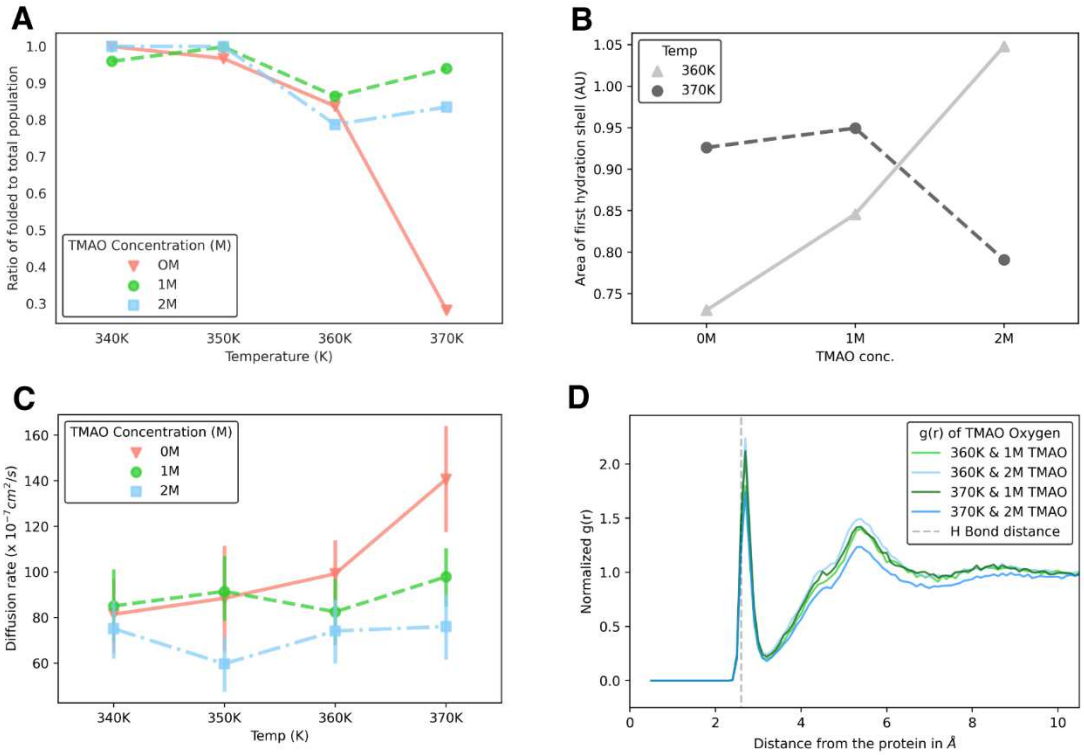

Fig. 5. Overview of the effect of temperature on stability of PRB and effect of TMAO. (A) Ratio of folded PRB population to total population gives an idea of the destabilization effect due to temperature and how TMAO (peach for 0M TMAO, light green for 1M TMAO and sky blue for 2M TMAO) can reduce this at higher temperatures. (B) The amount of water present in the first hydration shell plotted as area under the g(r) curve till 4 Å from the protein surface plotted in absolute units. depends on the TMAO concentration and temperature. Extreme temperatures lead to unfolding of the protein and increasing water at the surface. (C) At high temperatures the rate of diffusion is slower in the presence of TMAO. This plot is averaged over 4 shells (3A, 4A, 5A and 6A around the protein) for 100 ps short runs at each temperature starting out from a folded structure. (D) The normalized g(r) TMAO oxygen around the protein surface corresponding to two temperatures and two concentrations.

TMAO does have a dramatic effect on its own hydration shell: TMAO on average forms 3.2 hydrogen bonds with water (Fig. S10). Though earlier studies have supported TMAO accepting 3 hydrogen bonds from water (62–64), a recent study suggests that TMAO can form up to 4 hydrogen bonds with water at high pressure, and we see this effect on a smaller scale in our variable temperature simulations as well (6). The ability of TMAO to form up to four hydrogen bonds with water may be one of the reasons it is able to disrupt water structure and herd more water molecules towards the protein surface, and why the effect quickly saturates with TMAO concentration.

