Entropic Stabilization of Proteins by TMAO

Samuel S. Cho,†‡ Govardhan Reddy,† John E. Straub,§ and D. Thirumalai*†

†Department of Chemistry and Biochemistry and Biophysics Program, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, United States
§Department of Chemistry, Boston University, Boston, Massachusetts 02215, United States

ABSTRACT: The osmolyte trimethylamine N-oxide (TMAO) accumulates in the cell in response to osmotic stress and increases the thermodynamic stability of folded proteins. To understand the mechanism of TMAO induced stabilization of folded protein states, we systematically investigated the action of TMAO on several model dipeptides (leucine, L2, serine, S2, glutamine, Q2, lysine, K2, and glycine, G2) in order to elucidate the effect of residue-specific TMAO interactions on small fragments of solvent-exposed conformations of the denatured states of proteins. We find that TMAO preferentially hydrogen bonds with the exposed dipeptide backbone but generally not with nonpolar or polar side chains. However, interactions with the positively charged Lys are substantially greater than with the backbone. The dipeptide G2 is a useful model of the pure amide backbone; interacts with TMAO by forming a hydrogen bond between the amide nitrogen and the oxygen in TMAO. In contrast, TMAO is depleted from the protein backbone in the hexapeptide G6, which shows that the length of the polypeptide chain is relevant in aqueous TMAO solutions. These simulations lead to the hypothesis that TMAO-induced stabilization of proteins and peptides is a consequence of depletion of the solute from the protein surface provided intramolecular interactions are more favorable than those between TMAO and the backbone. To test our hypothesis, we performed additional simulations of the action of TMAO on an intrinsically disordered Aβ16–22 (KLVFFAE) monomer. In the absence of TMAO, Aβ16–22 is a disordered random coil. However, in aqueous TMAO solution, Aβ16–22 monomer samples compact conformations. A transition from random coil to α-helical secondary structure is observed at high TMAO concentrations. The coil to α-helix transition is highly cooperative especially considering the small number of residues in Aβ16–22. Our work highlights the potential similarities between the action of TMAO on long polypeptide chains and entropic stabilization of proteins in a crowded environment due to excluded volume interactions. In this sense, the chemical chaperone TMAO is a nanocrowding agent.

INTRODUCTION

Trimethylamine N-oxide (TMAO) is a naturally occurring osmolyte that accumulates in organisms to counteract the destabilizing effect of urea1 on folded protein conformations. A number of experiments have shown that TMAO stabilizes proteins,2–4 but the precise molecular mechanism has not been firmly established.5–7 The stabilization of proteins by TMAO can be qualitatively understood from the perspective of an entropic stabilization mechanism introduced in the context of crowding effects on protein stability.8 Depletion of an osmolyte from the vicinity of proteins results in compact conformations, which stabilizes the native states.8–12 On the other hand, if an osmolyte were to interact directly with the protein, as is the case with denaturants such as urea and guanidinium hydrochloride, the native basin of the protein would be destabilized.13–15 These arguments, while rationalizing the different roles of protective and denaturing osmolytes, do not provide a molecular explanation of their actions.

The structure of TMAO (Figure 1) suggests that there are two main types of intermolecular interactions that are possible between TMAO and proteins. The oxygen atom on TMAO (O7) can act as a hydrogen bond acceptor. Three methyl groups in TMAO can participate in hydrophobic interactions with the side chains of proteins. From transfer free energy calculations, it has been deduced that TMAO has no significant preference for hydrophobic moieties, but TMAO interacts with the backbone, as well as charged and polar side chains.2 It is necessary to extend such studies to systems in which chain connectivity and sequence effects are explicitly taken into account.

