Biophysical Journal
Article

Biophysical Society

Using a FRET Library with Multiple Probe Pairs To Drive Monte Carlo Simulations of α-Synuclein

John J. Ferrie,[1] Conor M. Haney,[1] Jimin Yoon,[1] Buyan Pan,[1] Yi-Chih Lin,[1] Zahra Fakhraai,[1] Elizabeth Rhoades,[1] Abhinav Nath,[2] and E. James Petersson[1,*]

[1]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania and [2]Department of Medicinal Chemistry, School of Pharmacy, University of Washington, Seattle, Washington

ABSTRACT We describe a strategy for experimentally-constraining computational simulations of intrinsically disordered proteins (IDPs), using α-synuclein, an IDP with a central role in Parkinson's disease pathology, as an example. Previously, data from single-molecule Förster Resonance Energy Transfer (FRET) experiments have been effectively utilized to generate experimentally constrained computational models of IDPs. However, the fluorophores required for single-molecule FRET experiments are not amenable to the study of short-range (<30 Å) interactions. Using ensemble FRET measurements allows one to acquire data from probes with multiple distance ranges, which can be used to constrain Monte Carlo simulations in PyRosetta. To appropriately employ ensemble FRET data as constraints, we optimized the shape and weight of constraining potentials to afford ensembles of structures that are consistent with experimental data. We also used this approach to examine the structure of α-synuclein in the presence of the compacting osmolyte trimethylamine-*N*-oxide. Despite significant compaction imparted by 2 M trimethylamine-*N*-oxide, the underlying ensemble of α-synuclein remains largely disordered and capable of aggregation, also in agreement with experimental data. These proof-of-concept experiments demonstrate that our modeling protocol enables one to efficiently generate experimentally constrained models of IDPs that incorporate atomic-scale detail, allowing one to study an IDP under a variety of conditions.

INTRODUCTION

Intrinsically disordered proteins (IDPs) and proteins containing disordered regions are exceptionally responsive to changes in solution conditions, making them prone to misfolding and aggregation. One such IDP is α-synuclein (αS), a 140-amino-acid neuronal protein, the aggregation of which is implicated in Parkinson's disease pathogenesis (1,2). αS is primarily expressed at presynaptic termini and is suspected to play roles in regulating neurotransmitter release and maintaining synaptic function and plasticity (3). Likely the most recognized aspect of αS is its ability to misfold and self-associate, resulting in the production of toxic amyloid fibrils. These fibrils are the primary components of Lewy Bodies, which have been long been recognized as a postmortem hallmark of Parkinson's disease (1). The protein is comprised of three domains: the N-terminal domain (residues 1–60), comprised of four imperfect 11-amino-acid repeats featuring a KTKEGV motif found in amphipathic helices; the nonamyloid β-component, or NAC domain (residues 61–95), which contains two additional KTKEGV repeats and forms the β-sheet rich core of amyloid aggregates; and the acidic C-terminal domain, which is highly charged and is considered to be largely disordered. The first two domains adopt an α-helical structure in the presence of lipid membranes or detergent micelles whereas, in solution, the αS monomer is largely disordered (2,4–8). The structural plasticity of αS and most IDPs is attributed to a lack of hydrophobic residues and an excess of charged residues, producing systems that are self-repulsive and unable to form a collapsed hydrophobic core (6). Despite these characteristics, αS has been shown to be partially collapsed and to deviate from a true random coil structure (9–14). Moreover, the dynamic structure of αS has high environmental sensitivity that has largely hindered development of a cohesive characterization of the structural ensemble of monomeric αS in solution (6,10–18). Solution conditions including temperature, salt concentration, and pH all have substantial effects on the partial folding and collapse of αS. Furthermore, cosolvents and osmolytes have been shown to drive conformational readjustment of αS (6,19–22). In particular, seminal work by Uversky, Fink, and coworkers (23) demonstrated the ability of trimethylamine-*N*-oxide

Submitted August 15, 2017, and accepted for publication November 6, 2017.

*Correspondence: ejpetersson@sas.upenn.edu

Editor: David Eliezer.

https://doi.org/10.1016/j.bpj.2017.11.006

(TMAO) to promote compaction and partial helical folding of αS monomers, and in high concentrations drive the formation of helical oligomers. Intriguingly, these helical conformations are morphologically distinct from the helical structure formed on membranes, and the presence of osmolytes has been show to affect fibril formation (24).

TMAO is a naturally occurring amphiphilic osmolyte that is found in several aquatic organisms, where it counteracts the destabilizing effects of high concentrations of urea required for regulation of osmotic pressure (25–27). Experimental and theoretical efforts have afforded an effective model by which preferential exclusion of TMAO from the protein backbone and sequestration of water by TMAO promotes the formation of intramolecular hydrogen bonds and a reduction in exposed surface area for the protein (28–34). Solution studies of TMAO have revealed that the large 4.67 D dipole moment prompts significant water ordering around each molecule (35). Comprised of a total of ∼13 water molecules, direct coordination of water to the oxygen along with formation of a clathrate like structure about the methyl groups produces a first solvation shell with a 6 Å radius, and elicits an excluded volume effect that entropically drives protein compaction (28,35–37). In addition to the effects of excluded volume and water arrangement, TMAO has been proposed to act as a nanocrowder and also serves as a poor solvent of the peptide backbone (38,39). Single-molecule Förster resonance energy transfer (FRET) has been previously used by Deniz and coworkers (24,40) to investigate the compact structure of αS in TMAO. These studies were performed with a limited set of measurements and demonstrated that successive compaction of αS occurs with increasing concentrations of TMAO while maintaining a single, broadly distributed conformational state. Moreover, the work by Deniz and coworkers showed that despite the emergence of a partially helical secondary structure, the folding pathway and resulting structure were distinct from helices formed on membranes. Although a putative structural ensemble was not proposed in this work, our laboratory and others have demonstrated the ability to apply distances and distributions obtained from single molecule FRET data as constraints for Monte Carlo (MC) simulations (15). In the context of the present problem of TMAO compaction of αS, we wish to further develop a method for employing FRET constraints to generate experimentally constrained models of IDP ensembles.

Although the use of FRET data has not been extensively explored as a basis for generating structural ensembles of IDPs, the application of other long-range measurements has demonstrated success. Work by Forman-Kay, Zweckstetter, Blackledge, and others has shown the efficacy of utilizing data from paramagnetic relaxation enhancement (PRE) NMR experiments as restraints for simulating ensembles of IDPs (12–14,41,42). Dobson, Vendruscolo, and Eliezer have also applied NMR-based methods to study the disordered ensemble of αS (10,11). Moreover, efforts by Langen have demonstrated the similar usefulness of electron paramagnetic resonance (5,43). To date, most studies have focused on the application of PRE data, which lacks description of the underlying distribution of states (42). Single-molecule FRET has demonstrated an exceptional ability to visualize subpopulations of disordered ensembles (44). Furthermore, Best and Schuler have addressed some of the major concerns surrounding distance extraction from FRET data (17,18). Studies of chemically denatured ubiquitin have shown that FRET and small angle x-ray scattering (SAXS) data afford comparable molecular sizes, and that the inclusion of large hydrophobic probe molecules does not significantly impact the structural ensemble (17). Moreover, recent work by Schuler has verified that distributions extracted from single-molecule FRET data display a high degree of agreement with distributions compiled from structural ensembles generated from molecular dynamics simulations restrained with NMR and SAXS data (18).

In this report, we focus on elucidating the structural changes associated with TMAO-induced compaction of αS by combining MC simulations in PyRosetta with experimental constraints from ensemble FRET measurements. Although single-molecule FRET measurements provide additional information regarding the number of distributions as well as the distribution breadth associated with a particular average value, the photophysical requisites for these measurements (i.e., bright, visible wavelength fluorophore pairs, which tend to have 30–70 Å working ranges) generally preclude accurate measurement of distances below ∼30 Å. This limitation is important, as we wish to generate atomically detailed computational models of αS and therefore need short distance constraints for our simulations. Based on the aforementioned results from Deniz and coworkers (24,40,44), we assume that the observed ensemble FRET efficiencies are resultant from single distributions, well described by polymer physics models, allowing us to rely on distances obtained from single distribution analysis. Furthermore, we propose that the most effective set of constraints would encompass not only long distances, such as those traditionally afforded from single-molecule FRET, but also short distances closer to those obtained via PRE measurements, which should more effectively limit the conformational variability within the ensemble.

Previously, we have studied αS compaction using the *p*-cyanophenylalanine (Cnf) and thioamide probe pair, which has a short Förster radius (R_0, the distance of half-maximal energy transfer) of 18 Å (45–47). We showed that increases in FRET efficiency (E_{FRET}) indicating compaction are observable as αS is incubated with increasing concentrations of TMAO. However, these Cnf-thioamide-labeled proteins required semisynthesis through native chemical ligation, a method with insufficient throughput for FRET library construction. Herein, we employ Cnf-tryptophan (Trp) as a genetically incorporable, short-range probe pair allowing for more facile production

of a library of labeled constructs for measuring distinct intramolecular distances in the presence of TMAO (48,49). Additionally, we have incorporated a second, longer-range probe pair, fluorescein-5-maleimide (Fam) conjugated to cysteine and tetramethylrhodamine azide (Raz) conjugated to an *O*-propargyl tyrosine unnatural amino acid (50). The Fam-Raz pair has a working distance around 50 Å and provides accuracy in ranges equivalent to those accessible through single-molecule FRET.

By applying measurements from these two complementary FRET libraries as weighted constraints for atomically detailed MC simulations in PyRosetta, we construct model conformational ensembles that agree well with experimental data such as fluorescence correlation spectroscopy (FCS), reporting on the overall average size of the protein. This strategy of using FRET data to direct all atom models in MC simulations provides a means to generate structure-guided hypotheses for allosteric transitions in αS to understand the effects of changes in environment or interactions with ligands.

MATERIALS AND METHODS

Protein overexpression and purification

Protein expression was performed in *Escherichia coli*, where unnatural amino acids were incorporated via amber stop codon suppression and traceless purification was facilitated via attachment of a C-terminal intein containing a C-terminal His-tag (Fig. 1). For the Cnf-Trp library, all native tyrosine residues were mutated to phenylalanine to assure that all energy transfer occurred exclusively between Cnf and Trp. Details of postexpression dye attachment, purification, and construct confirmation are reported in the Supporting Material.

FRET measurements in TMAO

All labeled αS variants were dialyzed into 20 mM Tris, 100 mM NaCl (pH 7.5). TMAO containing buffers were also prepared with 20 mM Tris, 100 mM NaCl, and the pH was adjusted to 7.5 after the addition of TMAO. Concentrations for the Cnf-Trp library were determined using the Sigma-Aldrich FluoroProfile Quantification Kit, whereas concentrations for the Fam-Raz library were determined via ultraviolet-visible absorbance. Steady-state measurements for the Cnf-Trp library and time-correlated single photon counting (TCSPC) measurements of the Fam-Raz library were performed at a 1 μM concentration, whereas steady-state measurements for the Fam-Raz library were performed at a 100 nM concentration. Measurements were taken in triplicate in 0, 2, and 4 M TMAO. TCSPC measurements were performed for Fam-Raz constructs under solution conditions identical to the steady-state measurements. Intramolecular FRET measurements were performed at concentrations of 1 μM, whereas intermolecular FRET measurements were performed by mixing single-labeled αS containing Fam with single-labeled αS containing Raz, at a concentration of 1 μM for each labeled construct.

Analysis of steady-state and TCSPC data

E_{FRET} values from steady-state measurements were determined by minimizing the squared difference between the spectrum from the double-labeled construct and a linear weighted sum of two single-labeled constructs using the equation

$$\sum_{\lambda} \left(I(\lambda)_{\mathrm{DA}} - AI(\lambda)_{\mathrm{D}} - BI(\lambda)_{\mathrm{A}}\right)^2 \rightarrow \min. \qquad (1)$$

In Eq. 1, $I(\lambda)$ represents the intensity as a function of wavelength for the construct indicated by the subscript, where A, D, and DA refer to constructs containing the acceptor fluorophore, the donor fluorophore, and both fluorophores, respectively. The weighting values A and B were used to determine E_{FRET} values from donor quenching and donor-sensitized emission of the acceptor to produce an error-weighted E_{FRET} value as described in the Supporting Material. The interresidue distances were calculated from E_{FRET} using a polymer-scaled version of the Förster equation:

$$E_{FRET} = \sum_{r} P_n(r,x) \Big/ \left(1 + (r/R_0)^6\right), \qquad (2)$$

$$P_1(r,x) = 4\pi r \left(\frac{3}{2\pi x^2}\right)^{3/2} \exp\left(-\frac{3}{2}\frac{r^2}{x^2}\right), \qquad (3)$$

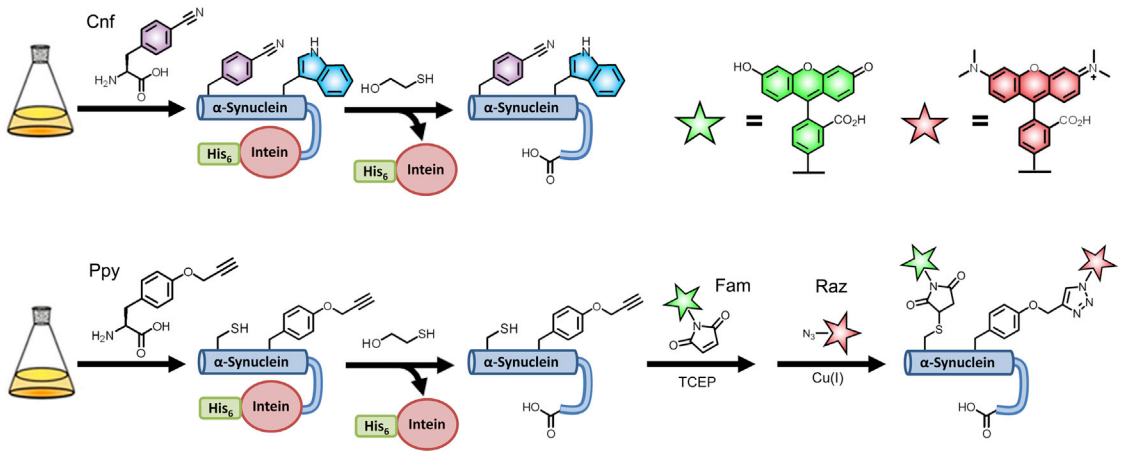

FIGURE 1 Scheme for labeled protein production. Top: Direct incorporation of *p*-cyanophenylalanine (Cnf) and tryptophan (Trp) via unnatural amino acid mutagenesis and conventional mutagenesis, respectively. Bottom: Incorporation of Cysteine (Cys) and *O*-propargyl tyrosine (Ppy) with subsequent labeling by fluorescein-5-maleimide (Fam) and tetramethylrhodamine azide (Raz), respectively. To see this figure in color, go online.

$$P_2(r,x)\,| \;=\; \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\frac{(r-x)^2}{\sigma^2}\right), \qquad (4)$$

where r and x represent the interresidue distance and the average interresidue distance. The reported distances determined from E_{FRET} values are the average interresidue distances. Two different polymer distributions, $P(r,x)$, based on the Gaussian chain model (Eq. 3) and a semiempirically derived normal distribution (Eq. 4), where the standard deviation, σ, was determined from polymer-liked simulation (detailed in the Supporting Material), were used to determine average distances. Additionally, fluorescence lifetimes from constructs containing only Fam (τ_D) and from constructs containing both Fam and Raz (τ_{DA}) were used to calculate the ratiometric change in lifetime (τ_{DA}/τ_D), and thus E_{FRET} values for comparison with steady-state data. Full details regarding TCSPC analysis are provided in the Supporting Material.

FCS

FCS measurements of 20 nM αS labeled with Alexa Fluor 488 at positions 9, 114, and 130 were taken in the presence of 0, 2, and 4 M TMAO using a laboratory-built instrument based on an Olympus IX71 microscope described in further detail in the Supporting Material. Before addition of labeled protein, chambers were incubated with 400 μM wild-type αS to prevent sticking of the labeled protein to the surface. Free Alexa Fluor 488 dye was measured in each concentration of TMAO to account for changes in viscosity. Additional information regarding FCS measurement and analysis can be found in the Supporting Material.

PyRosetta simulations

MC simulations were performed in PyRosetta on the University of Pennsylvania School of Arts and Sciences General Purpose Cluster. The simulation sampling includes ~1.5×10^6 backbone φ/ψ torsion angle changes in a simulated annealing score function gradient along with sample of side-chain rotamers. Constraints were applied either in the form of a harmonic potential or a potential derived from the Gaussian chain probability distribution. Output structures represent the lowest energy structures generated over the course of the simulation as determined by the sum of the "beta" Rosetta score function and the constraint energy. Additional details and the base PyRosetta script are provided in the Supporting Material.

RESULTS

We began by generating two libraries of proteins, labeled with either Cnf-Trp or Fam-Raz pairs, and making FRET measurements in varying concentrations of TMAO. For both FRET pairs, we observed changes in photophysical parameters impacting the extraction of distance information from FRET data that necessitated performing control fluorescence measurements using single-labeled proteins. When calculating R_0 (Eq. S8 in the Supporting Materials and Methods) for these experiments, there were four parameters of interest, Φ_D, J, κ^2 and n, which represent the quantum yield of the FRET donor, the overlap integral between donor emission and acceptor absorbance, the orientation factor between the two fluorophores, and the refractive index of the solution, respectively. Of these parameters, we determined that changes in Φ_D as a function of environment contributed most significantly to changes in R_0 for both Cnf and Fam. Additionally, Zheng et al. (18) previously demonstrated that for disordered proteins, the orientations of the fluorophores are sufficiently isotropic to warrant the approximation of $\kappa^2 = 2/3$, which is the value for an isotropic distribution of orientations. Fluorescence measurements made on donor-only and acceptor-only proteins enabled us to not only account for changes in Φ_D to enhance our accuracy in calculating R_0, but also allowed us to easily overcome the difficulties of interpreting highly overlapped spectra and extract distances from E_{FRET} measurements for use in modeling.

Cnf-Trp library

The Cnf-Trp construct library consisted of a total of 27 proteins (17 double-labeled and 10 single-labeled αS mutants) spanning 16 unique intramolecular distances. Acquiring concentration-matched emission spectra for each single-labeled construct in varying TMAO conditions allowed for tracking of changes in the quantum yield and spectral shape of Cnf and Trp emission. The sensitivity of these photophysical properties to increasing concentrations of TMAO was initially assessed for the free amino acids (Fig. S43). Although a decrease in the quantum yield was observed for both Cnf and Trp, no major change in emission maximum was observed in the Trp spectrum. In contrast to the free amino acid measurements, Trp-containing αS mutants successively blue-shifted in increasing concentrations of TMAO, as a result of changes in local environment upon compaction. Thus, measuring single-labeled spectra was not only crucial for tracking spectral changes due to environmental effects, but was also essential for deconvoluting the highly overlapped Cnf and Trp emission spectra in constructs containing both fluorophores. E_{FRET} values were obtained for each probe pair in buffer containing 0, 2, and 4 M TMAO. Fig. 2 shows the deconvolution of a 0 M measurement as well as a 2 M measurement for the Cnf_{125}-Trp_{94} pair. As expected, E_{FRET} increases, correlating with the expected decrease in interresidue distance resulting from compaction in TMAO. Without TMAO, in most cases this probe pair is not able to accurately capture the apparent intraresidue distance, as seen in low E_{FRET} values outside of the most reliable FRET range (E_{FRET} = 0.3–0.7). This is most clearly observed in Fig. 3, where the distances extracted from these data are largely invariant above sequence separations of 50 residues. However, in cases where the probes are relatively close in primary sequence, the utility of this short-range pair is clearly demonstrated. In the presence of 2 M, and especially 4 M TMAO, a significant number of the probed regions display E_{FRET} values within the optimal range for accurate distance determination.

Fam-Raz library

To complement our Cnf-Trp data, we also performed FRET measurements with a probe pair with a longer working

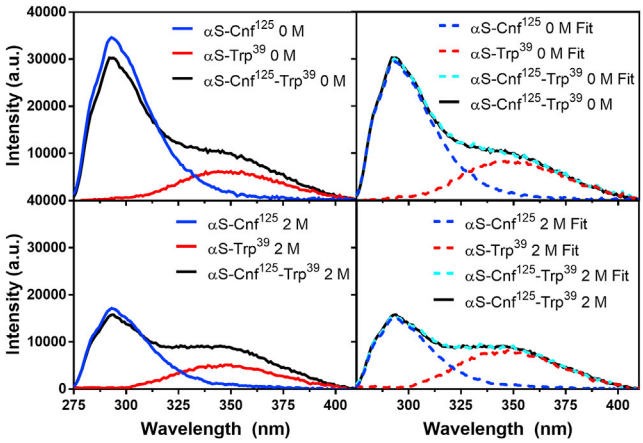

FIGURE 2 Determination of E_{FRET}. Left: Background subtracted fluorescence emission spectra of constructs labeled with Cnf, Trp, or both Cnf and Trp in 0 M (*top*) and 2 M TMAO (*bottom*). Right: Double-labeled Cnf-Trp spectrum compared to weighted sum of Cnf-only and Trp-only spectra, along with the contributions from each single-labeled spectrum shown for 0 M (*top*) and 2 M TMAO (*bottom*) spectra. To see this figure in color, go online.

range. A set of 21 constructs, consisting of 10 dual-labeled and 11 single-labeled analogs containing Fam and/or Raz, was produced to accurately monitor long-range distance changes during the compaction of αS by TMAO. In cases where exceedingly low E_{FRET} was observed with the short-range probe pair, we observed efficiencies much closer to the optimal efficiency range for the Fam-Raz pair. This was most significant in the absence of TMAO. However, our measurements approached the short end of the working range for the Fam-Raz pair in some instances in 2 M TMAO, and exhausted its utility in 4 M TMAO.