TMAO is thought to protect proteins from pressure by compacting the folded state because it pushes water towards the protein surface. In our simulations as a function of temperature, the protein void volume does not change by a large amount (Fig. S4) in the presence of TMAO: we do not observe compaction of the folded state due to TMAO with temperature as the variable. There is, however, a shift in helical arrangement giving rise to compact intermediates with high *Q* values (0.2-0.7) in the presence of TMAO, whereas a highly unfolded ensemble predominates in 0 M TMAO. This indicates a different mechanism of protein stabilization by TMAO with thermal denaturation as compared to

pressure denaturation. TMAO at high temperature acts by increasing the stability of partly structured states, not by mitigating destabilization of the native state due to reduced folded protein volume.

In summary, TMAO tends to point in opposite directions within its first and second shells above the protein surface, herding water molecules between the methyl groups. Much like a few shepherding dogs can keep a large number of sheep centered on a desired location, the cooperative co-solute effect of first- and second-shell TMAO has a small number of TMAO molecules herding water molecules into the two observed hydration shells at or near the protein surface. Protein stability, water hydration, and presence of TMAO near the protein saturate or turn over at 1 M TMAO especially at high temperature, pointing to a natural limit around 1 M TMAO above which organisms do not derive greater benefits from the energy-costly biosynthesis of TMAO.

Acknowledgments

The authors acknowledge supercomputing time provided by Extreme Science and Engineering Discovery Environment (XSEDE) grant TG-MCB 150089 awarded to S.S. and T.V.P. and School of Chemical Sciences Computer Center at University of Illinois Urbana-Champaign. M.B. and M.G. were supported by NSF grant MCB 2205665 and the NSF RCN on Extreme Biophysics. T.V.P. acknowledges support from the Department of Chemistry, University of Illinois at Urbana–Champaign.

References

1. Ball, P., and J.E. Hallsworth. 2015. Water structure and chaotropicity: their uses, abuses and biological implications. *Phys. Chem. Chem. Phys.* 17:8297–8305.

2. Sukenik, S., L. Sapir, R. Gilman-Politi, and D. Harries. 2013. Diversity in the mechanisms of cosolute action on biomolecular processes. *Faraday Discuss.* 160:225–237.

3. Yancey, P.H. 2001. Water Stress, Osmolytes and Proteins. *Integr Comp Biol*. 41:699–709.

4. Yancey, P.H., M.E. Gerringer, J.C. Drazen, A.A. Rowden, and A. Jamieson. 2014. Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. *PNAS*. 111:4461–4465.

5. Yancey, P.H. 2020. Cellular responses in marine animals to hydrostatic pressure. *J. Exp. Zool.* 333:398–420.

6. Kolling, I., C. Hölzl, S. Imoto, S.R. Alfarano, H. Vondracek, L. Knake, F. Sebastiani, F. Novelli, C. Hoberg, J.-B. Brubach, P. Roy, H. Forbert, G. Schwaab, D. Marx, and M. Havenith. 2021. Aqueous TMAO solution under high hydrostatic pressure. *Phys. Chem. Chem. Phys.* 23:11355–11365.

7. Munroe, K.L., D.H. Magers, and N.I. Hammer. 2011. Raman Spectroscopic Signatures of Noncovalent Interactions Between Trimethylamine N-oxide (TMAO) and Water. *J. Phys. Chem. B*. 115:7699–7707.

8. Larini, L., and J.E. Shea. 2013. Double resolution model for studying TMAO/water effective interactions. *Journal of Physical Chemistry B*. 117:13268–13277.