To dissect the molecular basis for the action of TMAO on peptides, we simulated five dipeptides in explicit water in 1 M TMAO. The dipeptides are ideal model systems for the study of TMAO–protein interactions because, like unfolded proteins, they are solvent exposed, and hence can freely interact with the surrounding solvent molecules. Each dipeptide was composed of one of the following types of amino acids: leucine (nonpolar), glutamine, serine, lysine, and glycine.
serine (polar, hydroxyl group), glutamine (polar, amino group), and lysine (basic). In addition, we also studied conformation changes in diglycine (G$_2$) and hexaglycine (G$_6$), which has been recently investigated using MD simulations in aqueous TMAO solution.$^{16}$ Comparison of the conformational changes in G$_2$ and G$_6$ in TMAO leads to the hypothesis that as the polypeptide chain length increases TMAO is expelled from the surface, which results in the collapse of the predominantly backbone construct. Consequently, G$_6$ adopts a conformation that maximizes the intrapeptide force for proteins and waters. In order to describe the interactions and the CHARMM22$^{18}$ force field parameters of Kast et al.$^{20}$ We first simulated five dipeptides and one hexa-glycine in order to dissect the interpeptide interactions. In order to further validate our hypothesis, we performed all-atom MD simulations of A$\beta$$_{16-22}$ (KLVFFAE) monomer, which aggregates to form amyloid fibrils, in various TMAO concentrations. The A$\beta$$_{16-22}$ peptide, which consists of a short sequence of hydrophobic residues flanked by two oppositely charged residues is disordered and adopts a random coil conformation that is devoid of secondary structure. Remarkably, A$\beta$$_{16-22}$ becomes helical upon interaction with increasing concentrations of TMAO. Analysis of the conformations of A$\beta$$_{16-22}$ shows that TMAO is depleted from the surface of the backbone, which establishes that TMAO-induced transition from random coil to $\alpha$-helix is due to the entropic stabilization mechanism. Thus, the stabilization of proteins by TMAO is akin to a mechanism by which crowding particles stabilize proteins, which suggests TMAO can be treated as a nanocrowding particle.

### METHODS

We performed MD simulations using the NAMD program$^{17}$ and the CHARMM22$^{18}$ force field with the CMAP modification$^{19}$ for proteins and waters. In order to describe the interactions between the osmolyte and the polypeptide chains, we used the TMAO force field parameters of Kast et al.$^{20}$ We first simulated five dipeptides and one hexa-glycine in order to dissect the interaction between TMAO and polypeptide chains. Each dipeptide was composed of one of the following types of amino acids: leucine (nonpolar), serine (polar, hydroxyl group), glutamine (polar, amino group), and lysine (basic). The diglycine and hexaglycine molecules were simulated in the absence and presence of TMAO.

As a starting point, the fully extended peptide was centered in a rectangular water box comprised of TIP3P water molecules, and all of the water molecules within 2.2 Å of the peptide molecule were deleted. The dimension of the water box was set to 10 Å, which is more than the length of the peptide. We performed 10 independent simulations where the TMAO positions were set by randomly replacing the TIP3P water molecules such that the concentration was 1.0 M, a value that is typically used in transfer experiments. For each independent trajectory, 100 initial configurations with different placements of the TMAO were generated, followed by 10 steps of steepest-descent and 25 steps of adopted-basis Newton–Raphson minimization with harmonic constraints on the peptide, and only the lowest energy configuration was used for simulations. Therefore, each independent simulation started from a unique, energy-minimized, random distribution of TMAO. We equilibrated the system by removing all harmonic constraints, applying 2000 steps of conjugate gradient minimization, and performing 50 ps of NVT MD simulations using a 2 fs time step. We then performed 10 ns production runs for NPT simulations of each di- and hexa-peptide at 298 K using the CHARMM force field. All analyses were performed for each production run.

For the A$\beta$$_{16-22}$ (KLVFFAE) monomer peptide simulations, the same protocol was used as with the di- and hexa-peptides except that the dimensions of the water box were cubic with each side of length 40 Å. The equilibration time period was increased to 5 ns, and the production run time length was increased to 100 ns at 300 K, of which the last 60 ns was used for analyses. Four sets of A$\beta$$_{16-22}$ monomer simulations were performed at TMAO concentrations of 0, 1.0, 2.5, and 5.0 M.