In 4 M TMAO, the long working range of this probe pair was effective for assessing the formation of oligomers, which were reported by Uversky et al. (23). When mixing Fam and Raz single-labeled species in 4 M TMAO, there was a significant reduction in the lifetime of the Fam labeled construct, consistent with intermolecular FRET (Figs. S48–S50). In buffer or 2 M TMAO, there was no observed change in the lifetime of the Fam-labeled protein due to the presence of the protein containing Raz. FCS and atomic force microscopy measurements also demonstrate the formation of oligomers in 4 M TMAO (Figs. S51 and S52). As a result of these observations, we decided to forego modeling the 4 M structure due to the complexity of deconvoluting intra- and intermolecular FRET in oligomers.

Interpretation of FRET data

Because αS is disordered, the interresidue distance separating any given label pair is widely distributed. To accurately obtain an average distance value, this distribution needs to be taken into account. Therefore, the corresponding interresidue distances were calculated using a polymer-scaled Förster equation, Eq. 2, taking into account changes in R_0 resulting from changes in quantum yield for each labeling position. Distances were calculated using both the Gaussian chain (Eq. 3) and semiempirical (Eq. 4) models, which produced significantly different sets of results. Conceptually, two

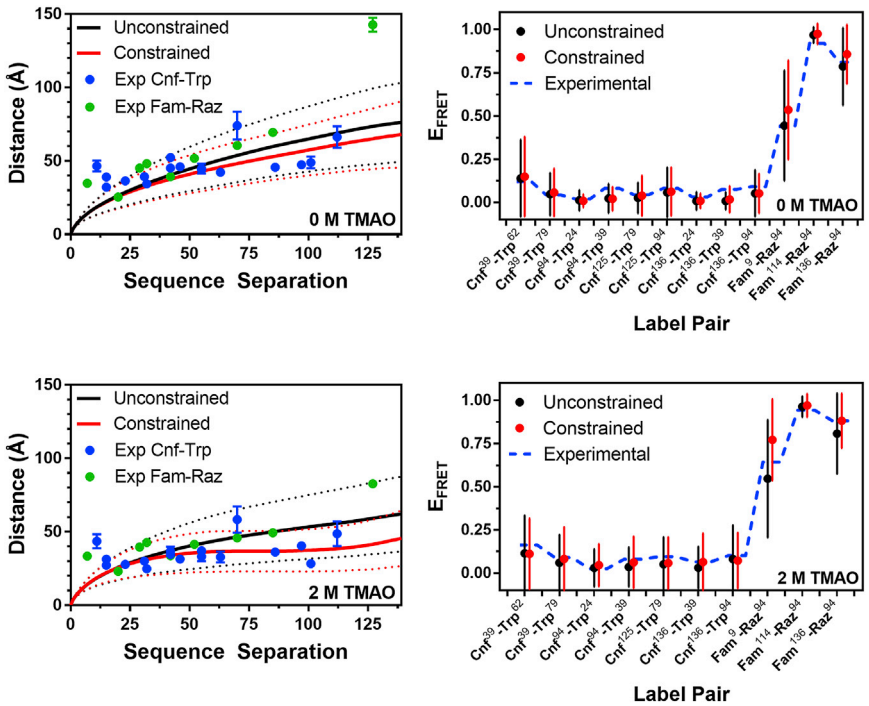

FIGURE 3 Comparison of experimental and simulated data. Left: Distances extracted from E_{FRET} measurements of the Cnf-Trp and Fam-Raz libraries shown with interresidue distances for a given primary sequence separation in unconstrained or constrained MC simulations. The average (*solid line*) and standard deviation (*dashed line*) of interresidue distances are shown for the simulated ensembles. Right: Experimental and simulated E_{FRET} values. Dashes represent average E_{FRET} values for each pair of labeled residues, with interconnecting lines to guide the eye. The average (*point*) and standard deviation (*line*) of E_{FRET} values obtained based on interresidue separations in the unconstrained or constrained simulated ensembles. To see this figure in color, go online.

positions labeled with two FRET pairs with different R_0 values should have distinct FRET efficiencies, but comparable extracted distances. Since the Cnf-Trp and Fam-Raz libraries contained a limited number of labeled positions in common, for comparison the distances were plotted in Fig. S44 as a function of probe pair primary sequence separation. We reasoned that the consistency between data from the Cnf-Trp and Fam-Raz libraries on a Flory scaling plot should help us to identify the more accurate probability distribution. In Fig. S44, it is clear that there is substantially better agreement between the Cnf-Trp and Fam-Raz data sets when applying the Gaussian chain distribution to extract distance data from the E_{FRET} values.

Simulations

After data collection, simulations were performed in PyRosetta to generate structural ensembles that represent the changes observed by FRET. Before introducing FRET-based constraints, we optimized a PyRosetta script to effectively produce ensembles of structures in general agreement with previously published radius of gyration (R_g) and radius of hydration (R_h) values (21,51). Simulations were performed where the weighting of the FRET constraints relative to other Rosetta energy terms (a, Eqs. S18 and S19 in the Supporting Materials and Methods) was varied, and different shapes for the FRET constraint function were assessed. These constraint function shapes were based on the Gaussian chain (Eq. 3) or semiempirical (Eq. 4) distance distributions used in the FRET analyses. Given the greater consistency between Cnf-Trp and Fam-Raz data when using the Gaussian chain probability distribution (above), we favored using this function for implementing constraints as well. However, we tested all combinations to ensure that no bias was introduced based on the pairing of distribution functions used in FRET data analysis and in constraint implementation. The results of these tests are detailed in the Supporting Material (Figs. S53–S64) and are discussed further below. We found that setting the weight of the constraints relative to the other score function components to unity allowed the FRET constraints to influence the structure without overconstraining, and that reasonable variations of the shape of the constraint function did not dramatically influence the simulation results. Thus, in the main text, we exclusively report simulations performed using the Gaussian chain distribution, with all other simulations reported in the Supporting Material. The 2 M TMAO simulations were performed with the solvation term removed in an effort to account for the significant change in solvation. Unconstrained simulations performed with this altered score function provided a more compact starting point for introducing FRET constraints, whereas constrained ensembles were noticeably more compact (Figs. S56–S58). Moreover, for simulations of the 2 M ensemble, fragment insertion was incorporated within the MC search to increase the amount of resultant secondary structure, which has been observed by circular dichroism studies of the TMAO-induced conformation (23). Although this significantly increased the number of helices, there was no marked improvement in the match of simulated structures with experimental FRET data (Figs. S56–S58).

Initially, we compared the constrained simulations to the FRET data using Flory scaling plots, where the average interresidue distance (between α-carbons) as a function of sequence separation was plotted for the ensemble of 1000 lowest energy structures with experimental FRET data overlaid as discrete points (Fig. 3, *left*). All experimental data were close to the average simulation distance or within the standard deviation for both 0 and 2 M ensembles, with the notable exception of the Fam_9-Raz_{136} FRET data. It is notable that the distance determined for this FRET pair using the harmonic potential (Fig. S44) does not show such dramatic discord with the other distances, potentially suggesting a limitation of the Gaussian chain model for very long distance ranges. We also predicted FRET values for all of the pairs of label sites for comparison to the experimental FRET data (Fig. 3, *right*). Distances between the α-carbons of the amino acids at the label sites were extracted from each structure in a simulated structural ensemble and converted to E_{FRET} values using the classical Förster equation (Eq. 2, where $P(r,x) = 1$). Since the variations in conformation in the ensemble explicitly capture the distribution of interresidue distances, the average E_{FRET} was computed as a simple average of the values extracted from each structure with no further correction for polymer scaling of the distance distribution. We found good agreement, with an average absolute E_{FRET} difference of 0.09 and all experimental values falling within the standard deviation of the simulated values. As discussed below, the close match between these values helps to validate our choice of a polymer-scaled distance distribution function in interpreting E_{FRET} values to input constraints in the simulations.

Fig. 4 shows the 10 lowest energy structures from the 0 and 2 M simulations, aligned and with a single structure darkened for clarity. One can see significant compaction of the 2 M ensemble, which can be quantified using histograms of the R_g for each structure in the ensemble. The average radius of gyration in 0 M TMAO is 29.3 $\pm$ 4.6 Å, which shrinks to 23.5 $\pm$ 3.4 Å in 2 M TMAO. Interestingly, despite this compaction, αS remains disordered in 2 M TMAO. The disorder can be visualized using a plot of the distribution of interresidue distances as a two-dimensional "heat map," with interresidue distances for the entire ensemble plotted above the diagonal and interresidue distances for the single darkened structure below the diagonal. As seen in the middle plots in Fig. 4, in both 0 and 2 M TMAO, the average interresidue distance scales with primary sequence separation, despite the fact that long-range contacts occur in individual structures within the ensembles (Fig. S67). This combined information indicated that, in the presence of 2 M TMAO, αS populates a

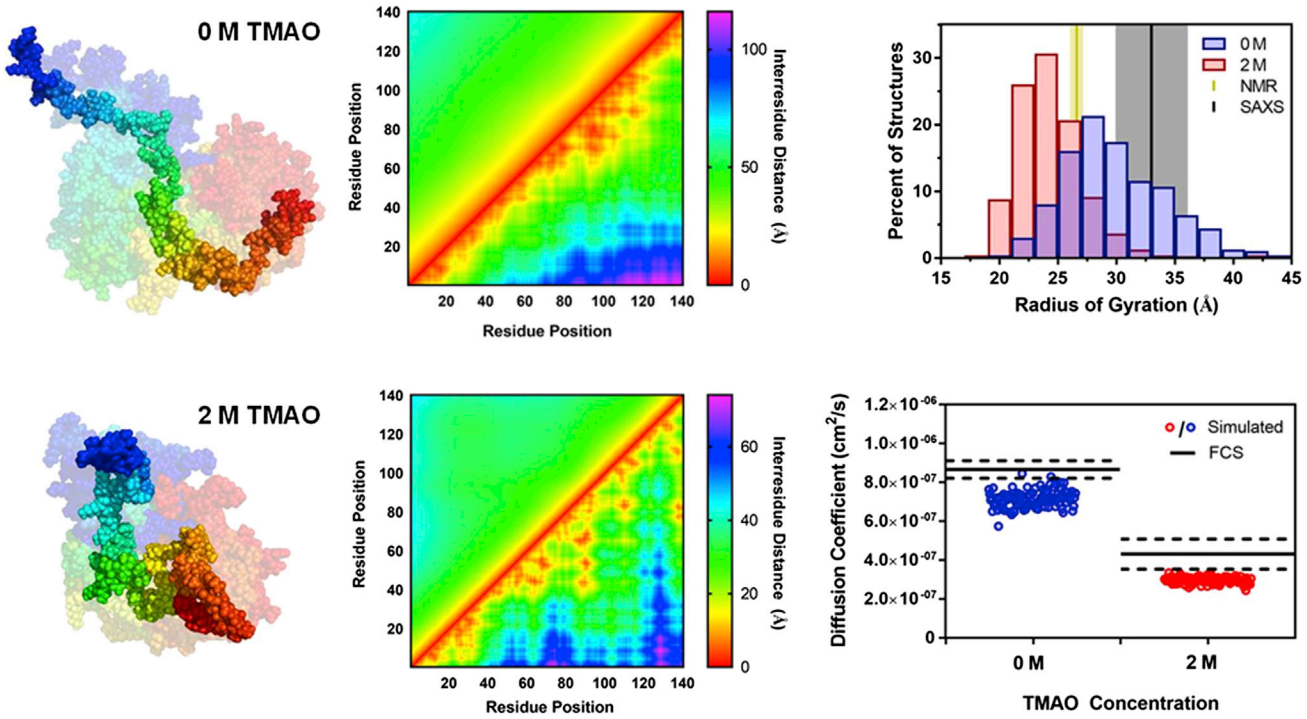

FIGURE 4 Analysis of αS structural ensembles. Left: Representative structures from 0 and 2 M simulations. The darkened structure in the foreground is the lowest energy structure and the faded ensemble in the background includes the 10 lowest energy structures. Middle: Heat maps showing the average interresidue distances in the simulated ensembles. Each heat map shows the average distance for the full simulated ensemble of structures above the diagonal and distances from the single lowest energy structure below the diagonal. Top right: Histograms of radii of gyration of structures from simulations, plotted with literature values of R_g determined from SAXS data and the hydrodynamic radius from NMR studies. Bottom right: Diffusion coefficients from FCS and simulated structures. To see this figure in color, go online.

compacted disordered ensemble that still maintains a high degree of structural heterogeneity. To confirm that the structure ensembles were consistent with other experimental observables, we calculated the diffusion coefficient for each structure of the 0 and 2 M ensembles using HydroPro (52). As seen in Fig. 4, the calculated diffusion coefficients are slightly (16%) smaller, but in reasonable agreement with those measured in our FCS experiments.

Comparison of structure ensembles to data from literature

We compared our structural models in the absence of TMAO to data from the literature as well as other published models. The average R_g of our 0 M ensemble matched well with reported values of R_g from SAXS experiments and R_h (typically 30% larger than R_g) from NMR experiments (21,51). Furthermore, our models agreed remarkably well with both the primary NMR PRE data (Fig. S66) and Flory scaling plots of the resulting models (Fig. S65) reported by Allison et al. (10). We do observe rare contacts between regions of the protein that are distant in primary sequence (Fig. S67), consistent with reports of transient C-terminal contacts with the NAC and N-terminus (14). Although there is limited information on the structure of αS in 2 M TMAO, we find that our models agree with the available information. Similar to previous studies, our circular dichroism measurements made in 0 or 2 M TMAO show little change in αS helicity (Fig. S73), consistent with our models which show that the 2 M ensembles are still highly disordered. Previously, Uversky and coworkers monitored fibrillization of αS in TMAO using thioflavin T (ThT) fluorescence, and found that aggregation was accelerated, but that the final level of ThT fluorescence was lower. Our own studies with Congo Red (Fig. S71) corroborated this observation, although, intriguingly, changes in fluorescence polarization (Fig. S72) were slower in 2 M TMAO, possibly indicating that the ThT and Congo Red spectroscopic properties are changed in TMAO. Thus, although compacted, αS remains disordered and able to sample conformations that lead to fibril formation. This can be seen in Figs. S69 and S70 where, despite compaction observed by changes in solvent-accessible surface area, metrics such as numbers of intramolecular hydrogen bonds and backbone dihedral angles do not indicate the presence of persistent structure.

DISCUSSION

Our previous study of αS structures using MC simulations with only a repulsive Lennard-Jones potential and

harmonic constraints based on single-molecule FRET data gave structural ensembles with global properties that matched well to experimental measurements such as R_g. This modeling protocol, with constrained simulations utilizing exclusively a repulsive van der Waals potential, was extremely efficient in its simplicity, but the resulting models lacked atomic-scale details that could be used in generating hypotheses or interpreting mechanisms of conformational change. We and others have also performed unconstrained simulations which included the amino acids using single-sphere "centroid" representations of the side chains, as well as all-atom molecular dynamics simulations of αS, which do provide such detailed information, but are much more computationally intensive and generally limited in the conformational space explored. We wished to find an intermediate level of simulation wherein we could include side chains in MC simulations and maintain efficient sampling. However, we reasoned that the previous long-range constraints (>30 Å) derived from single-molecule FRET data would not provide sufficient information on short-range interactions to properly direct these simulations. Thus, we here included data from the short-range Cnf-Trp pair along with data from the Fam-Raz pair, which has a comparable FRET range to the previous single-molecule FRET probes. These short-range probe pairs may be crucial for identifying contacts or collapsed regions that can exist under varied solution conditions or in the presence of allosteric molecules.

Assessing the validity of $P(r)$

We reasoned at the outset that the Gaussian chain $P(r)$ was more accurate because it led to greater consistency between the distances obtained from the Cnf-Trp and Fam-Raz libraries. However, this method of analysis was complicated by the limited numbers of intramolecular distances in a range that could be accurately captured by both pairs, as these probes were selected specifically for their efficacy across different distance ranges. Furthermore, one can observe that the Cnf-Trp distances are largely invariant after reaching a sequences separation of ∼50 residues due to a working range that extends to only ∼35 Å. Therefore, our assignment of the Gaussian chain $P(r)$ as the more accurate distribution function arises from visual inspection of the Flory scaling plots in Fig. S44 in the 25–45 Å range, which encompass the upper and lower bounds of the Cnf-Trp and Fam-Raz probe pairs, respectively. It is important to note that at very short (<15 residues) and very long (>115 residues) sequence separations, the calculated distances are likely unable to be accommodated by any conformation of the protein. Furthermore, the functional form of the distribution may also be dependent on the number of residues between two probes (e.g., short sequence separations cannot have truly polymer-like behavior and may need to be treated with classical FRET equations) or heterogeneity not accounted for by polymer-scaling behaviors (53).

Differences between constraint methods

Since IDPs have relatively flat energetic landscapes, improper introduction of constraints can easily result in ensembles where the resultant conformations are not sufficiently diverse, especially proximal to constrained sites. This arises primarily from constraining potentials that are too deep or too narrow, but can also occur when conformational space is not effectively sampled. The latter problem could be caused by poor parameterization of the Metropolis criterion (such as selection of a kT value that is too low) or by not having an appropriate score function to produce the multiple local minima present within a single pair-wise constraint potential. By expanding our score function from a single repulsive van der Waals term to the current optimal score function utilized by Rosetta, we assure that output states populate local minima that allow us to extract high-resolution information and provide an effective buffer for constraint introduction. Achieving unconstrained simulations of this quality required significant adjustment of PyRosetta scripts, which normally are used to model folded proteins rather than IDPs. These optimization efforts will be reported in detail elsewhere. In addition to improved modeling scripts, we hypothesized that issues related to the constraint functional form could be circumvented by directly employing appropriate distance distributions for disordered ensembles to analyze the FRET data.

Nonetheless, we performed simulations with all four combinations of the two $P(r)$ functions used to obtain distances and the two corresponding constraining functions. This allowed us to consider how the ensembles were influenced by the method of constraining in addition to the sets of distances used as constraints. Figs. S56–S63 demonstrate that neither the method of constraint nor the $P(r)$ utilized for the distance determination through Eq. 2 had a dramatic impact on the resulting structures. This likely arises from the fact that distances from the unconstrained simulation are already very close to the distances obtained from the experiments. Despite the relative agreement between the ensembles produced across all constraining methods for a given concentration of TMAO, there are several noteworthy observations. Figs. S60–S63 show that in all cases in which a harmonic potential was employed, the structural diversity, most clearly observable in the dispersion of R_g values, is visibly decreased. Moreover, harmonic constraints consistently produced deviations in the scaling behavior, where the heat maps (Figs. S60–S64) reveal distances between the N-terminal region and the NAC domain that are, on average, longer than those between the C-terminal region and the NAC. This observation is independent of the input data, as these nuances are observed for both data sets, with constraints from data obtained utilizing $P_2(r,x)$ further

enhancing these trends. These observations demonstrate that there are significant differences between these ensembles which result exclusively from the functional form of the constraining potential.

The relative impacts of changes in the score function or in the constraints are evident in comparisons between 2 M simulations performed using Gaussian chain derived data with or without constraints or solvation term modification (Figs. S56, S60–S63). FRET constraints alone compacted the 2 M ensemble (2 M GC-GC) relative to the 0 M ensemble (0 M GC-GC), but were not sufficient to generate ensembles that were consistent with the 2 M experimental data. However, it is important to note the marked differences between the 0 and 2 M ensembles, as these demonstrate that simulations in PyRosetta can be significantly influenced by constraints, which was not evident from the 0 M simulations alone. The fact that these constraints alone were unable to produce ensembles in full agreement with the 2 M data suggests that αS in the presence of molar quantities of TMAO populates a subset of conformations that are not accessible with the standard score function. With the solvation term removed, comparison of the 2 M constrained (2M GC-GC†, see Fig. S62) and unconstrained (Unconstrained†, see Fig. S61) simulations demonstrate that the constraints serve to exclude extended structures. This further suggests that the application of constraints does not produce new conformations not present in the unconstrained population, but can effectively remove unfavorable structures from an ensemble based on experimental data. Overall, we see this as a favorable compromise as the necessity to modulate the unconstrained population through changes to the PyRosetta scoring is likely driven by the dramatic change in buffer conditions, which would not be present in many other applications. Systems where dramatic changes in solvent conditions occur (as in this study) can be effectively represented by adjusting the score function, whereas intramolecular (or intermolecular) contacts can be accounted for by changing the constraint functional form to accurately represent experimental data.

Comparison of experimental and simulated data

Initial comparisons between the experimental and simulated distances in Figs. 3 and S56 demonstrate that most of the obtained distances fall within 1 SD of the average interresidue distance for the given sequence separation. The observed disagreements at high and low sequence separation are likely due to inaccuracies in the determination of the experimental distance arising from either the working range of the probe or the polymer-scaling function, as previously discussed. Moreover, agreement in simulated and experimental FRET efficiencies in Figs. 3, S57, and S58 demonstrate that the underlying conformation dispersion is accurately captured, since appropriate efficiencies are obtained for both probe pairs. The FCS measurements herein also demonstrate that the simulated ensembles are of approximately the correct overall size and that the degree of compaction is qualitatively accurate. The quantitative disagreement between the simulated diffusion coefficients observed in Fig. 4, where the simulated diffusion coefficients are systematically lower than the experimental values, could arise from several different factors. Given that the overall size agrees with previously published NMR and SAXS data, it is possible that we encountered small, systematic inaccuracies in determining diffusion coefficients resulting from a combination of the treatment of the disordered ensemble in HydroPro, the need for an empirical conversion factor, or accounting for the change in viscosity induced by TMAO. Importantly, we have demonstrated in Figs. S65 and S66, respectively, that our 0 M ensembles agree with previously published PRE data and the structural ensembles generated from those data (10,11).