9. Zetterholm, S.G., G.A. Verville, L. Boutwell, C. Boland, J.C. Prather, J. Bethea, J. Cauley, K.E. Warren, S.A. Smith, D.H. Magers, and N.I. Hammer. 2018. Noncovalent Interactions between Trimethylamine *N* -Oxide (TMAO), Urea, and Water. *J. Phys. Chem. B*. 122:8805–8811.

10. Imoto, S., H. Forbert, and D. Marx. 2018. Aqueous TMAO solutions as seen by theoretical THz spectroscopy: hydrophilic *versus* hydrophobic water. *Phys. Chem. Chem. Phys.* 20:6146–6158.

11. Canchi, D.R., and A.E. García. 2013. Cosolvent Effects on Protein Stability. *Annual Review of Physical Chemistry*. 64:273–293.

12. Lin, T.-Y., and S.N. Timasheff. 1994. Why do some organisms use a urea-methylamine mixture as osmolyte? Thermodynamic compensation of urea and trimethylamine N-oxide interactions with protein. *Biochemistry*. 33:12695–12701.

13. Bolen, D.W., and G.D. Rose. 2008. Structure and Energetics of the Hydrogen-Bonded Backbone in Protein Folding. *Annu. Rev. Biochem.* 77:339–362.

14. Canchi, D.R., P. Jayasimha, D.C. Rau, G.I. Makhatadze, and A.E. Garcia. 2012. Molecular mechanism for the preferential exclusion of TMAO from protein surfaces. *Journal of Physical Chemistry B*. 116:12095–12104.

15. Wolde-Kidan, A., Q.D. Pham, A. Schlaich, P. Loche, E. Sparr, R.R. Netz, and E. Schneck. 2019. Influence of polar co-solutes and salt on the hydration of lipid membranes. *Phys. Chem. Chem. Phys.* 21:16989–17000.

16. Smolin, N., V. Voloshin, A. Anikeenko, A. Geiger, R. Winter, and N.N. Medvedev. 2017. TMAO and urea in the hydration shell of the protein SNase. *Phys. Chem. Chem. Phys.*

17. Yancey, P., M. Clark, S. Hand, R. Bowlus, and G. Somero. 1982. Living with water stress: evolution of osmolyte systems. *Science*. 217:1214–1222.

18. Wang, A., and D.W. Bolen. 1997. A Naturally Occurring Protective System in Urea-Rich Cells: Mechanism of Osmolyte Protection of Proteins against Urea Denaturation †. *Biochemistry*. 36:9101–9108.

19. Mukherjee, M., and J. Mondal. 2020. Unifying the Contrasting Mechanisms of Protein-Stabilizing Osmolytes. *J. Phys. Chem. B*. 124:6565–6574.

20. Liao, Y.-T., A.C. Manson, M.R. DeLyser, W.G. Noid, and P.S. Cremer. 2017. Trimethylamine *N* -oxide stabilizes proteins via a distinct mechanism compared with betaine and glycine. *Proc Natl Acad Sci USA*. 114:2479–2484.

21. Ganguly, P., J. Polak, N.F.A. van der Vegt, J. Heyda, and J.-E. Shea. 2020. Protein Stability in TMAO and Mixed Urea-TMAO Solutions. *J. Phys. Chem. B*. acs.jpcb.0c04357.

22. Boob, M., Y. Wang, and M. Gruebele. 2019. Proteins: “Boil ’Em, Mash ’Em, Stick ’Em in a Stew”. *J. Phys. Chem. B*. 123:8341–8350.

23. Privalov, P.L., and N.N. Khechinashvili. 1974. A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. *Journal of Molecular Biology*. 86:665–684.

24. Lins, L., and R. Brasseur. 1995. The hydrophobic effect in protein folding. *FASEB j.* 9:535–540.

25. Liu, F., M. Nakaema, and M. Gruebele. 2009. The transition state transit time of WW domain folding is controlled by energy landscape roughness. *J. Chem. Phys.* 131:195101.