### RESULTS AND DISCUSSION

TMAO Preferentially Interacts with the Peptide Backbone and Basic Dipeptide Side Chains. For the simulations of the leucine dipeptide (Figure 2a), the radial distribution function, $g(r)$, between O$_T$ and the peptide backbone nitrogen (N), which is the only hydrogen bond donor, resulted in a peak with $g(r)\approx 3$ Å$^2$. At the typical hydrogen bond distance, $r \approx 3$ Å, the most probable value of the angle between N, the amide hydrogen, and O$_T$ is $150^\circ$ (Figure 3a). The $g(r)$ between the TMAO methyl carbon (C$_T$) and the carbon atom of the terminal methyl group in the leucine side chain has a modest peak where $g(r) \sim 1.0$ at $r \approx 4$ Å, which shows that the dispersion interactions with the side chain are negligible (Figure 2a). Since leucine is the most preferable interaction...
The preferential interaction of TMAO with the backbone and other hydrophobic side chain must be less favorable in comparison to hydrogen bonding with the backbone. Serine and glutamine dipeptides (Figure 2b,c), which have hydroxyl and amino hydrogen bond donors, respectively, in the side chains, also resulted in $g(r) \approx 1.5$ for the backbone nitrogen but $g(r) \approx 1.0$ for the side chain oxygen ($O_\epsilon$ in serine) and nitrogen atoms ($N_\gamma$ in glutamine). Interestingly, there is a greater preference for TMAO to hydrogen bond with the backbone nitrogen than the side chain hydrogen bond donors, including the amino nitrogen of glutamine (Figure 3bc).

The $g(r\approx3\text{Å}) \sim 1.5$ peak is independent of the polarity of the side chains (Figure 2). In lysine dipeptide, however, there is a pronounced peak ($g(r\approx3\text{Å}) \sim 2.5$) between $O_\epsilon$ and the charged side chain amino nitrogen ($N_\gamma$) without compromising the hydrogen bond formation with amide nitrogen (Figures 2d and 3d). Clearly, it is possible for charged side chain hydrogen bond donors to have significant interactions with TMAO. We expect similar results for TMAO interactions with other basic amino acids, but their relative abundance in proteins suggests that the overall significance of TMAO—side chain interactions may not be significant. Interestingly, the radial distribution functions in Figure 2 also show that the size of the side chain does not affect the extent of interactions with the backbone nitrogen atom, as long as the backbone is solvent-exposed.

The preferential interaction of TMAO with the backbone hydrogen bond donor over the uncharged side chain hydrogen bond donor can be understood in structural terms. The peptide backbone forms a resonance interaction between the nitrogen atom and the carbonyl group. Thus, the peptide bond not only has a partial double bond character, it also leaves the nitrogen with a partial positive charge (and the carbonyl oxygen with a partial negative charge). Therefore, TMAO would form more favorable hydrogen bond interactions with the partially charged backbone nitrogen than an uncharged side chain hydrogen bond acceptor and an even greater favorable interaction with fully charged side chain amino nitrogen, as explicitly shown for lysine. Of course, asparagine and glutamine side chain amides can also participate in resonance stabilization such that the amide nitrogen has a partial charge but the electronegativity of its neighbors makes its overall partial charge less than that of the peptide amide nitrogen.

**Figure 3.** Distribution of angles formed by hydrogen bonds between the peptide backbone amide $N$ and $H$ with the TMAO oxygen ($O_\gamma$) for the dipeptide constructs of (a) leucine, (b) serine, (c) glutamine, (d) lysine, and (e) glycine, as well as the hexaglycine construct. The angles exceeding 150°, which are characteristic of a perfect hydrogen bond, are most probable. Only interactions for which the distance between the backbone amide and TMAO is less than 3.5 Å, which corresponds to the first solvation shell, are considered. Thus, at the distance when $g(r)$ has a first peak in all dipeptides, $O_\gamma$ forms a hydrogen bond with the amide proton. See Supporting Information Figure S2 for the corresponding hydrogen bond distribution of angles for the dipeptide and hexapeptide constructs of glycine.