It is worth noting that in our effort to establish an efficient, intermediate level of simulation, we chose not to pursue some elements that could have further improved the accuracy or tested the boundaries of our simulations. For example, we elected not to incorporate probe/linker spatial exploration. Although we acknowledge that simulations would be made more accurate by including representations of the probes, with FRET based on distances between fluorophore transition dipoles instead of α-carbons, this would have significantly increased the computational time, undermining our intention of creating an efficient approach. Additionally, we envision that coupling our method with the strategy previously employed by Dobson and Vendruscolo, where constraining functions were iteratively updated, could produce a more refined, but computationally intensive, version of our modeling protocol (11). Moreover, this method circumvents the need to assume a probability distribution function when interpreting the FRET data, as the constraints would be applied directly as FRET efficiencies. Finally, since introduction of a new constraint functional form is as simple as writing a new function in Python, we hypothesize that our method of constraint introduction, where assumptions regarding underlying distributions are directly converted into a potential, is not only useful for applying FRET data from disordered systems, but would also allow for efficient incorporation of other types of experimental data such as PRE data (10,13,41).

CONCLUSIONS

We have developed a Rosetta modeling protocol using explicit protein side chains and sophisticated score functions in combination with appropriately weighted distance constraints to generate models of IDPs. By performing simulations that were constrained with experimental FRET data from two libraries containing different FRET pairs, we were able to model the ensemble of αS in buffer

and in the presence of 2 M TMAO. Our models agreed well with independent measurements of αS structure from FCS, NMR, PRE, and SAXS data, and were computationally less taxing than traditional molecular dynamics simulations. In future work, we intend to explore the degree to which ensembles generated in PyRosetta capture residual secondary structure and the accuracy of values computed from these ensembles, such as chemical shifts or J-couplings in comparison with NMR experiments. The modifications made to our Rosetta modeling protocol allowed us to produce reasonable αS starting models in the absence of constraints. The quality of these unconstrained models gives us greater confidence in interpreting the interactions observed in the constrained structural ensembles, and makes the simulations more robust to the inclusion of an inconsistent constraint. After careful consideration of constraint function shape and weight, we have found a form that allows them to influence the structural ensemble without overconstraining. Future investigations could employ different weights for different FRET pairs, or different functional forms for different distance ranges.

Our modeling protocol for IDPs incorporates atomic detail relevant to the study of chemical-, ligand-, or environment-induced conformational changes, and yet is sufficiently rapid both in data collection and simulation time to be applied in a moderate throughput fashion. Once libraries of labeled proteins have been generated, FRET measurements of the type reported here could easily be acquired under a variety of conditions, potentially even in a multiwell format. The collection of FRET constraints could then be used to generate structural ensembles to explore hypotheses for mechanisms of conformational change or to rationalize trends among molecules and modifications that modulate IDP conformation. For example, Kakish et al. (54) have recently shown that bis-heterocycles linked by a flexible tether are able to bind to αS and induce conformational changes in the monomer that inhibit its propensity to aggregate. The procedures used here to study the effects of TMAO on αS structure could be applied to study such molecules with therapeutic potential. Furthermore, recent work has shown that modifications such as serine glycosylation and tyrosine phosphorylation affect aggregation and membrane binding respectively (55,56). Subsequent modification of the current labeled library, or production of a new library, would allow one to utilize the method reported herein to visualize the effect of these and other posttranslational modifications on the disordered ensemble. Although certain classes of molecules or solution conditions may spectroscopically interfere with our FRET probes, it is important to note that we can easily vary these probes using essentially the same labeling strategies. For example, we have recently reported a methoxycoumarin-acridonylalanine FRET pair that can be introduced by a combination of cysteine modification and unnatural amino acid mutagenesis (57). Finally, it is important to note that our methods are not restricted to pure IDPs like monomeric αS, but can be applied to disordered regions of folded proteins or ordered aggregates such as the N- and C-terminal regions of fibrillar αS (58,59).

SUPPORTING MATERIAL

Supporting Materials and Methods, seventy-three figures, and eight tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(17)31213-4.

AUTHOR CONTRIBUTIONS

J.J.F. prepared the Cnf-Trp library and performed all steady-state and lifetime measurements and simulations. C.M.H. prepared the Fam-Raz library. J.Y. assisted in preparation of the Cnf-Trp library. B.P. performed FCS measurements (advised by E.R.). Y.-C.L. performed atomic force microscopy measurements (advised by Z.F.). A.N. and J.J.F. designed and performed computational modeling. J.J.F. and E.J.P. drafted the manuscript. All authors contributed to manuscript editing.

ACKNOWLEDGMENTS

We would like to thank Profs. Michele Vendruscolo and Chris Dobson for sharing PRE data.

This work was supported by the University of Pennsylvania, the National Institutes of Health (grants NS081033 to E.J.P. and NS079955 to E.R.), the Alfred P. Sloan Foundation (to Z.F.), and the University of Washington Royalty Research Fund (to A.N.). Instruments supported by the National Science Foundation include matrix-assisted laser desorption mass spectrometry and circular dichrosim (NSF MRI-0820996) and CD (DMR05-20020). J.J.F., C.M.H., and J.Y. thank the following for fellowship support: National Science Foundation (DGE-1321851), Age Related Neurodegenerative Disease training grant (NIH T32 AG000255), and the University of Pennsylvania Center for Undergraduate Research and Fellowships.

REFERENCES

1. Lashuel, H. A., C. R. Overk, …, E. Masliah. 2013. The many faces of α-synuclein: from structure and toxicity to therapeutic target. *Nat. Rev. Neurosci.* 14:38–48.
2. Auluck, P. K., G. Caraveo, and S. Lindquist. 2010. α-Synuclein: membrane interactions and toxicity in Parkinson's disease. *Annu. Rev. Cell Dev. Biol.* 26:211–233.
3. Bendor, J. T., T. P. Logan, and R. H. Edwards. 2013. The function of α-synuclein. *Neuron.* 79:1044–1066.
4. Ulmer, T. S., A. Bax, …, R. L. Nussbaum. 2005. Structure and dynamics of micelle-bound human α-synuclein. *J. Biol. Chem.* 280:9595–9603.
5. Jao, C. C., B. G. Hegde, …, R. Langen. 2008. Structure of membrane-bound alpha-synuclein from site-directed spin labeling and computational refinement. *Proc. Natl. Acad. Sci. USA.* 105:19666–19671.
6. Uversky, V. N. 2009. Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J.* 28:305–325.
7. Trexler, A. J., and E. Rhoades. 2009. α-synuclein binds large unilamellar vesicles as an extended helix. *Biochemistry.* 48:2304–2306.
8. Eliezer, D., E. Kutluay, …, G. Browne. 2001. Conformational properties of alpha-synuclein in its free and lipid-associated states. *J. Mol. Biol.* 307:1061–1073.

9. Smith, W. W., C. F. Schreck, …, C. S. O'Hern. 2012. Molecular simulations of the fluctuating conformational dynamics of intrinsically disordered proteins. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 86:041910.
10. Allison, J. R., P. Varnai, …, M. Vendruscolo. 2009. Determination of the free energy landscape of α-synuclein using spin label nuclear magnetic resonance measurements. *J. Am. Chem. Soc.* 131:18314–18326.
11. Dedmon, M. M., K. Lindorff-Larsen, …, C. M. Dobson. 2005. Mapping long-range interactions in α-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J. Am. Chem. Soc.* 127:476–477.
12. Jeschke, G. 2016. Ensemble models of proteins and protein domains based on distance distribution restraints. *Proteins.* 84:544–560.
13. Salmon, L., G. Nodet, …, M. Blackledge. 2010. NMR characterization of long-range order in intrinsically disordered proteins. *J. Am. Chem. Soc.* 132:8407–8418.
14. Bertoncini, C. W., Y.-S. Jung, …, M. Zweckstetter. 2005. Release of long-range tertiary interactions potentiates aggregation of natively unstructured α-synuclein. *Proc. Natl. Acad. Sci. USA.* 102:1430–1435.
15. Nath, A., M. Sammalkorpi, …, E. Rhoades. 2012. The conformational ensembles of α-synuclein and tau: combining single-molecule FRET and simulations. *Biophys. J.* 103:1940–1949.
16. Sung, Y. H., and D. Eliezer. 2007. Residual structure, backbone dynamics, and interactions within the synuclein family. *J. Mol. Biol.* 372:689–707.
17. Aznauryan, M., L. Delgado, …, B. Schuler. 2016. Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule FRET, NMR, and SAXS. *Proc. Natl. Acad. Sci. USA.* 113:E5389–E5398.
18. Zheng, W., A. Borgia, …, R. B. Best. 2016. Probing the action of chemical denaturant on an intrinsically disordered protein by simulation and experiment. *J. Am. Chem. Soc.* 138:11702–11713.
19. Ruzza, P., R. Hussain, …, G. Siligardi. 2015. Effects of trehalose on thermodynamic properties of alpha-synuclein revealed through synchrotron radiation circular dichroism. *Biomolecules.* 5:724–734.
20. McNulty, B. C., A. Tripathy, …, G. J. Pielak. 2006. Temperature-induced reversible conformational change in the first 100 residues of alpha-synuclein. *Protein Sci.* 15:602–608.
21. Morar, A. S., A. Olteanu, …, G. J. Pielak. 2001. Solvent-induced collapse of α-synuclein and acid-denatured cytochrome c. *Protein Sci.* 10:2195–2199.
22. Bai, J., M. Liu, …, C. Li. 2017. Macromolecular and small molecular crowding have similar effects on α-synuclein structure. *ChemPhysChem.* 18:55–58.
23. Uversky, V. N., J. Li, and A. L. Fink. 2001. Trimethylamine-N-oxide-induced folding of alpha-synuclein. *FEBS Lett.* 509:31–35.
24. Moosa, M. M., A. C. Ferreon, and A. A. Deniz. 2015. Forced folding of a disordered protein accesses an alternative folding landscape. *ChemPhysChem.* 16:90–94.
25. Yancey, P. H. 2005. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J. Exp. Biol.* 208:2819–2830.
26. Canchi, D. R., and A. E. Garcia. 2013. Cosolvent effects on protein stability. *In* Annual Review of Physical Chemistry, *Volume 64*. M. A. Johnson, and T. J. Martinez, eds.. Annual Reviews, Palo Alto, CA, pp. 273–293.
27. 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.
28. Knake, L., G. Schwaab, …, M. Havenith. 2015. Solvation dynamics of trimethylamine N-oxide in aqueous solution probed by terahertz spectroscopy. *J. Phys. Chem. B.* 119:13842–13851.
29. Chang, Y.-C., and T. G. Oas. 2010. Osmolyte-induced folding of an intrinsically disordered protein: folding mechanism in the absence of ligand. *Biochemistry.* 49:5086–5096.
30. Kokubo, H., C. Y. Hu, and B. M. Pettitt. 2011. Peptide conformational preferences in osmolyte solutions: transfer free energies of decaalanine. *J. Am. Chem. Soc.* 133:1849–1858.
31. Hu, C. Y., H. Kokubo, …, B. M. Pettitt. 2010. Backbone additivity in the transfer model of protein solvation. *Protein Sci.* 19:1011–1022.
32. Hu, C. Y., G. C. Lynch, …, B. M. Pettitt. 2010. Trimethylamine N-oxide influence on the backbone of proteins: an oligoglycine model. *Proteins.* 78:695–704.
33. Holthauzen, L. M. F., J. Rösgen, and D. W. Bolen. 2010. Hydrogen bonding progressively strengthens upon transfer of the protein urea-denatured state to water and protecting osmolytes. *Biochemistry.* 49:1310–1318.
34. Baskakov, I., and D. W. Bolen. 1998. Forcing thermodynamically unfolded proteins to fold. *J. Biol. Chem.* 273:4831–4834.
35. Larini, L., and J.-E. Shea. 2013. Double resolution model for studying TMAO/water effective interactions. *J. Phys. Chem. B.* 117:13268–13277.
36. Doi, H., Y. Watanabe, and M. Aida. 2014. Influence of trimethylamine *N*-Oxide (TMAO) on the three-dimensional distribution and alignment of solvent molecules in aqueous solution. *Chem. Lett.* 43:865–867.
37. Sagle, L. B., K. Cimatu, …, P. S. Cremer. 2011. Methyl groups of trimethylamine N-oxide orient away from hydrophobic interfaces. *J. Am. Chem. Soc.* 133:18707–18712.
38. Ma, J., I. M. Pazos, and F. Gai. 2014. Microscopic insights into the protein-stabilizing effect of trimethylamine N-oxide (TMAO). *Proc. Natl. Acad. Sci. USA.* 111:8476–8481.
39. Cho, S. S., G. Reddy, …, D. Thirumalai. 2011. Entropic stabilization of proteins by TMAO. *J. Phys. Chem. B.* 115:13401–13407.
40. Ferreon, A. C., M. M. Moosa, …, A. A. Deniz. 2012. Counteracting chemical chaperone effects on the single-molecule α-synuclein structural landscape. *Proc. Natl. Acad. Sci. USA.* 109:17826–17831.
41. Marsh, J. A., and J. D. Forman-Kay. 2012. Ensemble modeling of protein disordered states: experimental restraint contributions and validation. *Proteins.* 80:556–572.
42. Konrat, R. 2014. NMR contributions to structural dynamics studies of intrinsically disordered proteins. *J. Magn. Reson.* 241:74–85.
43. Rao, J. N., C. C. Jao, …, T. S. Ulmer. 2010. A combinatorial NMR and EPR approach for evaluating the structural ensemble of partially folded proteins. *J. Am. Chem. Soc.* 132:8657–8668.
44. Schuler, B., A. Soranno, …, D. Nettels. 2016. Single-molecule FRET spectroscopy and the polymer physics of unfolded and intrinsically disordered proteins. *In* Annual Review of Biophysics, *Volume 45*. K. A. Dill, ed.. Annual Reviews, Palo Alto, CA, pp. 207–231.
45. Wissner, R. F., S. Batjargal, …, E. J. Petersson. 2013. Labeling proteins with fluorophore/thioamide Förster resonant energy transfer pairs by combining unnatural amino acid mutagenesis and native chemical ligation. *J. Am. Chem. Soc.* 135:6529–6540.
46. Wissner, R. F., A. M. Wagner, …, E. J. Petersson. 2013. Efficient, traceless semi-synthesis of α-synuclein labeled with a fluorophore/thioamide FRET pair. *Synlett.* 24:2454–2458.
47. Batjargal, S., Y. J. Wang, …, E. J. Petersson. 2012. Native chemical ligation of thioamide-containing peptides: development and application to the synthesis of labeled α-synuclein for misfolding studies. *J. Am. Chem. Soc.* 134:9172–9182.
48. Glasscock, J. M., Y. Zhu, …, F. Gai. 2008. Using an amino acid fluorescence resonance energy transfer pair to probe protein unfolding: application to the villin headpiece subdomain and the LysM domain. *Biochemistry.* 47:11070–11076.
49. Taskent-Sezgin, H., J. Chung, …, I. Carrico. 2009. Interpretation of p-cyanophenylalanine fluorescence in proteins in terms of solvent exposure and contribution of side-chain quenchers: a combined fluorescence, IR and molecular dynamics study. *Biochemistry.* 48:9040–9046.
50. Haney, C. M., R. F. Wissner, …, E. J. Petersson. 2016. Comparison of strategies for non-perturbing labeling of α-synuclein to study amyloidogenesis. *Org. Biomol. Chem.* 14:1584–1592.

51. Araki, K., N. Yagi, …, H. Mochizuki. 2016. A small-angle X-ray scattering study of alpha-synuclein from human red blood cells. *Sci. Rep.* 6:30473.

52. Ortega, A., D. Amorós, and J. García de la Torre. 2011. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* 101:892–898.

53. Song, J., G.-N. Gomes, …, H. S. Chan. 2017. Conformational heterogeneity and FRET data interpretation for dimensions of unfolded proteins. *Biophys. J.* 113:1012–1024.

54. Kakish, J., K. J. H. Allen, …, J. S. Lee. 2016. Novel dimer compounds that bind α-synuclein can rescue cell growth in a yeast model overexpressing α-synuclein. A possible prevention strategy for Parkinson's disease. *ACS Chem. Neurosci.* 7:1671–1680.

55. Dikiy, I., B. Fauvet, …, D. Eliezer. 2016. Semisynthetic and in vitro phosphorylation of alpha-synuclein at Y39 promotes functional partly helical membrane-bound states resembling those induced by PD mutations. *ACS Chem. Biol.* 11:2428–2437.

56. Lewis, Y. E., A. Galesic, …, M. R. Pratt. 2017. O-GlcNAcylation of α-synuclein at serine 87 reduces aggregation without affecting membrane binding. *ACS Chem. Biol.* 12:1020–1027.

57. Ferrie, J. J., N. Ieda, …, E. J. Petersson. 2017. Multicolor protein FRET with tryptophan, selective coumarin-cysteine labeling, and genetic acridonylalanine encoding. *Chem. Commun. (Camb.).* 53:11072–11075.

58. Tuttle, M. D., G. Comellas, …, C. M. Rienstra. 2016. Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. *Nat. Struct. Mol. Biol.* 23:409–415.

59. Del Mar, C., E. A. Greenbaum, …, V. L. Woods, Jr. 2005. Structure and properties of α-synuclein and other amyloids determined at the amino acid level. *Proc. Natl. Acad. Sci. USA.* 102:15477–15482.

Biophysical Journal, Volume *114*

Supplemental Information

Using a FRET Library with Multiple Probe Pairs To Drive Monte Carlo Simulations of α-Synuclein

John J. Ferrie, Conor M. Haney, Jimin Yoon, Buyan Pan, Yi-Chih Lin, Zahra Fakhraai, Elizabeth Rhoades, Abhinav Nath, and E. James Petersson

Using a Large Ensemble FRET Library with Multiple Probe Pairs to Drive Monte Carlo Simulations of Disordered α-Synuclein

Supplementary Information

John J. Ferrie,[a] Conor M. Haney,[a] Jimin Yoon,[a] Buyan Pan,[a] Yi-Chih Lin,[a] Zahra Fakhraai,[a] Elizabeth Rhoades,[a] Abhinav Nath,[b] and E. James Petersson[a,*]

[a]Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104-6323, USA

[b]Department of Medicinal Chemistry, School of Pharmacy, University of Washington, Box 357610, Seattle, WA 98195-7610, USA

Table of Contents

General Information:

Materials: *E. coli* BL21(DE3) cells were purchased from Stratagene (La Jolla, CA, USA). Milli-Q filtered (18 MΩ) water was used for all solutions (Millipore; Billerica, MA, USA). Bradford reagent assay kits were purchased from BioRAD (Hercules, CA, USA). The unnatural amino acid *p*-cyanophenylalanine was purchased from PepTech (Bedford, MA, USA; Cat. # AL240-2). Fluorescein-5-maleimide (Fam) was purchased from TCI America (Portland, OR, USA; Cat. # F0810); carboxytetramethylrhodamine-5-azide (Raz) was ordered from Lumiprobe (Hallendale Beach, FL, USA; Cat # D7130) and AlexaFluor488 maleimide, C5 isomer, was purchased from ThermoFisher Scientific (Pittsburgh, PA, USA; Cat. # A10254). *O*-Propargyltyrosine (Ppy) was synthesized as previously described(1). Amicon Ultra centrifugal filter units (3 kDa MWCO) were purchased from EMD Millipore. All other reagents and solvents were purchased from Fisher Scientific (Pittsburgh, PA, USA) or Sigma-Aldrich unless otherwise specified. DNA sequencing was performed at the University of Pennsylvania DNA sequencing facility.

Instruments: Matrix assisted laser desorption/ionization with time-of-flight detector (MALDI-TOF) mass spectra were acquired on a Bruker Ultraflex III instrument (Billerica, MA, USA). UV-Vis absorption spectra were acquired on a Hewlett-Packard 8452A diode array spectrophotometer (currently Agilent Technologies; Santa Clara, CA, USA). Fluorescence spectra were collected on a Photon Technologies International (PTI) QuantaMaster40 fluorometer (currently Horiba Scientific, Edison, NJ, USA). Details regarding the fluorescence correlation spectroscopy (FCS) equipment are contained within the FCS Measurements section of the SI. Atomic force microscopy (AFM) measurements were made using a Keysight 5500 AFM instrument (Keysight Technologies; Santa Rosa, CA, USA) equipped with a closed-loop scanner**.**

Protein Synthesis

DNA Oligomers Used for αS Quikchange Mutagenesis:

Mutation	Direction	Primer Sequence
Y39F	Forward	5'-GGA AAG ACA AAA GAG GGT GTT CTC TTT GTA GGC TCC AAA-3'
	Reverse	5'-TTT GGA GCC TAC AAA GAG AAC ACC CTC TTT TGT CTT TCC-3'
Y125F	Forward	5'-GGA TCC TGA CAA TGA GGC TTT TGA AAT GCC TTC TGA-3'
	Reverse	5'-TCA GAA GGC ATT TCA AAA GCC TCA TTG TCA GGA TCC-3'
Y133F-Y136F	Forward	5'-CCT TCT GGA GAA GGG TTT CAA GAC TTC GAA CCT GAA GCC-3'
	Reverse	5'-GGC TTC AAG TTC GAA GTC TTG AAA CCC TTC CTC AGA AGG-3'
Y24W	Forward	5'-TCT GCC ACA CCC CAT TTG GTT TTC TCA GCA GCA GCC A-3'
	Reverse	5'-TGG CTG CTG CTG AGA AAA CCA AAT GGG GTG TGG CAG A-3'
Y39W	Forward	5'-ACA AAA GAG GGT GTT CTC TGG GTA GGC TCC AAA ACC AAG G-3'
	Reverse	5'-CCT TGG TTT TGG AGC CTA CCC AGA GAA CAC CCT CTT TTG T-3'
Q62W	Forward	5'-TGC TCC TCC AAC ATT TGT CAC CCA CTC TTT GGT CTT CTC AGC CAC-3'
	Reverse	5'-GTG GCT GAG AAG ACC AAA GAG TGG GTG ACA AAT GTT GGA GGA GCA-3'
Q79W	Forward	5'-CTC CCT CCA CTG TCT TCC AGG CTA CTG CTG TCA CAC-3'
	Reverse	5'-GTG TGA CAG CAG TAG CCT GGA AGA CAG TGG AGG GAG-3'
F94W	Forward	5'-TGC AGC AGC CAC TGG CTG GGT CAA AAA GGA CCA GTT C-3'
	Reverse	5'-CAA CTG GTC CTT TTT GAC CCA GCC AGT GGC TGC TGC A-3'
Y125W	Forward	5'-GAT CCT GAC AAT GAG GCT TGG GAA ATG CCT TCT GAG GAA G-3'
	Reverse	5'-CTT CCT CAG AGG GCA TTT CCC AAG CCT CAT TGT CAG GAT C-3'
Y39TAG	Forward	5'-AAA AGA GGG TGT TCT CTA GGT AGG TCT CAA AAC CAA G-3'
	Reverse	5'-CTT GGT TTT GGA GCC TAC CTA GAG AAC ACC CTC TTT T-3'
F94TAG	Forward	5'-GCA TTG CAG CAG CCA CTG GCT AGG TCA AAA AGG ACC AGT TGG G-3'
	Reverse	5'-CCC AAC TGG TCC TTT TTG ACC TAG CCA GTG GCT GCA ATG C-3'
Y125TAG	Forward	5'-CCT GAC AAT GAG GCT TAG GAA ATG CCT TCT GAG GA-3'
	Reverse	5'-TCC TCA GAA GGC ATT TCC TAA GCC TCA TTG TCA GG -3'
Y133F-Y136TAG	Forward	5'-TGC CTT CTG AGG AAG GGT TTC AAG ACT AGG AAC C-3'
	Reverse	5'-GGT TCC TAG TCT TGA AAC CCT TCC TCA GAA GGC A-3'

All of constructs and primers required for production of the Fam-Raz library have been previously reported. The construct containing the Y136TAG mutation, primers for all of Cys mutation primers have also been previously reported (2-4).