26. Wimbush, M. 1970. Temperature Gradient above the Deep-Sea Floor. *Nature*. 227:1041–1043.

27. Treberg, J.R., J.S. Bystriansky, and W.R. Driedzic. 2005. Temperature effects on trimethylamine oxide accumulation and the relationship between plasma concentration and tissue levels in smelt (Osmerus mordax). *Journal of Experimental Zoology Part A: Comparative Experimental Biology*. 303A:283–293.

28. Johansson, M.U., M. de Château, M. Wikström, S. Forsén, T. Drakenberg, and L. Björck. 1997. Solution structure of the albumin-binding GA module: a versatile bacterial protein domain. *Journal of Molecular Biology*. 266:859–865.

29. Lindorff-Larsen, K., S. Piana, R.O. Dror, and D.E. Shaw. 2011. How Fast-Folding Proteins Fold. *Science*. 334:517–520.

30. Voloshin, V., N. Smolin, A. Geiger, R. Winter, and N.N. Medvedev. 2019. Dynamics of TMAO and urea in the hydration shell of the protein SNase. *Phys. Chem. Chem. Phys.* 21:19469–19479.

31. Mondal, J., G. Stirnemann, and B.J. Berne. 2013. When does TMAO fold a polymer chain and urea unfold it? .

32. Guin, D., S. Mittal, B. Bozymski, D. Shukla, and M. Gruebele. 2019. Dodine as a Kosmo-Chaotropic Agent. *J. Phys. Chem. Lett.* 10:2600–2605.

33. Abraham, M.J., T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, and E. Lindahl. 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 1–2:19–25.

34. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*. 14:33–38.

35. Essmann, U., L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.

36. Nosé, S. 1984. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*. 52:255–268.

37. Hoover, W.G. 1985. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*. 31:1695–1697.

38. Huang, J., and A.D. MacKerell. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem*. 34:2135–2145.

39. Mark, P., and L. Nilsson. 2001. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A*. 105:9954–9960.

40. Kast, K.M., J. Brickmann, S.M. Kast, and R.S. Berry. 2003. Binary Phases of Aliphatic N-Oxides and Water: Force Field Development and Molecular Dynamics Simulation. *J. Phys. Chem. A*. 107:5342–5351.

41. Schneck, E., D. Horinek, and R.R. Netz. 2013. Insight into the Molecular Mechanisms of Protein Stabilizing Osmolytes from Global Force-Field Variations. *J. Phys. Chem. B*. 117:8310–8321.

42. Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A.D. MacKerell. 2010. CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem*. 31:671–690.

43. 2020. Anaconda Software Distribution. Anaconda, Inc.

44. McGibbon, R.T., K.A. Beauchamp, M.P. Harrigan, C. Klein, J.M. Swails, C.X. Hernández, C.R. Schwantes, L.-P. Wang, T.J. Lane, and V.S. Pande. 2015. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical Journal*. 109:1528–1532.

45. Chen, C.R., and G.I. Makhatadze. 2015. ProteinVolume: calculating molecular van der Waals and void volumes in proteins. *BMC Bioinformatics*. 16:101.

46. Best, R.B., G. Hummer, and W.A. Eaton. 2013. Native contacts determine protein folding mechanisms in atomistic simulations. *Proceedings of the National Academy of Sciences*. 110:17874–17879.

47. Sukenik, S., T.V. Pogorelov, and M. Gruebele. 2016. Can Local Probes Go Global? A Joint Experiment–Simulation Analysis of λ 6–85 Folding. *The Journal of Physical Chemistry Letters*. 7:1960–1965.

48. Michaud-Agrawal, N., E.J. Denning, T.B. Woolf, and O. Beckstein. 2011. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* 32:2319–2327.

49. Gowers, R., M. Linke, J. Barnoud, T. Reddy, M. Melo, S. Seyler, J. Domański, D. Dotson, S. Buchoux, I. Kenney, and O. Beckstein. 2016. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. Austin, Texas: . pp. 98–105.