**Figure 4.** Conformations adopted by hexaglycine in the absence (a, c) and presence of 1 M TMAO (b, d). The Ramachandran free energy profiles are shown with the four major basins labeled (a, b). The normalized histograms of the radius of gyration, $R_g$, are shown with the $R_g$ of ideal $\alpha$-helices and PPII $\beta$-sheets labeled.
that of an ideal $\alpha$-helical $G_6$ (Figure 4c,d). Interestingly, $\langle R_g \rangle$ is similar to the value obtained by Shortle and co-workers from their NMR and SAXS measurements, which show a value intermediate between ideal $\alpha$-helices and PPII $\beta$-sheets.\textsuperscript{23}

The radial distribution functions involving the atomic interactions of water are almost quantitatively identical in pure water and 1 M TMAO (see Figure 5), even in the presence of hexaglycine. These results show that the structural changes observed in $G_6$ have to be related to the depletion of TMAO with the polypeptide chain. Since $\alpha$-helical $G_6$ conformations can have up to two intramolecular hydrogen bonds of the backbone nitrogen (with the backbone carbonyl oxygen that is separated by four amino acids earlier in sequence), we calculated the $g(r)$ between the N and OT. $\langle R_g \rangle$ for $G_6$ is much lower than that for diglycine ($G_2$) (Figure 6a,b), which we use as a control because it is too short to form intramolecular hydrogen bonds. The enhancement of the $\alpha$-helical basin in 1 M TMAO in $G_6$ is due to its depletion from around the peptide backbone, which is in accord with previous studies.\textsuperscript{2} Thus, the formation of $\alpha$-helical structure disfavors backbone hydrogen bond formation with TMAO, but the backbone amide nitrogen still favors hydrogen bonding with the carbonyl oxygen because of its larger dipole moment and proximity as compared to the TMAO oxygen. The differences in $g(r)$ between N and O$_T$ in $G_2$ and $G_6$ highlight the role of chain length in TMAO–peptide interaction. For $G_2$, the amide nitrogen is accessible to O$_T$, which is consistent with simulations of cyclic $G_2$.$^5$

**TMAO is a Nano-Crowding Agent.** Observations of ordered secondary structure in 1 M TMAO in $G_6$ can be rationalized using the depletion theory used to predict stability changes in the folded states in a crowded environment. Depletion of TMAO around $G_6$ essentially induces an osmotic pressure\textsuperscript{24} on $G_6$ which results in chain compaction, as was previously observed in MD simulations of the longer $G_{15}$ in TMAO.\textsuperscript{25} Thus, the polypeptide is forced to adopt conformations that maximize intramolecular interactions. In the $G_6$ case, this results in an increase in the population of $\alpha$-helical structure (Figure 4b). The exclusion of TMAO from $G_6$ is vividly illustrated using a number of pair functions involving water, TMAO, and the amide nitrogen (Figure 7a,b). Thus, in 1 M TMAO, $G_6$ is localized in a region that is devoid of both water and TMAO. Such a mechanism is exactly the one invoked to quantitatively predict the native state stabilization in the excluded volume dominated crowding agents!

Two remarks of caution are in order. (1) For small peptides such as $G_2$, the amide nitrogen interacts favorably with O$_w$ and O$_T$ (Figure 7c). Thus, polypeptide chain length is important in observing TMAO-induced structure formation. (2) More importantly, it is known from crowding theory\textsuperscript{11,12} that the nature of the structures adopted depends on $q = (R_g)/R_c$ where $R_c$ is the size of the crowding particles. In our study of $G_6$ in TMAO,
The Journal of Physical Chemistry B

ARTICLE

Figure 9. Ramachandran free energy profiles of $\text{A}\beta_{16-22}$ at TMAO concentrations of (a) 0 M, (b) 1 M, (c) 2.5 M, and (d) 5 M. The cooperative transition from random coil to $\alpha$-helix transition is shown in part e. This panel also shows that at all TMAO concentrations there is negligible $\beta$-strand content.