General Expression Protocol for αS-intein-H_6 Fusion Proteins: A pTXB1 plasmid encoding for the αS-intein-H_6 fusion protein was transformed into *E. coli* BL21-Gold (DE3) cells. The DNA GyrA intein from Mxe fusion construct has been previously reported (4). Transformed cells were selected based on ampicillin (Amp) (100 mg/L) resistance, encoded on the pTXB1 plasmid containing the protein of interest. Single colonies were selected and used to inoculate 5 mL primary cultures of LB media, and were grown at 37 °C, shaking at 250 rpm in the presence of 100 μg/mL Amp. A secondary culture of LB media containing 100 mg/L Amp was inoculated with a single saturated primary culture, and was grown at 37 °C while shaking at 250 rpm. After reaching an OD_{600} of 0.7-1.0, 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the culture to induce expression. Following induction the culture continued to grow overnight at 18 °C overnight.

Expression of αS-intein-H_6 Proteins Containing Unnatural Amino Acids: The procedure for the expression of mutants containing Cnf or Ppy was identical to the general protocol above with the following noted exceptions. The pTXB1 plasmid coding for the protein of interest contained an amber stop codon (TAG) at the intended site for unnatural amino acid introduction was co-transformed with a plasmid encoding a pDULE2-pFX plasmid containing an orthogonal aminoacyl synthetase/tRNA pair (4). pDULE2-pFX encodes streptomycin (Strep) resistance, and cells containing both plasmids were selected for based on Strep (100mg/L) resistance in addition to Amp resistance. Primary cultures were also grown in the presence of both Strep (100 μg/mL) and Amp (100 μg/mL). In place of LB, secondary cultures were grown in M9 minimal media containing 6 g Na_2HPO_4, 3 g KH_2PO_4, 0.5 g NaCl and 1 mL of 2 M $MgSO_4$, 1 mL of 1mg/mL $FeCl_2$ (in 1.0 M HCl), 1 mL of 15 mg/mL, 2 mL of 10% Yeast Extract, 12.5 mL 40% glucose (w/v) in 1L of autoclaved water, along with 100 mg/L Amp and 100 mg/L Strep. Lastly, once the secondary culture reached an OD_{600} of 0.7-1.0, 0.8 mM of the unnatural amino acid was added to culture prior to induction with IPTG. Expression was performed overnight at 18 °C.

General Purification Protocol for Cnf-Trp containing αS-intein-H_6 Fusion Proteins: Cells were harvested via centrifugation at 5000 rpm for 20 minutes at 4 °C with a GS3 rotor on a Sorvall RC-5 centrifuge. The supernatant was decanted away from the cell pellet, and the pellet was resuspended in 20 mL resuspension buffer containing 40 mM Tris, 5 mM EDTA, pH 8.3 and a protease inhibitor tablet (Roche cOmplete mini tabs, EDTA free). Cells were lysed by sonication

at an amplitude of 30 for 5 minutes with 1 second on 1 second off. Lysate was subsequently centrifuged at 14000 rpm for 20 minutes using an SS-34 rotor on a Sorvall RC-5 centrifuge. His tagged protein was isolated from the supernatant via nickel affinity. Ni-NTA resin (3 mL CV) was incubated with the supernatant on ice for 1 hour. Following incubation the mixture was loaded into a fritted column, and the lysate was drained. The resin was initially washed with 15 mL of 50 mM HEPES, pH 7.5 followed by 20 mL of 50 mM HEPES, 5 mM imidazole, pH 7.5, prior to elution of the protein of interest with 12 mL of 50 mM HEPES, 300 mM imidazole, pH 7.5. Cleavage of the intein was performed via addition of β-mecaptoethanol (βME) to a final concentration of 200 mM and incubation at room temperature for 18 hours on a rotisserie. Following cleavage of the intein, removal of imidazole and βME was facilitated by dialysis into 20 mM tris base pH 8.0 at 4 °C overnight. The undesired intein was removed by a second Ni column (3 mL CV). The flow though containing the protein of interest was collected after 1 hour incubation with the Ni resin on ice.

General Labeling Protocol for Fam-Raz containing αS-intein-H_6 Fusion Proteins: αS variants containing Cys and/or Ppy were expressed as described above and purified by Ni-NTA affinity. Following the second purification subsequent to intein cleavage, the desired protein was dialyzed into 20 mM Tris, pH 8.0 overnight at 4 °C. The protein was then labeled in this semi-crude state in one (donor-only or acceptor-only) or two (doubly-labeled) labeling steps as described below. Proteins lacking an unnatural amino acid (αS-Cys) were quantified by UV-Vis spectroscopy using ε_{280} = 5120 M^{-1} cm^{-1} with the exception of αS-Cys_{136}, which was quantified using ε_{280} = 3840 M^{-1} cm^{-1} (5). Protein variants containing Ppy were quantified using the BCA assay with bovine serum albumin (BSA) standards generated by two-fold serial dilution in water from 2 mg/mL to 0.125 mg/mL. Labeling reactions of proteins containing Cys were carried out by addition of BondBreaker® TCEP solution to a final concentration of 1 mM and incubation at room temperature for 10-15 minutes. Following this time, fluorescein-5-maleimide (Fam) was added from a 25 mM stock in DMSO to a final concentration equal to five-fold excess relative to protein (typical dye concentration was ~500-800 μM). The labeling reaction was shielded from light by wrapping in aluminum foil and incubated at 37 °C. The labeling reaction was monitored by MALDI-TOF MS following 5-fold dilution of the reaction into water. Following complete disappearance of unlabeled protein as determined by MALDI, the labeled product was dialyzed

overnight at 4 °C in 20 mM Tris pH 8.0. Proteins containing Ppy were labeled via copper(I) catalyzed azide-alkyne cycloaddition as described below; doubly-labeled proteins were first labeled with Fam as described above and subsequently by Raz. Azide-alkyne labeling reactions were performed by preparation of copper(I)-tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) catalyst mixture by using 80 mM $CuSO_4$ in water (1.25 μL/mL protein solution) to which was added 50 mM THPTA in water (30 μL/mL protein solution) and then 100 mM sodium ascorbate in water (30 μL/mL protein solution). This solution was incubated at room temperature for 10-15 minutes prior to addition to the protein solution. Protein was labeled by addition of five equivalents of Raz from a 50 mM stock in DMSO (typical dye concentrations were ~300-500 μM), followed by addition of the Cu(I) catalyst mixture. Labeling reactions were then shielded from light and incubated at 37 °C and reaction progress monitored by MALDI-TOF MS following 5-fold dilution of the reaction in water. Upon completion of the reaction, the labeled product was dialyzed into 20 mM Tris pH 8.0 overnight at 4 °C to remove excess dye.

Purification Protocol for All αS Proteins: All proteins were then purified by anion-exchange chromatography on an ÄKTA FPLC system using HiTrap Q HP columns and elution between 20 mM Tris pH 8.0 (buffer A) and 20 mM Tris, 1 M NaCl pH 8.0 (buffer B). Fractions containing the desired protein were identified by MALDI-TOF MS and were pooled. All proteins from the Fam-Raz library as well as poorly expressing member of the Cnf-Trp library were concentrated using Amicon Ultra 3 kDa filters to a total volume of ≤ 1 mL. Further purification was then performed using a Varian HPLC system and a Vydac C4 TP reverse-phase semi-preparative column with a flow rate of 4 mL/min and gradient between 0.1% TFA in water (solvent A) and 0.1% TFA in acetonitrile (solvent B). Fractions containing the desired product were identified by MALDI-TOF MS and judged to be of high purity were combined and diluted a minimum of 5-fold by addition of 20 mM Tris, 100 mM NaCl pH 7.5 and subsequently concentrated via Amicon μLtra 3 kDa centrifugal filters to a total volume of ≤ 5 mL, diluted 3-5 fold and re-concentrated; the final dilution and concentration was repeated twice. All proteins were then assessed for purity and integrity of the fluorescent label by MALDI-TOF analysis of whole protein and tryptic fragments.

Protein Library Characterization

Table S1. MALDI Masses from Cnf-Trp Library

	Full Length		Tyrpsin Digest				
Construct	Exp. [M+H]+	Obs. [M+H]+	Fragment	Exp. [M+H]+	Obs. [M+H]+	Exp. [M+Na]+	Obs. [M+Na]+
αS-Cnf^{39}	14421	14422.4	35-43	960.5	960.7	982.5	982.7
αS-Cnf^{94}	14421	14421.4	81-96	1504.7	1504.0	1526.7	1526.0
αS-Cnf^{125}	14421	14421.7	103-140	4266.4	4265.9	4288.4	4287.9
αS-Cnf^{136}	14421	14421	103-140	4266.4	4265.9	4288.4	4287.9
αS-Trp^{24}	14454	14455.6	24-32	888.5	888.6	910.5	910.6
αS-Trp^{39}	14435	14434.7	35-43	974.5	974.8	996.5	996.8
αS-Trp^{62}	14454	14454	61-80	1986.1	1986.5	2008.1	2008.5
αS-Trp^{79}	14454	14454.3	61-80	1986.1	1986.4	2008.1	2008.4
αS-Trp^{94}	14435	14435.4	81-96	1517.8	1518.1	1539.8	1540.1
αS-Trp^{125}	14435	14434.8	103-140	4280.5	4280.0	4302.5	4302.0
αS-Cnf^{39}-Trp^{24}	14479	14479.5	35-43	960.5	960.7	982.5	982.7
			24-32	888.5	888.7	910.5	910.7
αS-Cnf^{39}-Trp^{62}	14479	14480.1	35-43	960.5	960.7	982.5	982.7
			61-80	1986.1	1986.4	2008.1	2008.4
αS-Cnf^{39}-Trp^{79}	14479	14479.3	35-43	960.5	960.7	982.5	982.7
			61-80	1986.1	1986.4	2008.1	2008.4
αS-Cnf^{39}-Trp^{94}	14460	14461.6	35-43	960.5	960.7	982.5	982.7
			81-96	1517.8	1518.1	1539.8	1540.1
αS-Cnf^{94}-Trp^{24}	14479	14480.2	81-96	1504.7	1504.1	1526.7	1526.0
			24-32	888.5	888.6	910.5	910.6
αS-Cnf^{94}-Trp^{39}	14460	14461.4	81-97	1504.7	1504.1	1526.7	1526.0
			35-43	974.5	974.7	996.5	996.7
αS-Cnf^{94}-Trp^{62}	14479	14480.1	81-98	1504.7	1504.1	1526.7	1526.1
			61-80	1986.1	1986.4	2008.1	2008.4
αS-Cnf^{94}-Trp^{79}	14479	14480.1	81-99	1504.7	1504.1	1526.7	1526.1
			61-80	1986.1	1986.4	2008.1	2008.4
αS-Cnf^{125}-Trp^{24}	14479	14480.1	103-140	4266.4	4268.7	4288.4	4290.7
			24-32	888.5	888.6	910.5	910.6
αS-Cnf^{125}-Trp^{39}	14460	14461.5	103-140	4266.4	4265.8	4288.4	4287.8
			35-43	974.5	974.7	996.5	996.7
αS-Cnf^{125}-Trp^{62}	14479	14479.6	103-140	4266.4	4265.3	4288.4	4287.3
			61-80	1986.1	1986.4	2008.1	2008.4
αS-Cnf^{125}-Trp^{79}	14479	14480	103-140	4266.4	4265.5	4288.4	4287.5
			61-80	1986.1	1986.4	2008.1	2008.4
αS-Cnf^{125}-Trp^{94}	14460	14460	103-140	4266.4	4265.7	4288.4	4287.7
			81-96	1517.8	1518.1	1539.8	1540.0
αS-Cnf^{136}-Trp^{24}	14479	14478.5	103-140	4266.4	4265.9	4288.4	4287.9
			24-32	888.5	888.6	910.5	910.6
αS-Cnf^{136}-Trp^{39}	14460	14460.9	103-140	4266.4	4266.2	4288.4	4288.2
			35-43	974.5	974.7	996.5	996.7
αS-Cnf^{136}-Trp^{94}	14460	14461.6	103-140	4266.4	4265.5	4288.4	4287.5
			81-96	1517.8	1518.1	1539.8	1540.1
αS-Cnf^{136}-Trp^{125}	14460	14460.9	103-140	4305.5	4304.8	4327.5	4326.8

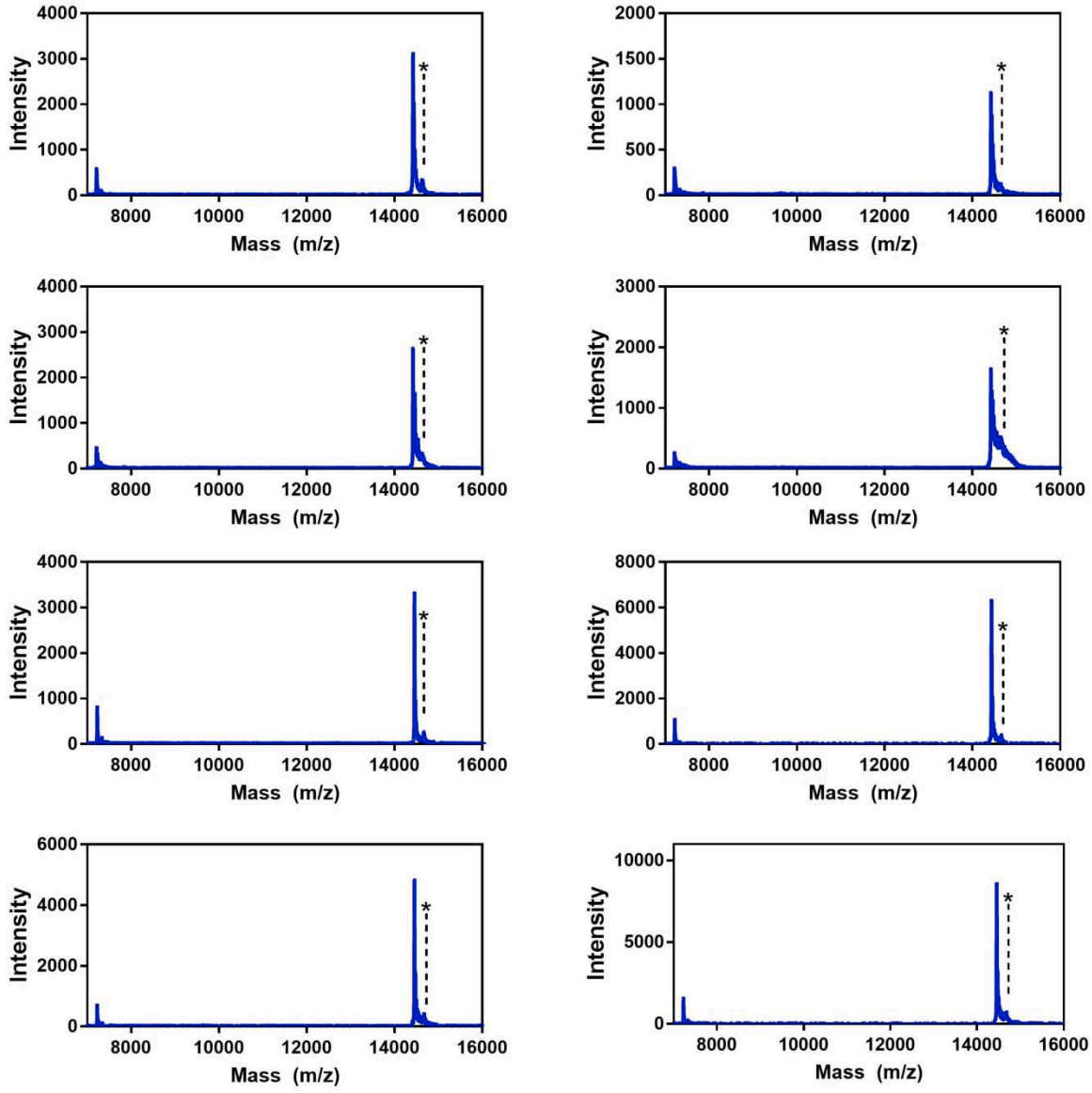

Figure S1: MALDI Spectra of Full-Length Cnf-Trp Constructs. Plots show αS-Cnf^{39} (top left), αS-Cnf^{94} (top right), αS-Cnf^{125} (upper middle left), αS-Cnf^{136} (upper middle right), αS-Trp^{24} (lower middle left), αS-Trp^{39} (lower middle right), αS-Trp^{62} (bottom left) and αS-Trp^{79} (bottom right). On each plot the matrix adduct is noted with *.

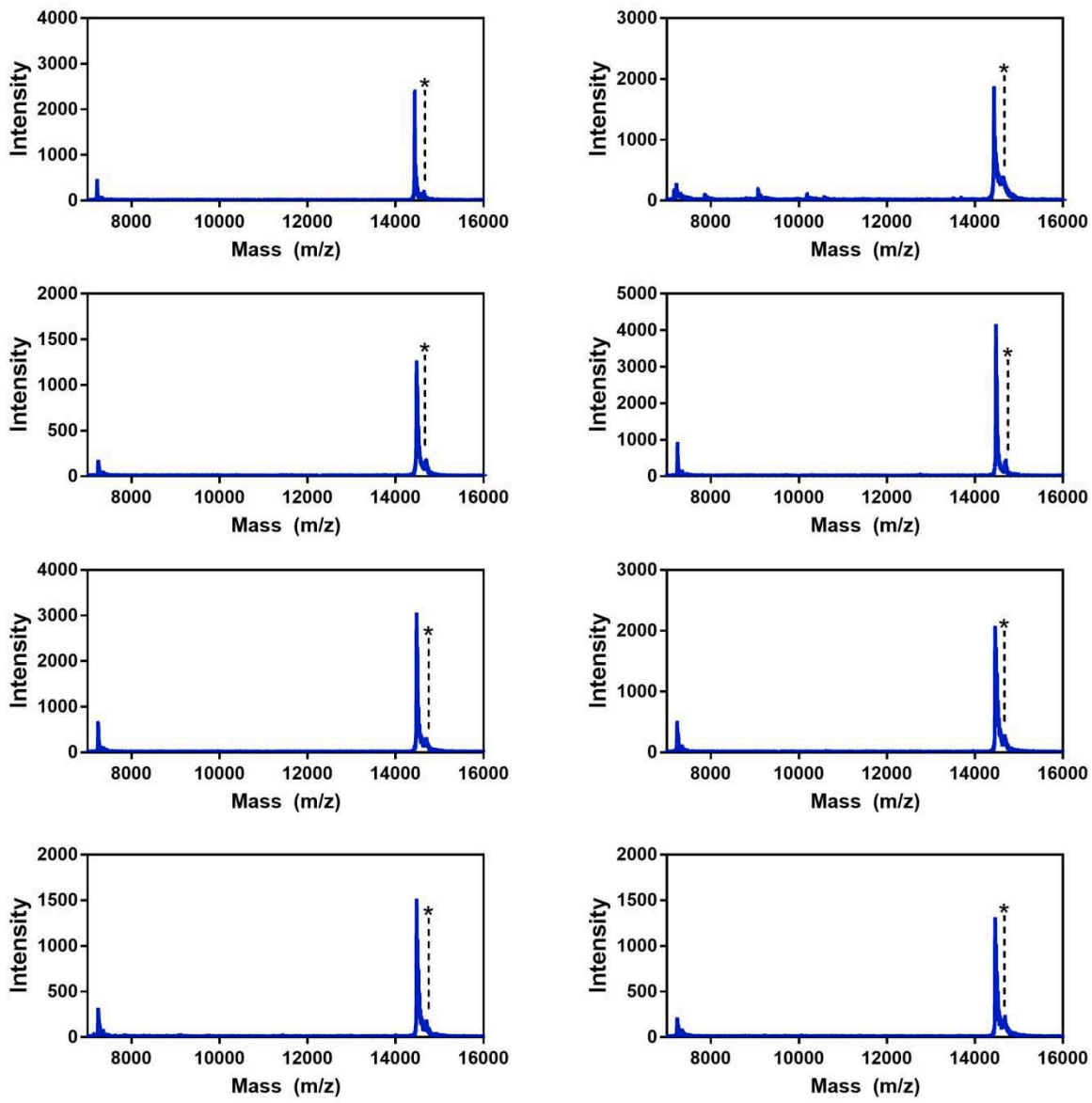

Figure S2: MALDI Spectra of Full-Length Cnf-Trp Constructs. Plots show αS-Trp^{94} (top left), αS-Trp^{125} (top right), αS-Cnf^{39}-Trp^{24} (upper middle left), αS-Cnf^{39}-Trp^{62} (upper middle right), αS-Cnf^{39}-Trp^{79} (lower middle left), αS-Cnf^{39}-Trp^{94} (lower middle right), αS-Cnf^{94}-Trp^{24} (bottom left) and αS-Cnf^{94}-Trp^{39} (bottom right). On each plot the matrix adduct is noted with *.