50. Prigozhin, M.B., Y. Zhang, K. Schulten, M. Gruebele, and T.V. Pogorelov. 2019. Fast pressure-jump all-atom simulations and experiments reveal site-specific protein dehydration-folding dynamics. *PNAS*. 116:5356–5361.

51. Auton, M., and D.W. Bolen. 2005. Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proceedings of the National Academy of Sciences*. 102:15065–15068.

52. Sagle, L.B., K. Cimatu, V.A. Litosh, Y. Liu, S.C. Flores, X. Chen, B. Yu, and P.S. Cremer. 2011. Methyl Groups of Trimethylamine N-Oxide Orient Away from Hydrophobic Interfaces. *J. Am. Chem. Soc.* 133:18707–18712.

53. Stanley, C., and D.C. Rau. 2008. Assessing the Interaction of Urea and Protein-Stabilizing Osmolytes with the Nonpolar Surface of Hydroxypropylcellulose †. *Biochemistry*. 47:6711–6718.

54. Sagle, L.B., K. Cimatu, V.A. Litosh, Y. Liu, S.C. Flores, X. Chen, B. Yu, and P.S. Cremer. 2011. Methyl Groups of Trimethylamine *N* -Oxide Orient Away from Hydrophobic Interfaces. *J. Am. Chem. Soc.* 133:18707–18712.

55. Rösgen, J., and R. Jackson-Atogi. 2012. Volume Exclusion and H-Bonding Dominate the Thermodynamics and Solvation of Trimethylamine- *N* -oxide in Aqueous Urea. *J. Am. Chem. Soc.* 134:3590–3597.

56. Teng, X., Q. Huang, C.C. Dharmawardhana, and T. Ichiye. 2018. Diffusion of aqueous solutions of ionic, zwitterionic, and polar solutes. *The Journal of Chemical Physics*. 148:222827.

57. Teng, X., and T. Ichiye. 2019. Dynamical Effects of Trimethylamine N-Oxide on Aqueous Solutions of Urea. *J Phys Chem B*. 123:1108–1115.

58. Ma, J., I.M. Pazos, and F. Gai. 2014. Microscopic insights into the protein-stabilizing effect of trimethylamine N-oxide (TMAO). *Proceedings of the National Academy of Sciences*. 111:8476–8481.

59. Folberth, A., J. Polák, J. Heyda, and N.F.A. van der Vegt. 2020. Pressure, Peptides, and a Piezolyte: Structural Analysis of the Effects of Pressure and Trimethylamine- *N* -oxide on the Peptide Solvation Shell. *J. Phys. Chem. B*. 124:6508–6519.

60. Rani, A., A. Jayaraj, B. Jayaram, and V. Pannuru. 2016. Trimethylamine- N -oxide switches from stabilizing nature: A mechanistic outlook through experimental techniques and molecular dynamics simulation. *Sci Rep*. 6:1–14.

61. Born, B., S.J. Kim, S. Ebbinghaus, M. Gruebele, and M. Havenith. 2008. The terahertz dance of water with the proteins: the effect of protein flexibility on the dynamical hydration shell of ubiquitin. *Faraday Discuss.* 141:161–173.

62. Hunger, J., K.-J. Tielrooij, R. Buchner, M. Bonn, and H.J. Bakker. 2012. Complex Formation in Aqueous Trimethylamine-N-oxide (TMAO) Solutions. *J. Phys. Chem. B*. 116:4783–4795.

63. Hunger, J., N. Ottosson, K. Mazur, M. Bonn, and H.J. Bakker. 2014. Water-mediated interactions between trimethylamine-N-oxide and urea. *Phys. Chem. Chem. Phys.* 17:298–306.

64. Xie, W.J., S. Cha, T. Ohto, W. Mizukami, Y. Mao, M. Wagner, M. Bonn, J. Hunger, and Y. Nagata. 2018. Large Hydrogen-Bond Mismatch between TMAO and Urea Promotes Their Hydrophobic Association. *Chem*. 4:2615–2627.