$\langle R_g \rangle = 5.01 \text{ Å} \text{ and } R_c = 1.32 \text{ Å}$, resulting in $q \approx 3.77$. It is unclear whether depletion theory also holds if $q$ and sequence are varied.

$\text{A}\beta_{16-22}$ Peptide Becomes Compact and $\alpha$-Helical with Increasing TMAO Concentration. In order to assess if the depletion mechanism leading to a shift in the population toward $\alpha$-helical structure is general, we used simulations to probe TMAO-induced changes in $\text{A}\beta_{16-22}$ (KLVFFAE) monomer peptide, which aggregates to form antiparallel fibrils $^{21}$ at high peptide concentration. We had shown earlier that $\text{A}\beta_{16-22}$ monomer is a random coil largely devoid of secondary structure. In particular, the population of $\alpha$-helical structure is less than about 1%. If TMAO acts as a nanocrowder, then we expect that $\text{A}\beta_{16-22}$ would be localized in a region devoid of TMAO. Under these conditions, $\text{A}\beta_{16-22}$ is expected to adopt an $\alpha$-helical conformation to maximize intramolecular interactions. $^{26}$ In order to test the applicability of depletion-induced structure formation, we performed simulations of $\text{A}\beta_{16-22}$ in various TMAO concentrations.

In addition to being a biologically relevant system, the intrinsically disordered $\text{A}\beta_{16-22}$ peptide is a very good model system to study the role of TMAO on conformational fluctuations of peptides, since its sequence consists of charged residues, a positive lysine (K), and a negative glutamic acid (E), that cap the ends of a short stretch of hydrophobic residues. The probability distribution of the radius of gyration, $P(R_g)$, of the $\text{A}\beta_{16-22}$ peptide shows that it becomes more compact with increasing TMAO concentration (Figure 8). The values of mean $R_g$ ($\langle R_g \rangle$) for the TMAO concentrations, [TMAO] = 0, 1.0, 2.5, and 5 M, are 6.9, 6.7, 5.9, and 5.7 Å, respectively. Thus, there is a 17% reduction in $\langle R_g \rangle$ as [TMAO] is changed from 0 to 5 M.
TMAO interacts with the Aβ16–22 Backbone and Lysine Side Chain. To determine the molecular interactions that induce the helical formation of the Aβ16–22 (KLVFFAE) peptide, we calculated the radial distribution function, g(r), between TMAO and the Aβ16–22 peptide. There is a stronger preference for terminal (K16, L17, A21, and E22) backbone amide N with TMAO (Figure 10a) compared with those in the interior (V18, F19, and F20) (Figure 10b), which may be a reflection of the bulky phenylalanine that effectively excludes interactions with the peptide backbone. The TMAO interactions with the backbone of these residues are more pronounced for polar interactions with the amide N (Figure 10c,d). The residence time of TMAO near the backbone atoms is approximately 55 ps, which is approximately twice that of water. The residence time is defined as the time during which any of the TMAO or water atoms are within 4 Å of any of the backbone atoms. Hydrophobic interactions with TMAO are modestly significant for the side chains (Figure 10e) and nonexistent in our simulations for the backbone (Figure 10f). The interactions between TMAO and the terminal positively charged lysine side chain, however, are pronounced (Figure 10e), even more than interactions with the terminal backbone amide N (Figure 10a). The affinity of TMAO to negatively charged side chains is minimal (Figure 10f).