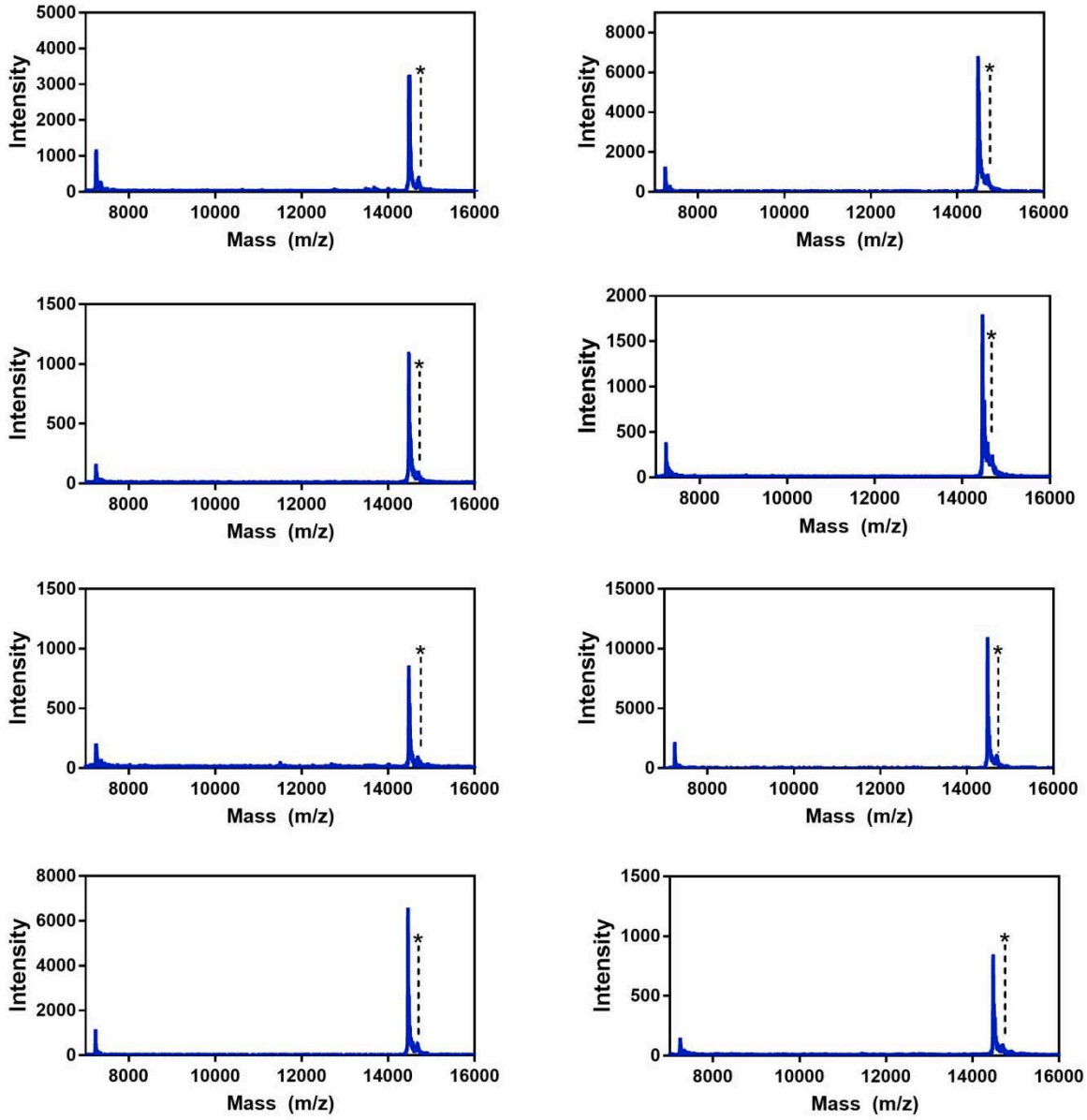

Figure S3: MALDI Spectra of Full-Length Cnf-Trp Constructs. Plots show αS-Cnf^{94}-Trp^{62} (top left), αS-Cnf^{94}-Trp^{79} (top right), αS-Cnf^{125}-Trp^{24} (upper middle left), αS αS-Cnf^{125}-Trp^{39} (upper middle right), αS-Cnf^{125}-Trp^{62} (lower middle left), αS-Cnf^{125}-Trp^{79} (lower middle right), αS-Cnf^{125}-Trp^{94} (bottom left) and αS-Cnf^{136}-Trp^{24} (bottom right). On each plot the matrix adduct is noted with *.

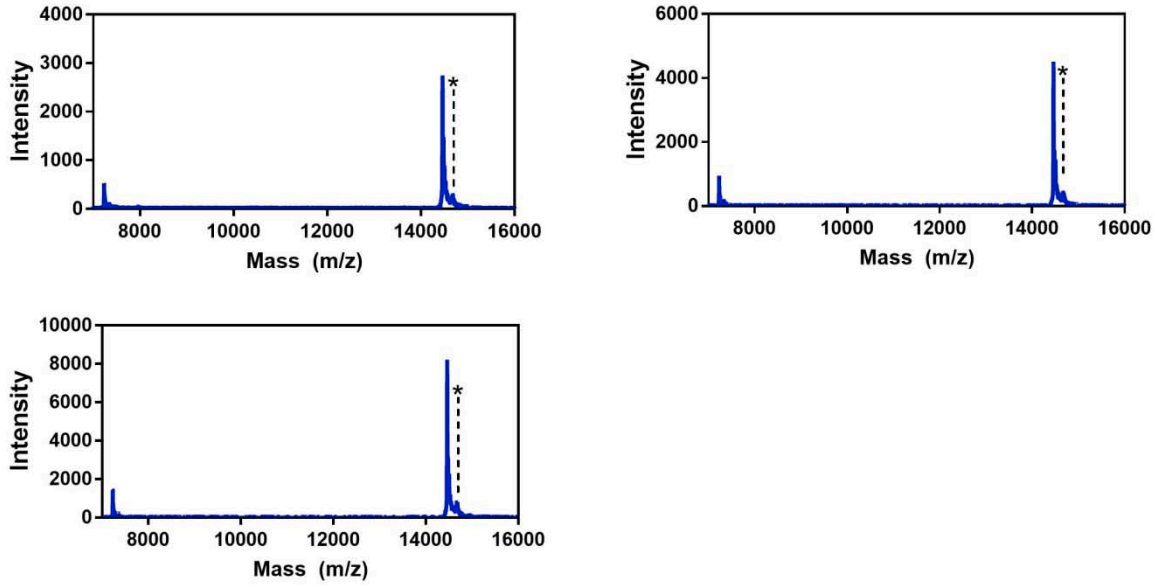

Figure S4: MALDI Spectra of Full-Length Cnf-Trp Constructs. Plots show αS-Cnf^{136}-Trp^{39} (top left), αS-Cnf^{136}-Trp^{94} (top right), αS-Cnf^{136}-Trp^{125} (bottom). On each plot the matrix adduct is noted with *.

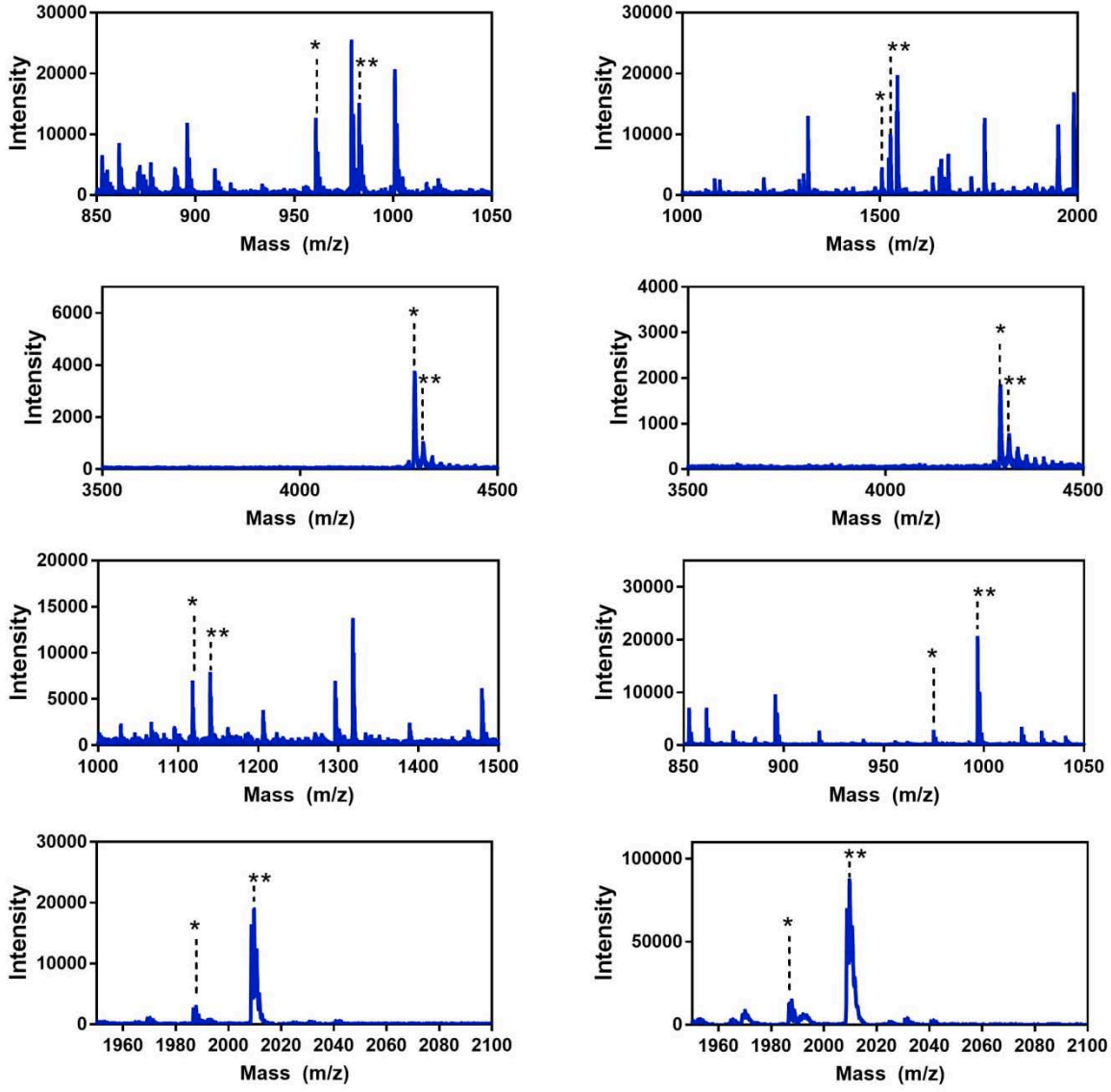

Figure S5: MALDI Spectra of Trypsin Digests of Cnf-Trp Constructs. Plots show Cnf containing fragments of αS-Cnf^{39} (top left), αS-Cnf^{94} (top right), αS-Cnf^{125} (upper middle left), αS-Cnf^{136} (upper middle right) and Trp containing fragments of αS-Trp^{24} (lower middle left), αS-Trp^{39} (lower middle right), αS-Trp^{62} (bottom left) and αS-Trp^{79} (bottom right). On each plot the M+H mass of the fragment of interest is indicated with * while M+Na masses are indicated with **.

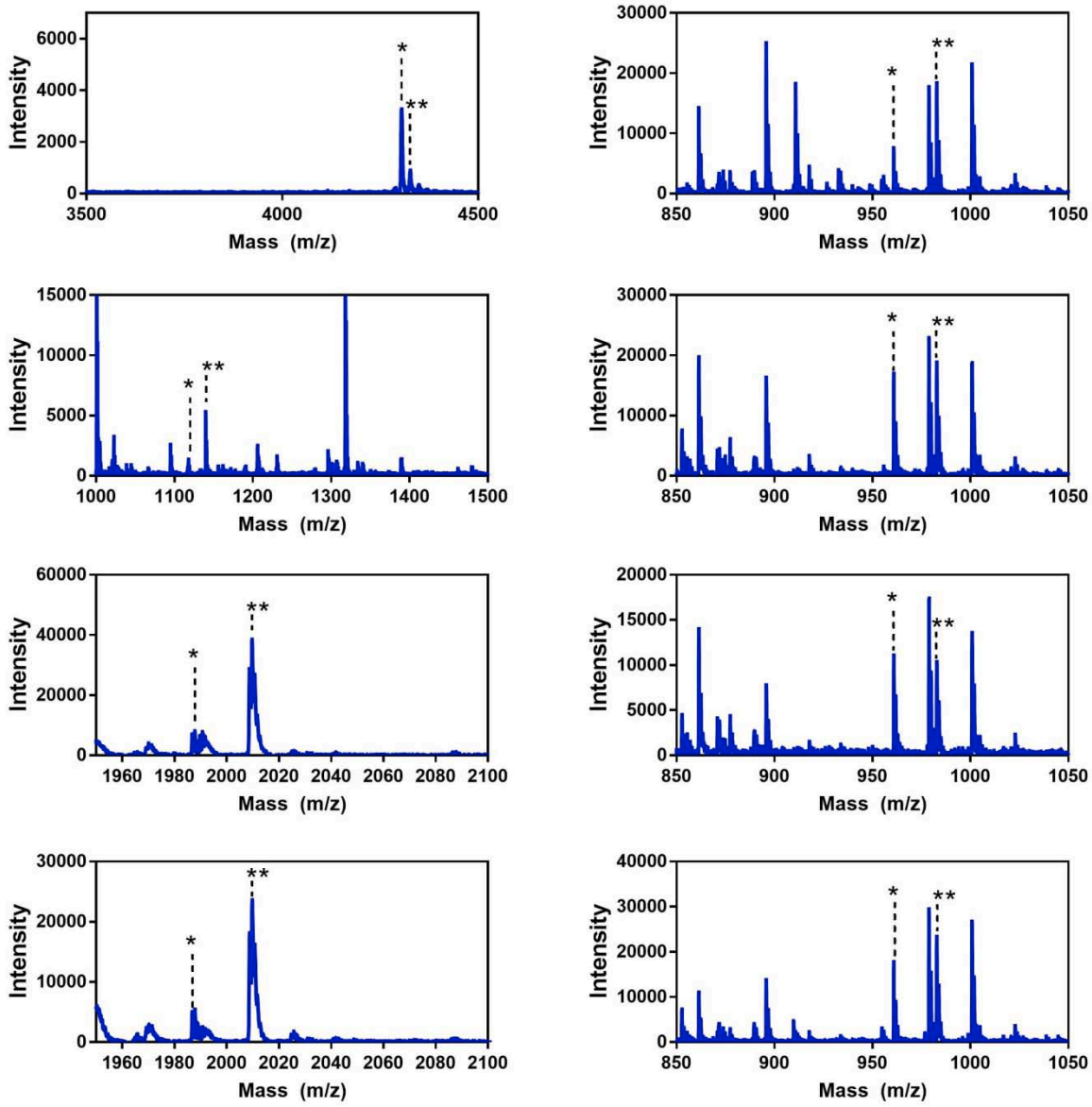

Figure S6: MALDI Spectra of Trypsin Digests of Cnf-Trp Constructs. Plots show Trp containing fragment of αS-Trp^{125} (top left), Cnf (top right) and Trp (upper middle left) containing fragments of αS-Cnf^{39}-Trp^{24}, Cnf (upper middle right) and Trp (lower middle left) containing fragments of αS-Cnf^{39}-Trp^{62}, Cnf (lower middle right) and Trp (bottom left) containing fragments of αS-Cnf^{39}-Trp^{79} and Cnf containing fragment of αS-Cnf^{39}-Trp^{94} (bottom right). On each plot the M+H mass of the fragment of interest is indicated with * while M+Na masses are indicated with **.

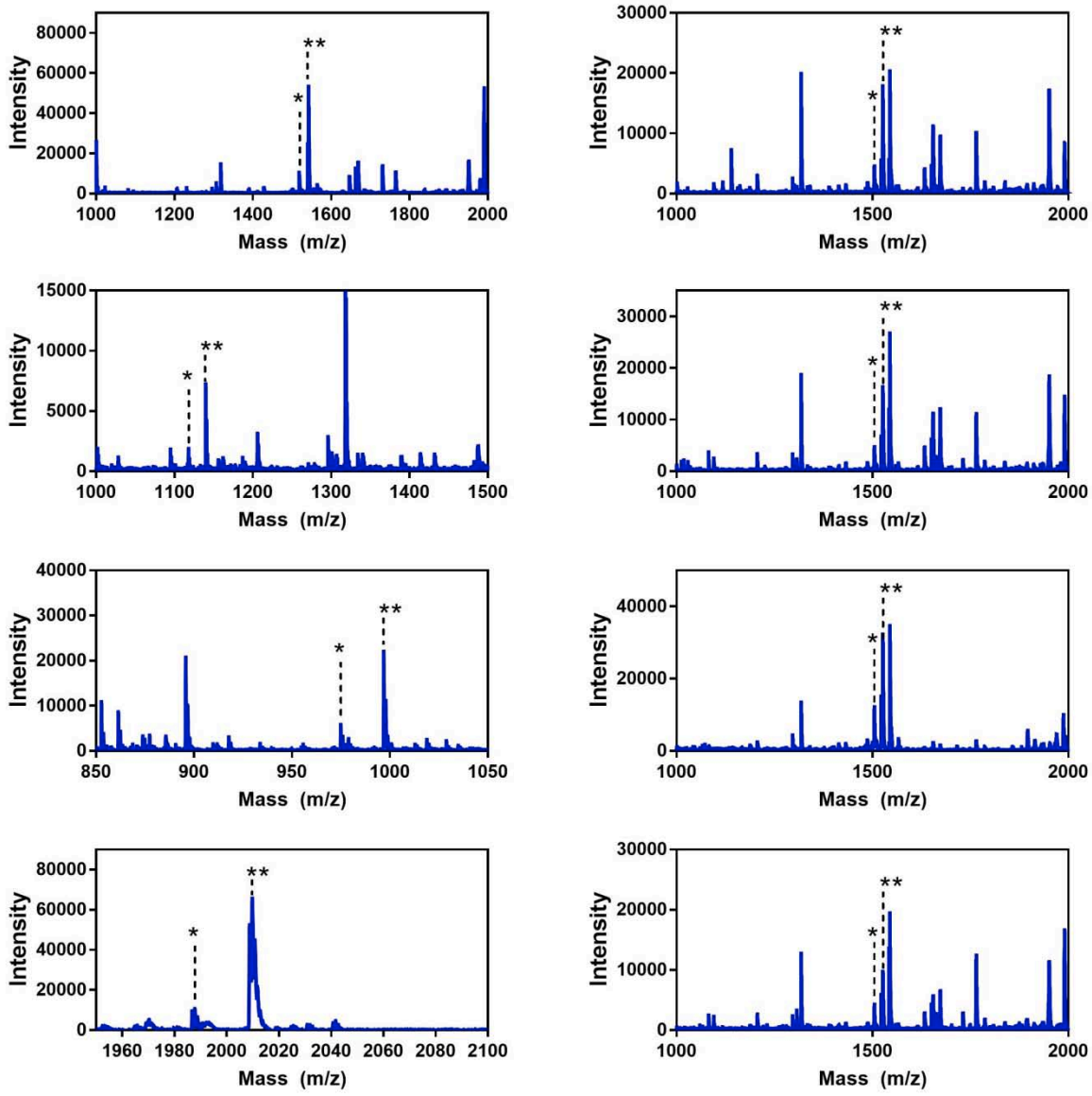

Figure S7: MALDI Spectra of Trypsin Digests of Cnf-Trp Constructs. Plots show Trp containing fragment of αS-Cnf^{39}-Trp^{94} (top left), Cnf (top right) and Trp (upper middle left) containing fragments of αS-Cnf^{94}-Trp^{24}, Cnf (upper middle right) and Trp (lower middle left) containing fragments of αS-Cnf^{94}-Trp^{39}, Cnf (lower middle right) and Trp (bottom left) containing fragments of αS-Cnf^{94}-Trp^{62} and Cnf containing fragment of αS-Cnf^{94}-Trp^{79} (bottom right). On each plot the M+H mass of the fragment of interest is indicated with * while M+Na masses are indicated with **.

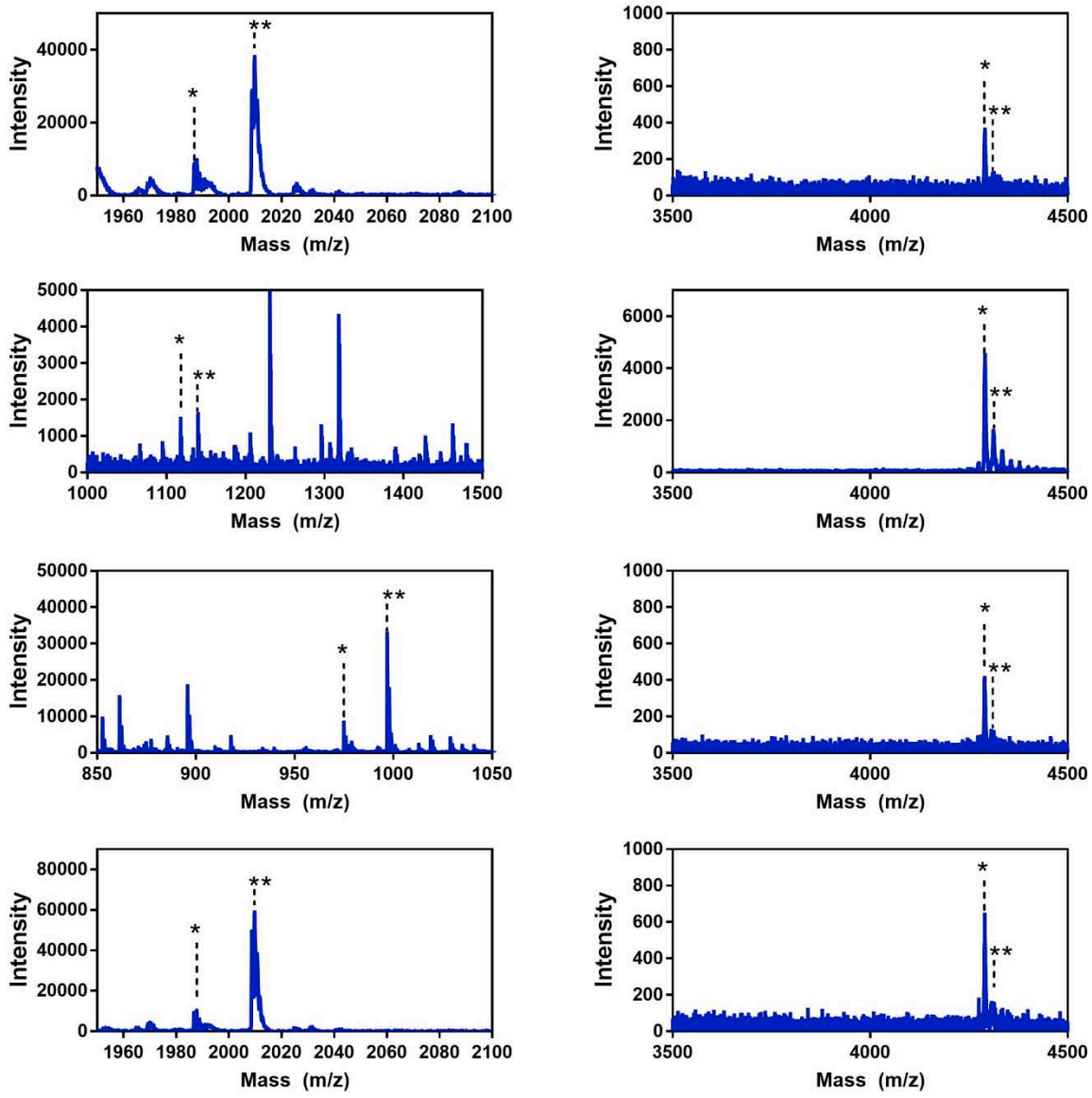

Figure S8: MALDI Spectra of Trypsin Digests of Cnf-Trp Constructs. Plots show Trp containing fragment of αS-Cnf^{94}-Trp^{79} (top left), Cnf (top right) and Trp (upper middle left) containing fragments of αS-Cnf^{125}-Trp^{24}, Cnf (upper middle right) and Trp (lower middle left) containing fragments of αS-Cnf^{125}-Trp^{39}, Cnf (lower middle right) and Trp (bottom left) containing fragments of αS-Cnf^{125}-Trp^{62} and Cnf containing fragment of αS-Cnf^{125}-Trp^{79} (bottom right). On each plot the M+H mass of the fragment of interest is indicated with * while M+Na masses are indicated with **.

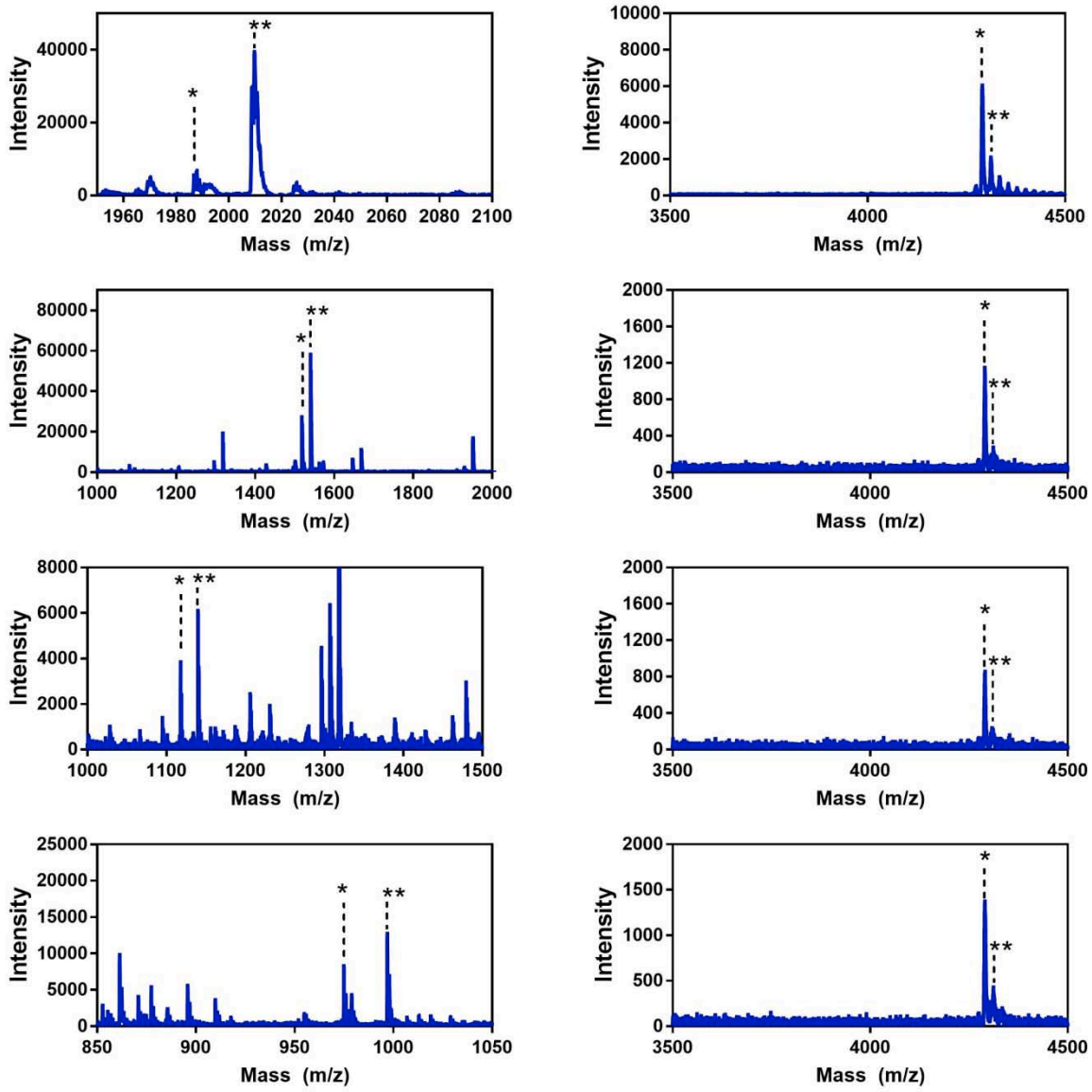

Figure S9: MALDI Spectra of Trypsin Digests of Cnf-Trp Constructs. Plots show Trp containing fragment of αS-Cnf^{125}-Trp^{79} (top left), Cnf (top right) and Trp (upper middle left) containing fragments of αS-Cnf^{125}-Trp^{94}, Cnf (upper middle right) and Trp (lower middle left) containing fragments of αS-Cnf^{136}-Trp^{24}, Cnf (lower middle right) and Trp (bottom left) containing fragments of αS-Cnf^{136}-Trp^{39} and Cnf containing fragment of αS-Cnf^{136}-Trp^{94} (bottom right). On each plot the M+H mass of the fragment of interest is indicated with * while M+Na masses are indicated with **.

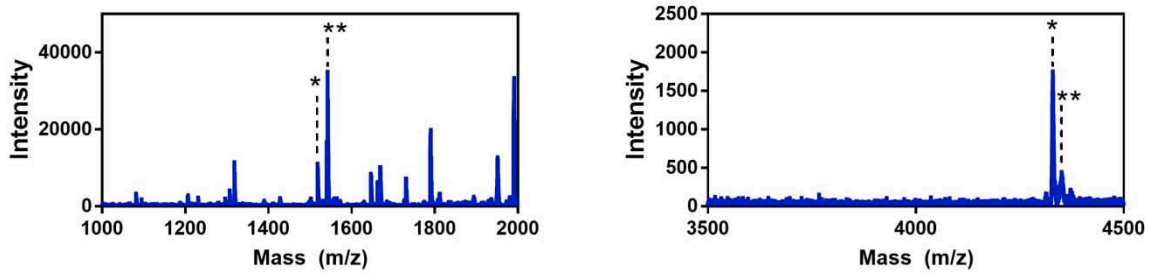

Figure S10: MALDI Spectra of Trypsin Digests of Cnf-Trp Constructs. Plots show Trp containing fragment of αS-Cnf^{136}-Trp^{94} (left) and Cnf and Trp containing fragment of αS-Cnf^{125}-Trp^{136} (top right). On each plot the M+H mass of the fragment of interest is indicated with * while M+Na masses are indicated with **.

Table S2. MALDI Masses from Fam-Raz Library

	Full Length		Tyrpsin Digest				
Construct	Exp. [M+H]+	Obs. [M+H]+	Fragment	Exp. [M+H]+	Obs. [M+H]+	Exp. [M+Na]+	Obs. [M+Na]+
αS-Raz^{94}	15027	15027.8	81-96	2044.0	2046.6	2066.0	2068.6
αS-Raz^{136}	15011	15013.3	103-140	4839.0	4837.4	4861.0	4859.4
αS-Fam^{9}-Raz^{136}	15474	15475.1	7-10	865.3	865.9	887.3	887.9
			103-140	4839.0	4837.1	4861.0	4859.1
αS-Fam^{9}-Raz^{94}	15488	15488.1	7-10	865.3	865.6	887.3	887.6
			81-96	2044.0	2044.6	2066.0	2066.6
αS-Fam^{24}-Raz^{94}	15447	15450.7	24-32	1250.5	1250.3	1272.4	1272.3
			81-96	2044.0	2044.8	2066.0	2066.8
αS-Fam^{42}-Raz^{94}	15488	15490.1	35-43	1412.6	1413.7	1434.6	1435.7
			81-96	2044.0	2045.3	2066.0	2067.3
αS-Fam^{62}-Raz^{94}	15447	15448.5	61-80	2348.1	-	2370.1	2371.7
			81-96	2044.0	2046.4	2066.0	2068.4
αS-Fam^{87}-Raz^{94}	15488	15488.4	81-96	2505.1	2506.6	2527.0	2528.6
αS-Fam^{114}-Raz^{94}	15446	15446.5	103-140	4704.8	-	4726.8	4728.8
			81-96	2044.0	2046.4	2066.0	2068.4
αS-Fam^{123}-Raz^{94}	15446	15446.1	103-140	4704.8	-	4726.8	4728.9
			81-96	2044.0	2046.4	2066.0	2068.4
αS-Fam^{136}-Raz^{94}	15410	15413.8	103-140	4668.8	-	4690.7	4694.8
			81-96	2044.0	2046.4	2066.0	2068.4

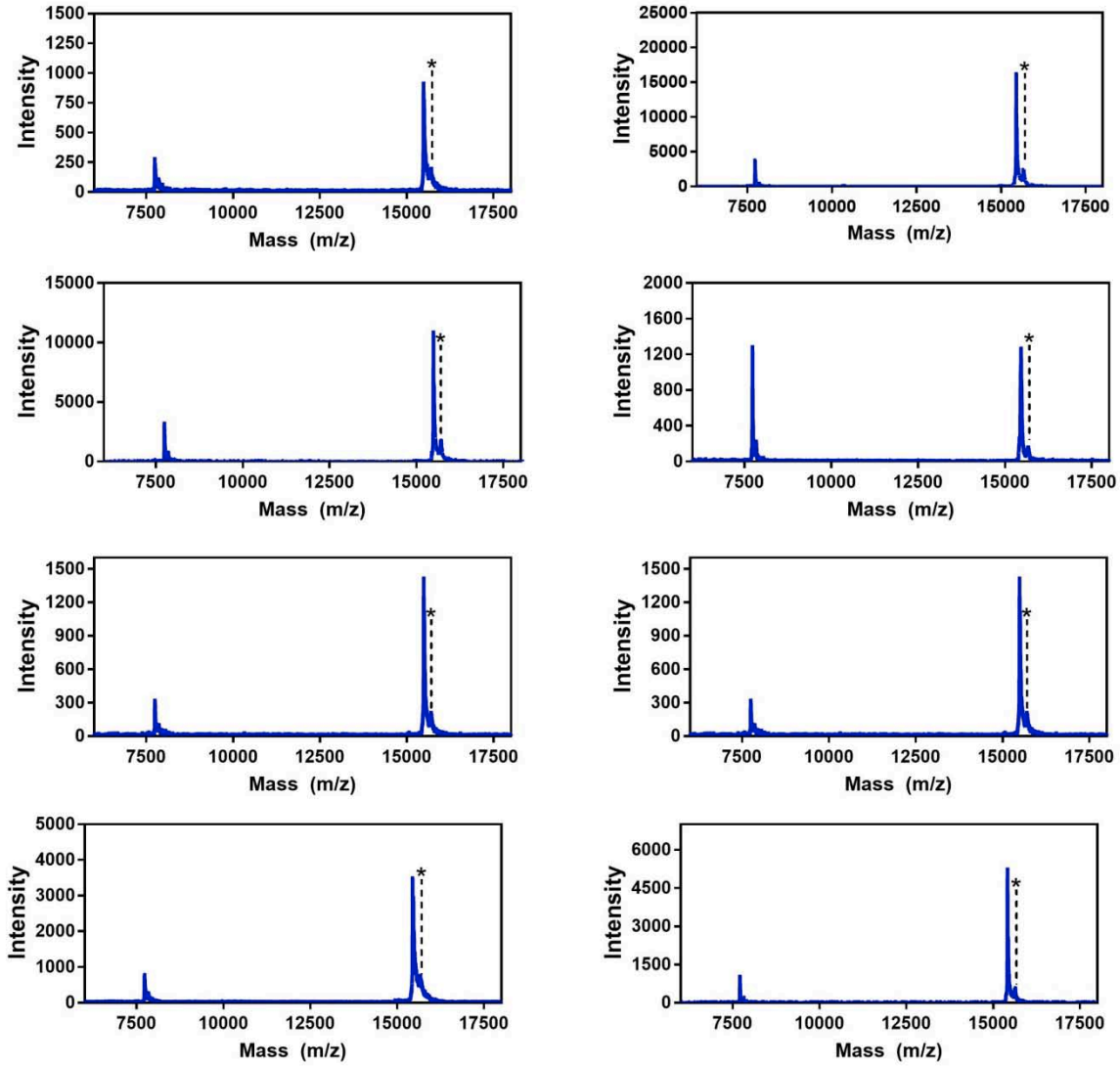

Figure S11: MALDI Spectra of Full-Length Fam-Raz Constructs. Plots show αS-Fam^{9}-Raz^{94} (top left), αS-Fam^{24}-Raz^{94} (top right), αS-Fam^{42}-Raz^{94} (upper middle left), αS-Fam^{62}-Raz^{94} (upper middle right), αS-Fam^{87}-Raz^{94} (lower middle left), αS-Fam^{114}-Raz^{94} (lower middle right), αS-Fam^{123}-Raz^{94} (bottom left) and αS-Fam^{136}-Raz^{94} (bottom right). On each plot the matrix adduct is noted with *.

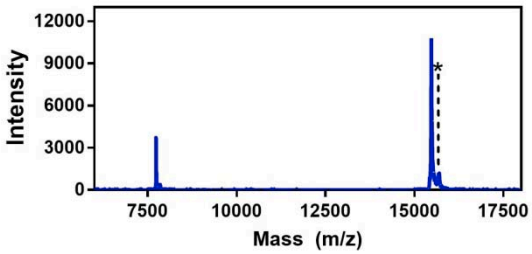

Figure S12: MALDI Spectra of Full-Length Fam-Raz Constructs. Plot shows αS-Fam^{9}-Raz^{136} with the matrix adduct is noted with *.

All single Fam and Raz labeled constructs have been previously reported (3, 4).

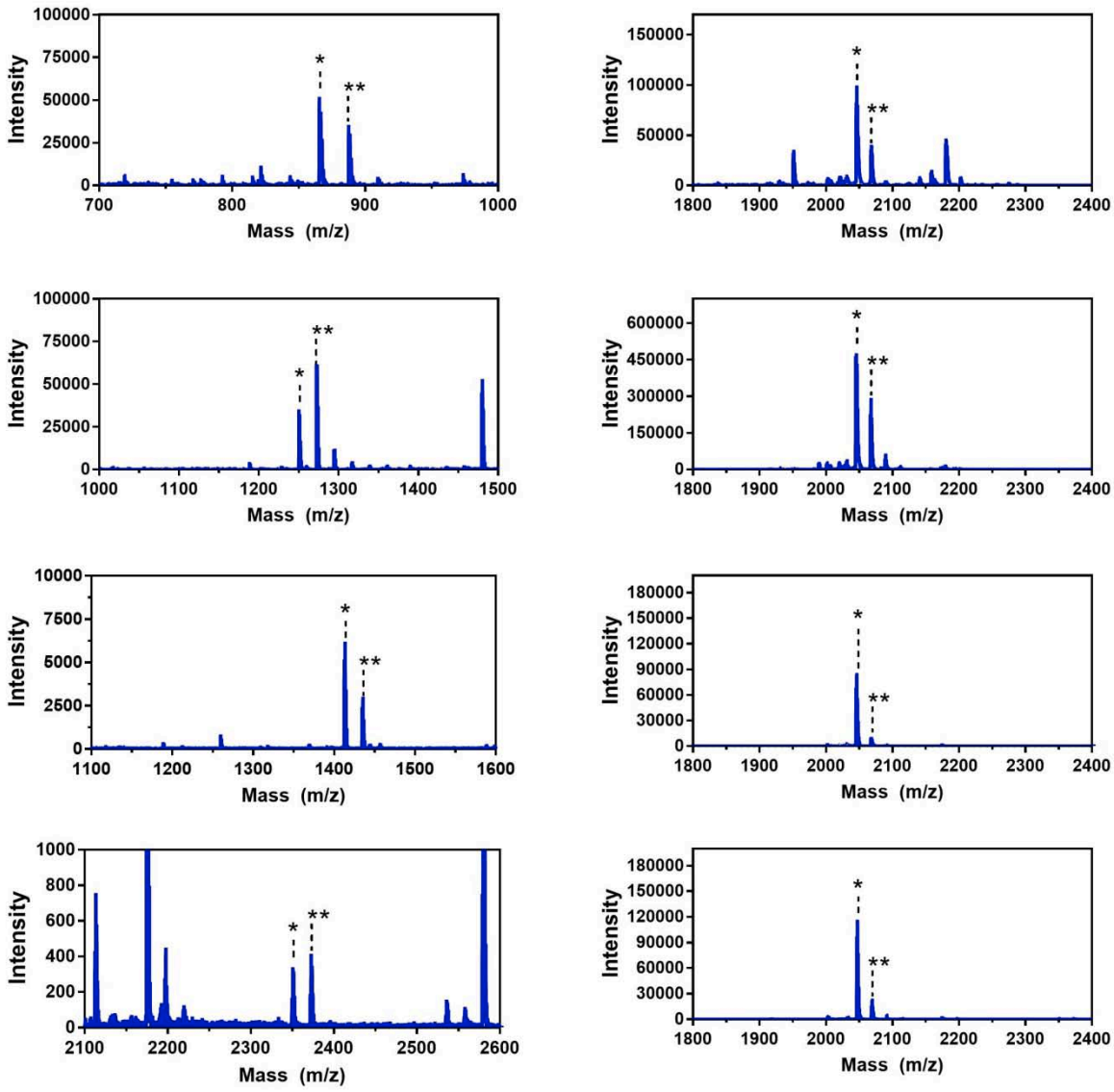

Figure S13: MALDI Spectra of Tryspin Digests of Fam-Raz Constructs. Plots show Fam (top left) and Raz (top right) containing fragments of αS-Fam^{9}-Raz^{94}, Fam (upper middle left) and Raz (upper middle right) containing fragments of αS-Fam^{24}-Raz^{94}, Fam (lower middle left) and Raz (lower middle right) containing fragments of αS-Fam^{42}-Raz^{94} and Fam (bottom left) and Raz (bottom right) containing fragments of αS-Fam^{62}-Raz^{94}. On each plot the M+H mass of the fragment of interest is indicated with * while M+Na masses are indicated with **.