**CONCLUSIONS**

Using all-atom MD simulations of a number of model peptide constructs and Aβ16–22 monomer in aqueous TMAO solution, we dissected the molecular mechanism of how TMAO stabilizes the native basin of proteins. By preferentially hydrogen bonding to the backbone nitrogen of the solvent exposed peptides, TMAO acts as a nano “crowder” that limits the degrees of freedom of the unfolded state and entropically destabilizes it. When the backbone nitrogen forms a secondary structure, it no longer available to hydrogen bond with TMAO, resulting in the depletion from the vicinity of the protein, which in turn results in native state stabilization. Comparisons between G2 and G6 show that polypeptide length is a relevant factor in determining the energetic balance between collapsed and extended structures. If the polypeptide chain exceeds a critical size, it is likely that aqueous TMAO would be a “poor” solvent for generic proteins, which would promote collapse and structure formation, as demonstrated for Aβ16–22 peptide. In particular, TMAO induces a highly cooperative coil to α-helix transition, a prediction that can be easily tested. Finally, our work also shows that in sequences that contain charged residues (e.g., intrinsically disordered proteins or fragments of Aβ peptides, such as Aβ16–22), the interactions between TMAO and positively charged side chains are significant. However, for generic proteins, TMAO is expelled from the surface. In this sense, TMAO behaves as a nanocrowding particle, thus stabilizing proteins by the entropic stabilization mechanism.10

**ASSOCIATED CONTENT**

3 Supporting Information. Radial distribution functions between amide N and H in dipptide constructs and oxygens of water and TMAO. Pair functions between amide N and H in

---

**Figure 10.** Radial distribution functions between atomic centers of TMAO and Aβ16–22. (a) The TMAO oxygen (O_T) interactions with the backbone amide N of the termini residues of the Aβ16–22 peptide. (b) The O_T interactions with backbone amide N of residues in the interior of the Aβ16–22 peptide. The only significant TMAO interactions with the backbone amide N are observed for those in the N-terminal residues. (c) A comparison of the total backbone polar amide N with O_T vs hydrophobic C_C interactions with the TMAO carbon C_T, as well as (d) the individual per residue hydrophobic C_C interactions. There are no significant hydrophobic backbone C_C interactions with TMAO. (e) Hydrophobic side chain interactions with C_T show that these interactions can be modestly significant. (f) The polar side chain interactions with O_T show a clear and significant preference of TMAO for the positively charged lysine side chain.

We further probed the structural changes by calculating the Ramachandran free energy profiles of the peptide. In the absence of TMAO, Aβ16–22 fluctuates among a number of distinct structures. Figure 9a,b shows that for Aβ16–22 the basins with (ϕ,ψ) angles that correspond to β-sheets and α-helices are populated, as would be expected from an intrinsically disordered peptide that is basically a random coil. In the absence of TMAO, Aβ16–22 has negligible ϕ-helical or β-strand content (Figure 9e), which accords well with our earlier study.26 At a modest concentration of TMAO (i.e., [TMAO] = 2.5 M), the β-strand basins disappear (Figure 9c). Remarkably, for [TMAO] greater than 2.5 M, only the right-handed α-helices remain (Figure 9c,d). Thus, TMAO induces a transition between a predominantly random coil state to α-helical structure. Considering the small size of Aβ16–22, the transition is relatively sharp, as assessed by the cooperativity measure.27 Similar results are experimentally observed where TMAO induces helical formation in alanine peptides.28

The tendency of Aβ16–22 to form ordered α-helical structures has implications for oligomer formation. The formation of stable helical structure could preclude amyloid formation, which requires β-structures as seeds. It is interesting to constrain TMAO-induced structure formation to the effects of urea on Aβ16–22. Molecular dynamics simulations showed that in urea Aβ16–22 monomer is extended and forms β-strands.29

---

Supporting Information. Radial distribution functions between amide N and H in dipptide constructs and oxygens of water and TMAO. Pair functions between amide N and H in


REFERENCES