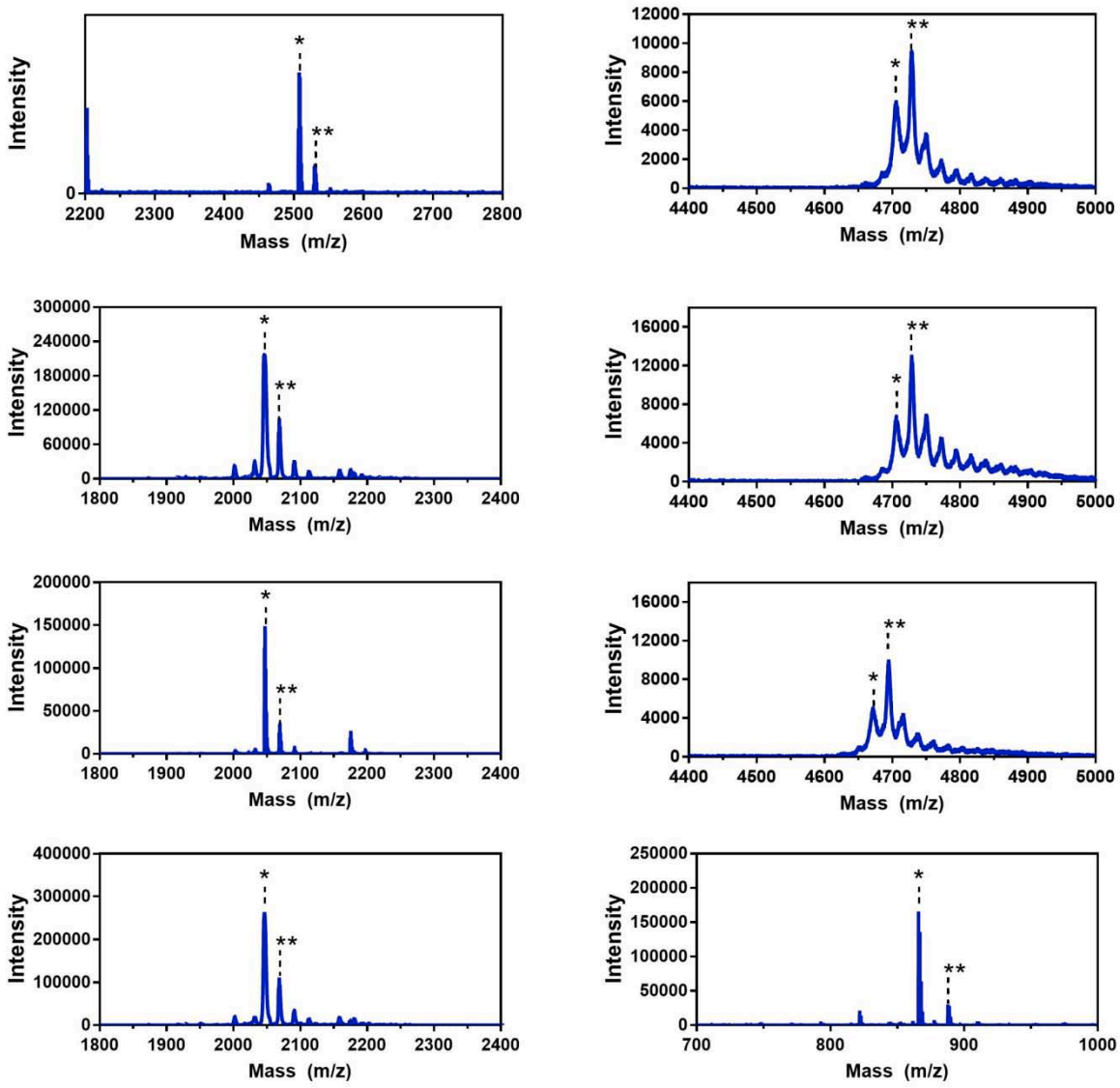

Figure S14: MALDI Spectra of Tryspin Digests of Fam-Raz Constructs. Plots show Fam and Raz containing fragement of αS-Fam^{87}-Raz^{94} (top left), Fam (top right) and Raz (upper middle left) containing fragments of αS-Fam^{114}-Raz^{94}, Fam (upper middle rightt) and Raz (lower middle left) containing fragments of αS-Fam^{123}-Raz^{94}, Fam (lower middle right) and Raz (bottom left) containing fragments of αS-Fam^{136}-Raz^{94} and Fam containing fragment of αS-Fam^{9}-Raz^{136} (bottom right). On each plot the M+H mass of the fragment of interest is indicated with * while M+Na masses are indicated with **.

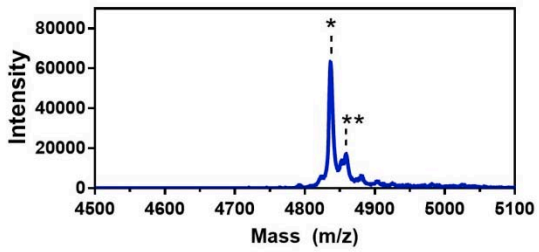

Figure S15: MALDI Spectra of Tryspin Digest of Fam-Raz Construct. Plots show Raz containing fragement of αS-Fam^{9}-Raz^{136}. The M+H mass of the fragment of interest is indicated with * while M+Na mass is indicated with **.

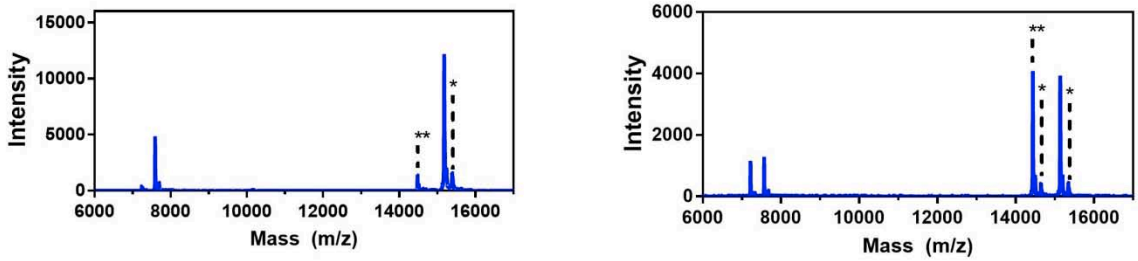

Figure S16: MALDI Spectra of Full-Length AF488 Constructs. Plots show αS-AF488^{9} (left) and αS-AF488^{114} (right). On each plot the matrix adduct is indicated with * while unlabeled Cys protein is indicated with **.

The αS-AF488^{130} construct has been previously reported (6).

FRET Data Collection and Analysis

TMAO Assay: All labeled αS variants were dialyzed into 20 mM tris, 100 mM NaCl, pH 7.5. Buffers containing TMAO were also prepared with 20 mM tris, 100 mM NaCl, and the pH was adjusted to 7.5 following the addition of TMAO. Concentrations for the Cnf-Trp library was determined using the Sigma-Aldrich FluoroProfile Quantification Kit, while concentrations for Fam-Raz constructs were determined via UV-Vis absorbance. Steady-state measurements for the Cnf-Trp library and time correlated single photon counting (TCSPC) measurements of the Fam-Raz library were performed at 1 μM, while steady-state measurements for the Fam-Raz library were performed at 100 nm. Prior to each measurement, labeled protein was mixed with TMAO containing buffer and briefly vortexed. Measurements were taken in triplicate at final concentrations of 0, 2 and 4 M TMAO at 20 °C. Steady-state measurements for the Cnf-Trp library were collected with an excitation of 240 nm over an emission range of 275 - 410 nm with excitation and emission slits set to 5 nm and 1 nm step size collecting for 0.75 seconds per step, exciting primary the Cnf fluorophore. Spectra were collected for direct excitation of Trp with an excitation of 280 nm over an emission range of 310 - 410 nm with excitation and emission slits set to 5 nm and 1 nm step size collecting for 0.75 seconds per step. Measurements of the Fam-Raz library were collected with an excitation of 486 nm over an emission range of 495 - 700 nm with excitation and emission slits set to 5 nm and 1 nm step size collecting for 0.25 seconds per step. Direct excitation of Raz was performed by exciting at 555 and measuring over an emission range of 565 - 700 nm with a 1 nm step size and a collection time of 0.25 seconds per step.

All TCSPC measurements of fluorescence lifetime decays were collected using a pulsed LED with a maximum emission at 486 nm. Fluorescence was collected at 515 nm with the slit widths adjusted for each measurement to keep the ACD value between 1 - 3 % of the SYNC value. The instrument response function (IRF) was collected for each slit width used for collection. For these experiments labeled αS was mixed with TMAO containing buffer to final protein concentrations of 2 μM and 0, 2 or 4 M TMAO. Additionally, formation of aggregates in buffer or TMAO were assayed by mixing single-labeled αS containing FAM and single-labeled αS containing Raz in equimolar concentrations with a final protein concentration of 2 μM in 0, 2 and 4 M TMAO.

Fitting Steady-State Data: Following data collection, the single labeled spectra were used to quantify the degree of energy transfer. First, the spectral overlap of the donor and acceptor were deconvoluted by fitting the double labeled spectrum with the linear sum of the individual donor and acceptor-labeled spectra. Fitting was performed by minimizing the total least squared difference using the Excel Solver feature to adjust the constants *A* and *B*:

$$\sum_{\lambda}\left(I(\lambda)_{DA} - AI(\lambda)_{D} - BI(\lambda)_{A}\right)^{2} \rightarrow \min \qquad \text{(Eq. 1)}$$

$$I(\lambda)_{DA} = AI(\lambda)_{D} + BI(\lambda)_{A} \qquad \text{(Eq. S1)}$$

Here, $I(\lambda)_{DA}$, $I(\lambda)_{D}$ and $I(\lambda)_{A}$ are the wavelength dependent fluorescence intensities of the double-labeled, single labeled protein containing the donor, and single-labeled protein containing the acceptor fluorophore, respectively. Solutions to Eq. S1 were obtained by utilizing the Excel Solver functionality. This procedure was performed at each TMAO concentration, thus accounting for any changes in quantum yield or spectral shifting. The linear contributions of the single-labeled construct containing the donor only, *A*, and the contribution from the single-labeled construct containing the acceptor only, *B*, were both used to independently calculate the E_{FRET} through Eq. S2 and S3 and combined in a weighted average via Eq. S4:

$$E_{D} = (1 - A) \qquad \text{(Eq. S2)}$$

$$E_{A} = (B - 1)\frac{\varepsilon_{A}}{\varepsilon_{D}} \qquad \text{(Eq. S3)}$$

$$E_{FRET} = \left(\frac{1}{S_{D}} + \frac{1}{S_{A}}\right)^{-1}\left(\frac{E_{D}}{S_{D}} + \frac{E_{A}}{S_{A}}\right) \qquad \text{(Eq. S4)}$$

In Eqs. S2-S4, E_D and E_A are E_{FRET} values calculated from the donor and acceptor weights respectively. Additionally, Eq. S3 requires ratio of the extinction coefficient for the acceptor, ε_A, to the donor, ε_D, which are detailed in Tables S3a and S3b and have been determined from the absorption spectrum of each fluorophores scaled using published extinction coefficients (Cnf ε_{240}=13,921 $M^{-1}cm^{-1}$ (7), Trp ε_{278}=5700 $M^{-1}cm^{-1}$ (7), FAM ε_{494}=68,000 $M^{-1}cm^{-1}$ (8) and Raz ε_{555}=87,000 $M^{-1}cm^{-1}$ as indicated by the manufacturer) . The E_{FRET} values from the donor and acceptor were then used to compute a weighted average E_{FRET} value by using the inverse of the

experimental error to weight each E_{FRET} value, where S_D and S_A represent the donor and acceptor propagated error respectively.

$$S_A = A \times \sqrt{\left(S_{I_A} \times B / I_D\right)^2 + \left(S_{I_{DL}} / I_D\right)^2 + \left(S_{I_D} \times \left(I_{DL} - B \times I_A\right) / I_D^2\right)^2} \qquad \text{(Eq. S5)}$$

$$S_B = B \times \sqrt{\left(S_{I_D} \times S / I_A\right)^2 + \left(S_{I_{DL}} / I_A\right)^2 + \left(S_{I_A} \times \left(I_{DL} - A \times I_D\right) / I_A^2\right)^2} \qquad \text{(Eq. S6)}$$

$$S_{EFRET} = \sqrt{2} \Big/ \left(\frac{1}{S_A} + \frac{1}{S_B}\right) \qquad \text{(Eq. S7)}$$

Lastly, S_{EFRET} represents the propagated error of the calculated weighted average E_{FRET} value.

Accurate calculation of R_0 was required for calculation of interresidue distances from the determined E_{FRET}. R_0 was calculated using the equation:

$$R_0^6 = \left(9 \ln(10) \kappa^2 \Phi_D J\right) / \left(128 \pi^2 N_A n^4\right) \qquad \text{(Eq. S8)}$$

Here, N_A is Avogadro's number, κ^2 is the dipole orientation factor, approximated at 2/3, Φ_D is the quantum yield of the donor, J is the spectral overlap integral between the emission of the donor and the absorbance of the acceptor and n is the refractive index of the medium. The overlap integral of the donor fluorescence and acceptor absorbance spectra for each fluorophore pair was determined empirically from the absorbance and emission spectra of the free fluorophores in buffer through application of the integral:

$$J = \int_0^\infty f_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda \qquad \text{(Eq. S9)}$$

where $f_D(\lambda)$ is the normalized donor emission, $\varepsilon_A(\lambda)$ is the molar extinction coefficient of the acceptor, at each wavelength (λ). The normalized donor emission is given by:

$$f_D(\lambda) = \frac{F_D(\lambda)}{\int_0^\infty F_D(\lambda) d\lambda} \qquad \text{(Eq. S10)}$$

where $F_D(\lambda)$ is the fluorescence emission spectrum of the donor dye. Following determination of R_0 average distance values from the polymer scaled Förster equation (Eq. 2), using either Eq 3 or

Eq 4 to determine $P_n(r,x)$), along with the associated error and constraint weights for each constraining function, were determined in Wolfram Mathematica.

$$E_{FRET} = \sum_r P_n(r,x) / \left(1 + (r/R_0)^6\right) \quad \text{(Eq. 2)}$$

$$P_1(r,x) = 4\pi r \left(\frac{3}{2\pi x^2}\right)^{3/2} \exp\left(-\frac{3}{2}\frac{r^2}{x^2}\right) \quad \text{(Eq. 3)}$$

$$P_2(r,x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(r-x)^2}{\sigma^2}\right) \quad \text{(Eq. 4)}$$

Fluorescence measurements of labeled proteins have shown that the quantum yield is sensitive to the local environment, as well as TMAO concentration. Therefore, changes in quantum yield must also be taken into account to effectively determine R_0 via Eq. S8. The quantum yield was calculated by fitting the emission spectrum of the free fluorophore in buffer, without TMAO, to the emission spectrum of each labeled analog at each concentration of TMAO, again using a linear least squared difference approach.

$$\sum_\lambda \left(I_D - C \times I_{Dye}\right)^2 \quad \text{(Eq. S11)}$$

The sum in Eq. S11 was minimized using the Excel solver function by adjusting the constant, C. Here, I_D and I_{Dye} represent the sum of the fluorescence intensity over all wavelengths of the labeled protein and the free fluorophore, respectively. The donor quantum yield is then defined as:

$$\Phi_D = C \times \Phi_0 \quad \text{(Eq. S12)}$$

where Φ_0 is the quantum yield of the free fluorophore. By using this empirically adjusted quantum yield in the calculation of R_0, we effectively reduce inaccuracies in the determined interchromophore distance arising from changes in quantum yield. Lastly, the distance of interest, R, is determined from the above variables as described in the main text.

Finally, error was propagated through the calculation of the interprobe distance. This was performed by determining the inverse function of $P_n(r,x)$, here represented as $F(E_{FRET})$, using Wolfram Mathematica as shown in Eq. S13.

$$F\left(E_{FRET}\right) = P_n^{-1}\left(r, x\right) \qquad \text{(Eq. S13)}$$

The error is then simply propagated following the determination of the partial derivative of $P_n^{-1}(E_{FRET})$ with respect to the E_{FRET} variable as shown in Eq. S14.

$$S_{Distance} = S_{EFRET} \times \frac{\delta F\left(E_{FRET}\right)}{\delta E_{FRET}} \qquad \text{(Eq. S14)}$$

Fitting Lifetime Data: Lifetime data were fit using PowerFit10 distributed by PTI. Each decay was fit to a single or double exponential decay where the time regime was selected to minimize the chi-squared values and the residuals. E_{FRET} was determined from $1-(\tau_{DA}/\tau_D)$ where τ_{DA} and τ_D are the lifetimes for double-labeled and donor-only constructs respectively. For biexponential decays both the amplitude average and the intensity average lifetimes were used to calculate E_{FRET} values for comparison to the E_{FRET} values extracted from steady-state measurement.

Steady-state Fluorescence Spectra and Fitting

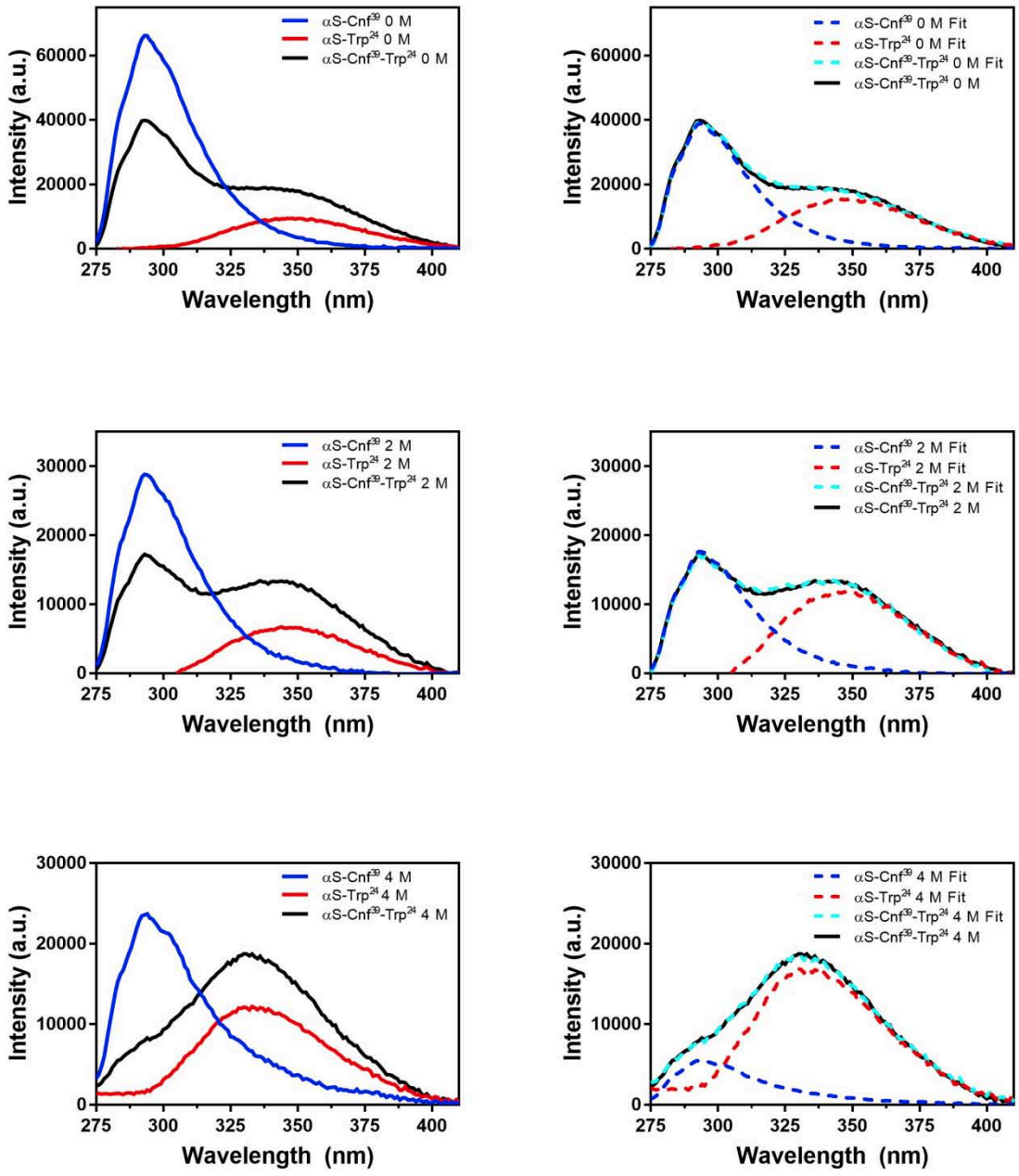

Figure S17: Steady-state Fluorescence Spectra of αS-Cnf^{39}-Trp^{24}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

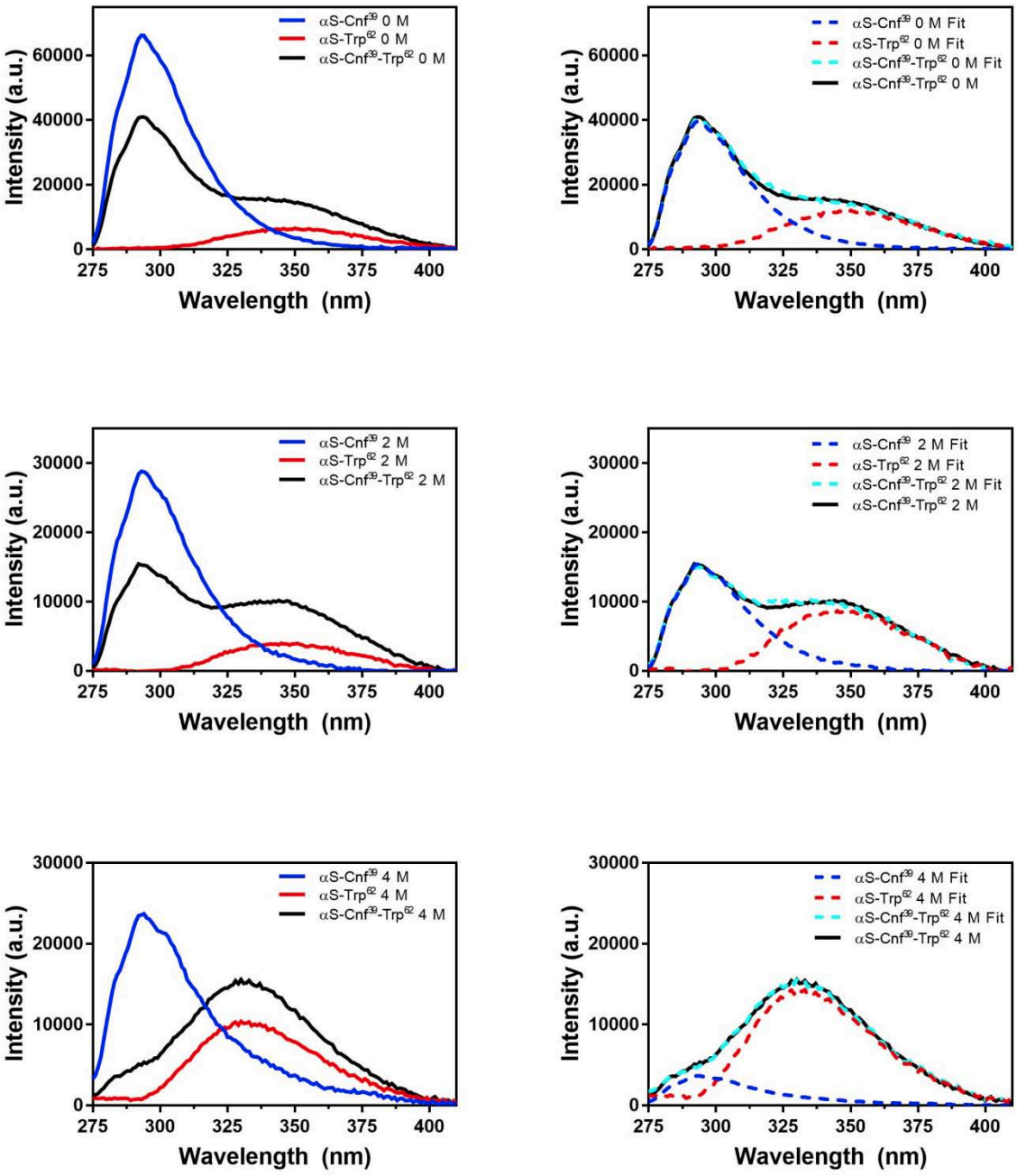

Figure S18: Steady-state Fluorescence Spectra of αS-Cnf^{39}-Trp^{62}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

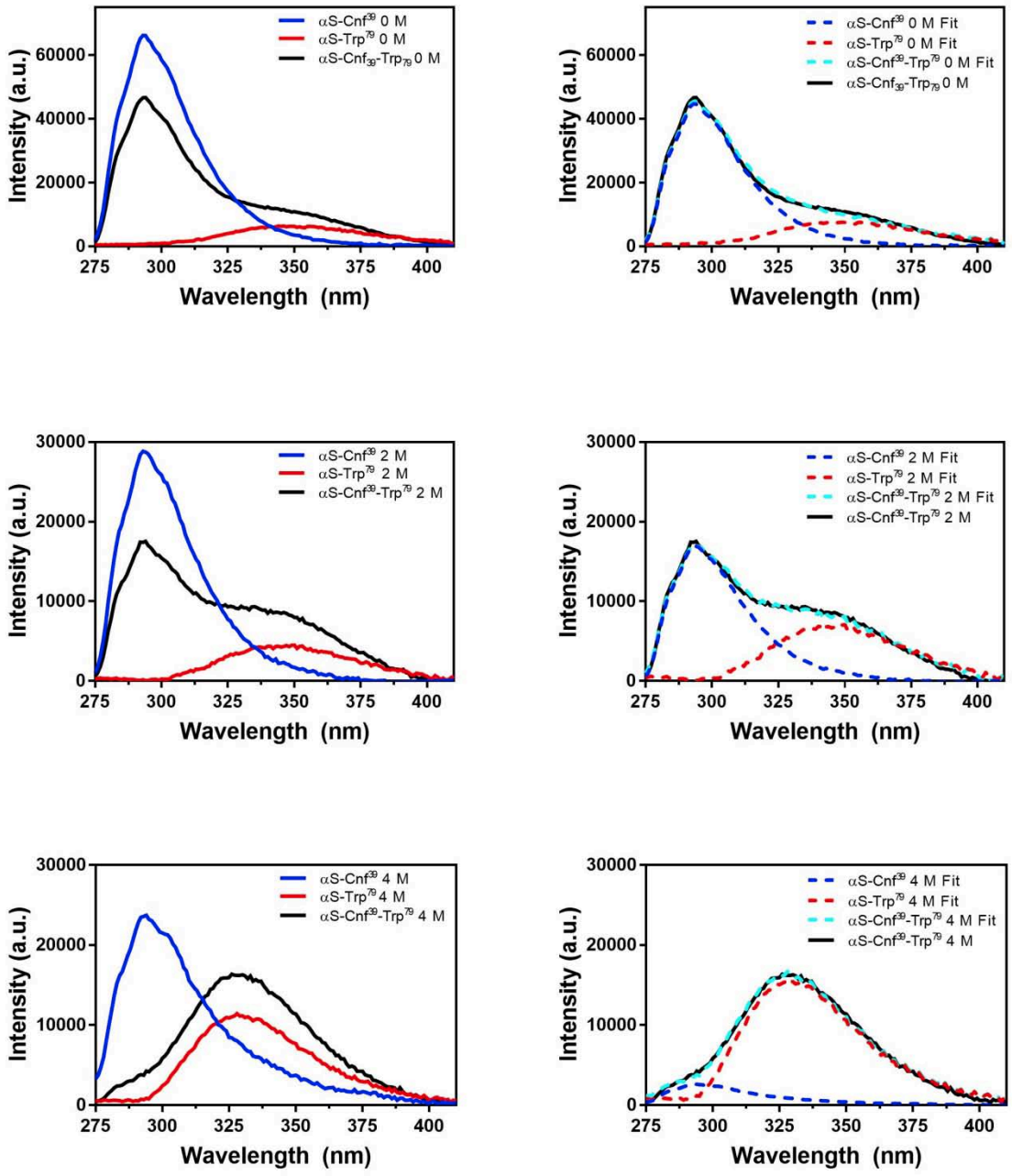

Figure S19: Steady-state Fluorescence Spectra of αS-Cnf^{39}-Trp^{79}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

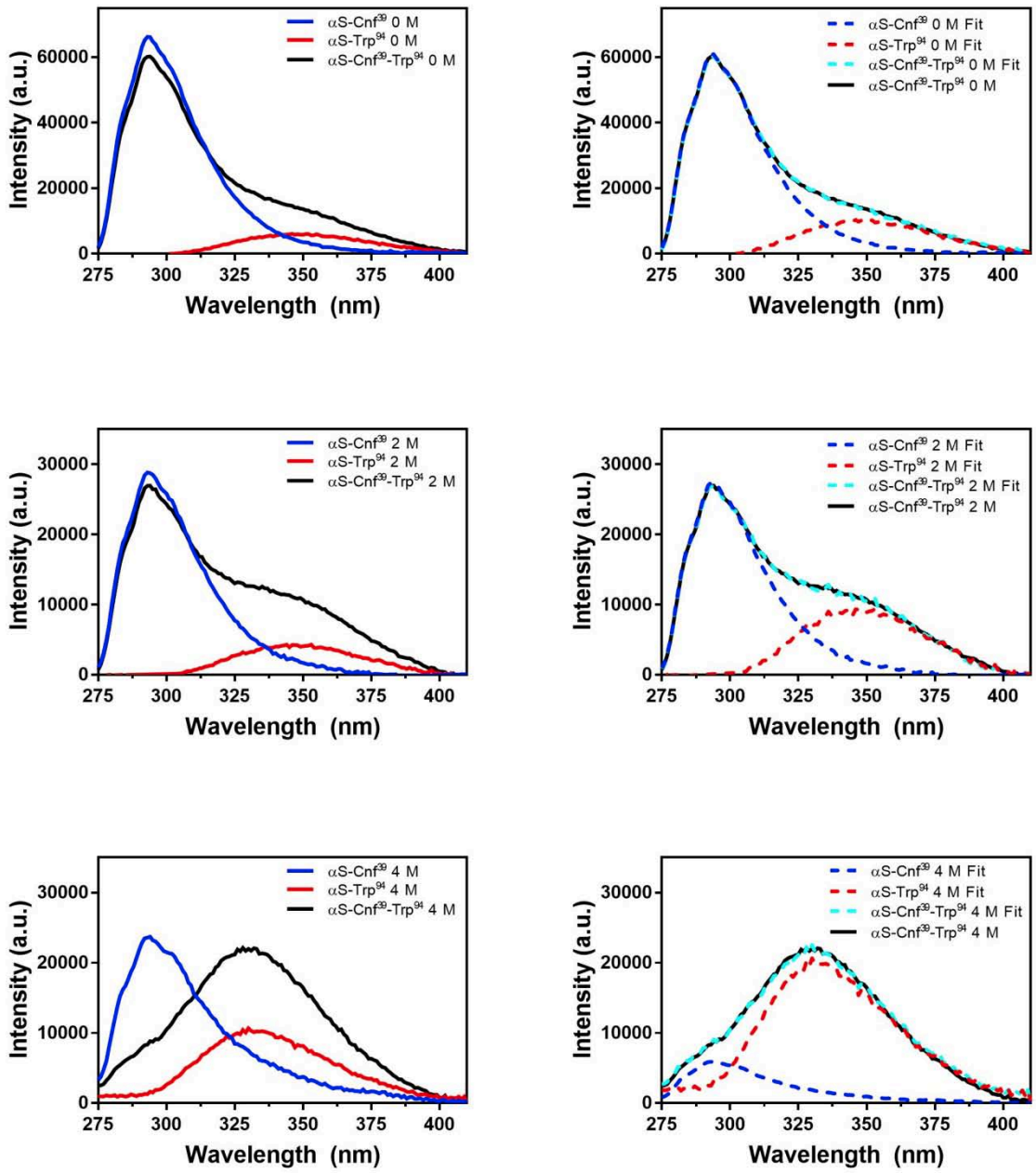

Figure S20: Steady-state Fluorescence Spectra of αS-Cnf^{39}-Trp^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

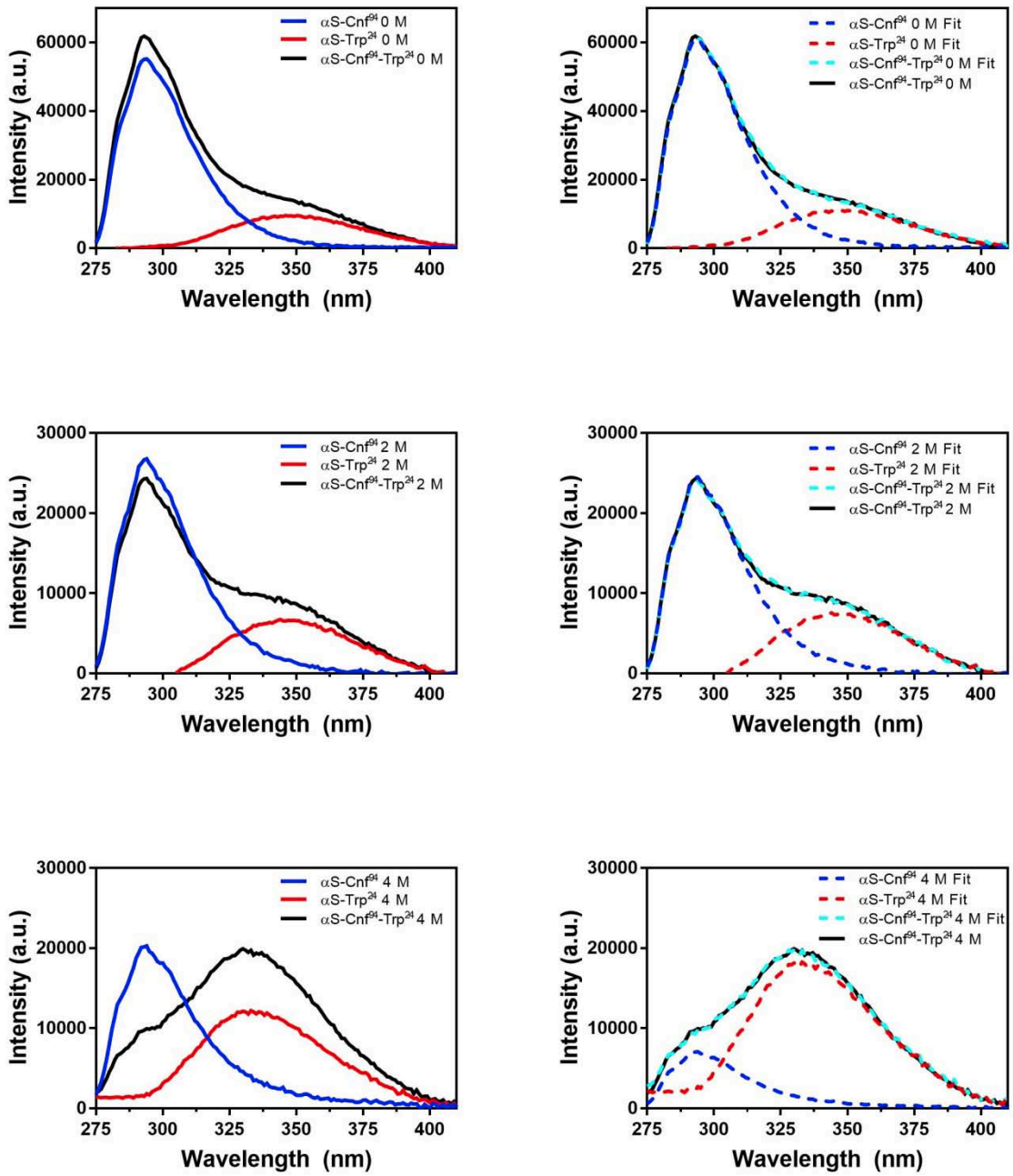

Figure S21: Steady-state Fluorescence Spectra of αS-Cnf^{94}-Trp^{24}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

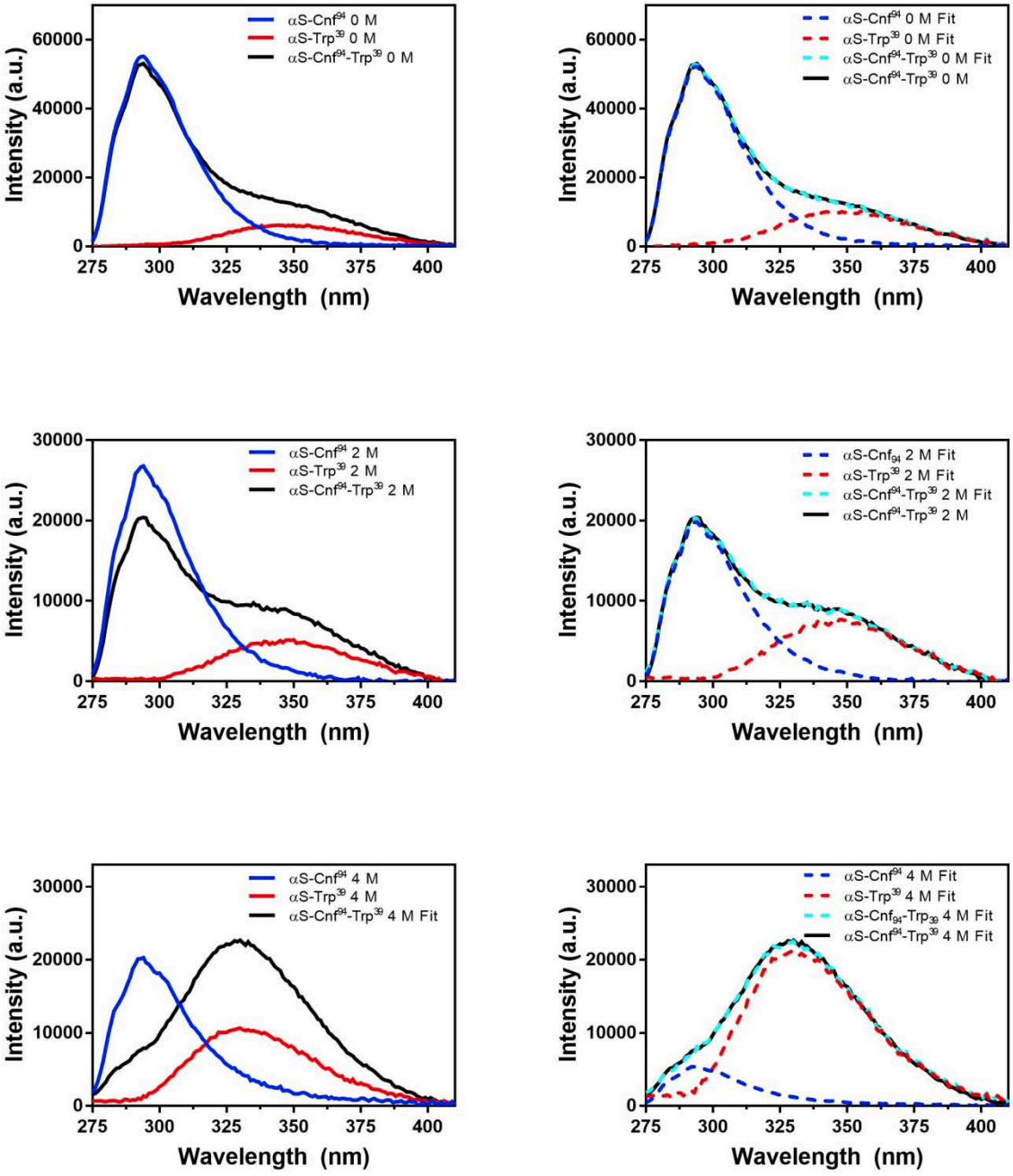

Figure S22: Steady-state Fluorescence Spectra of αS-Cnf^{94}-Trp^{39}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

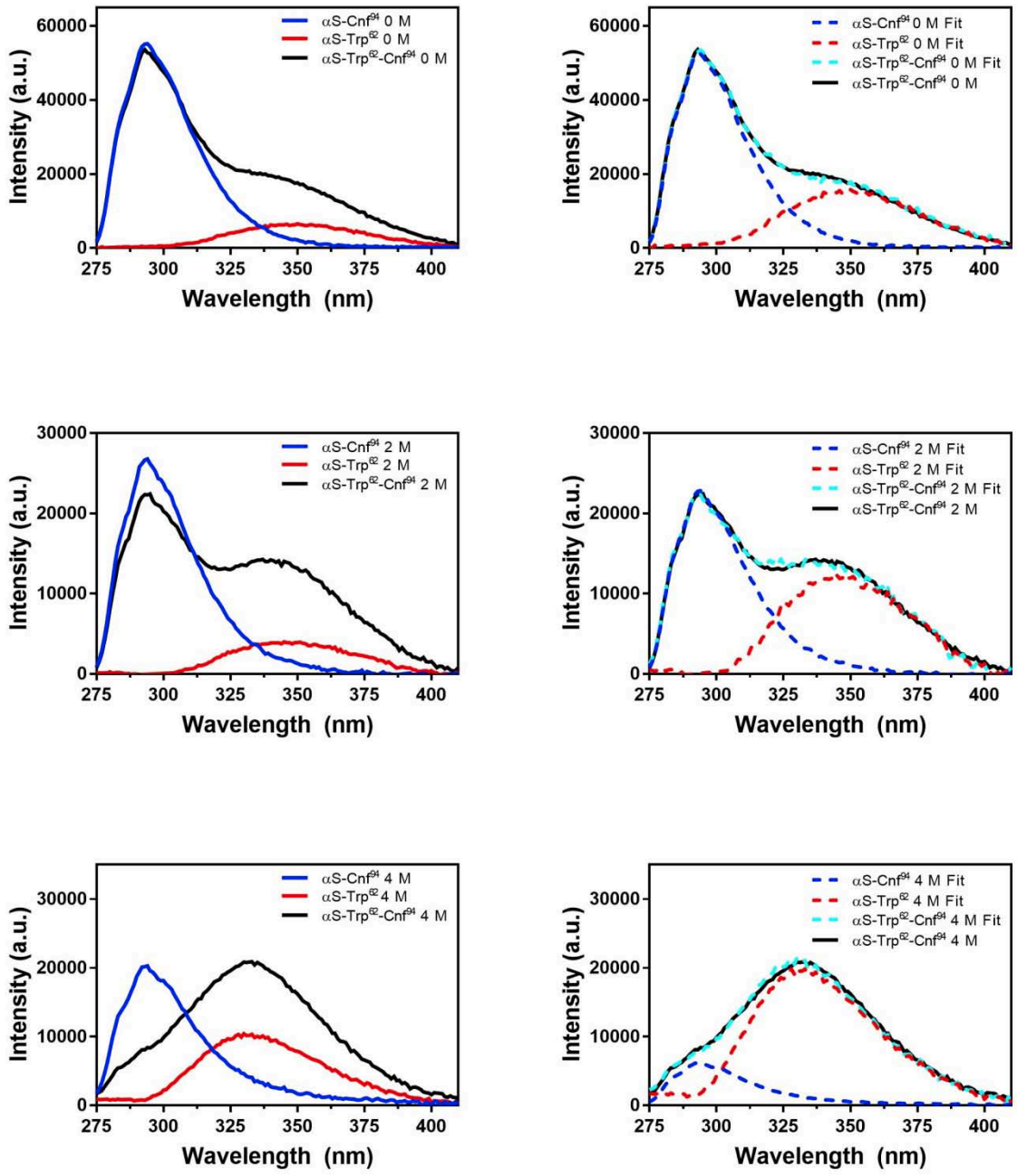

Figure S23: Steady-state Fluorescence Spectra of αS-Cnf^{94}-Trp^{62}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

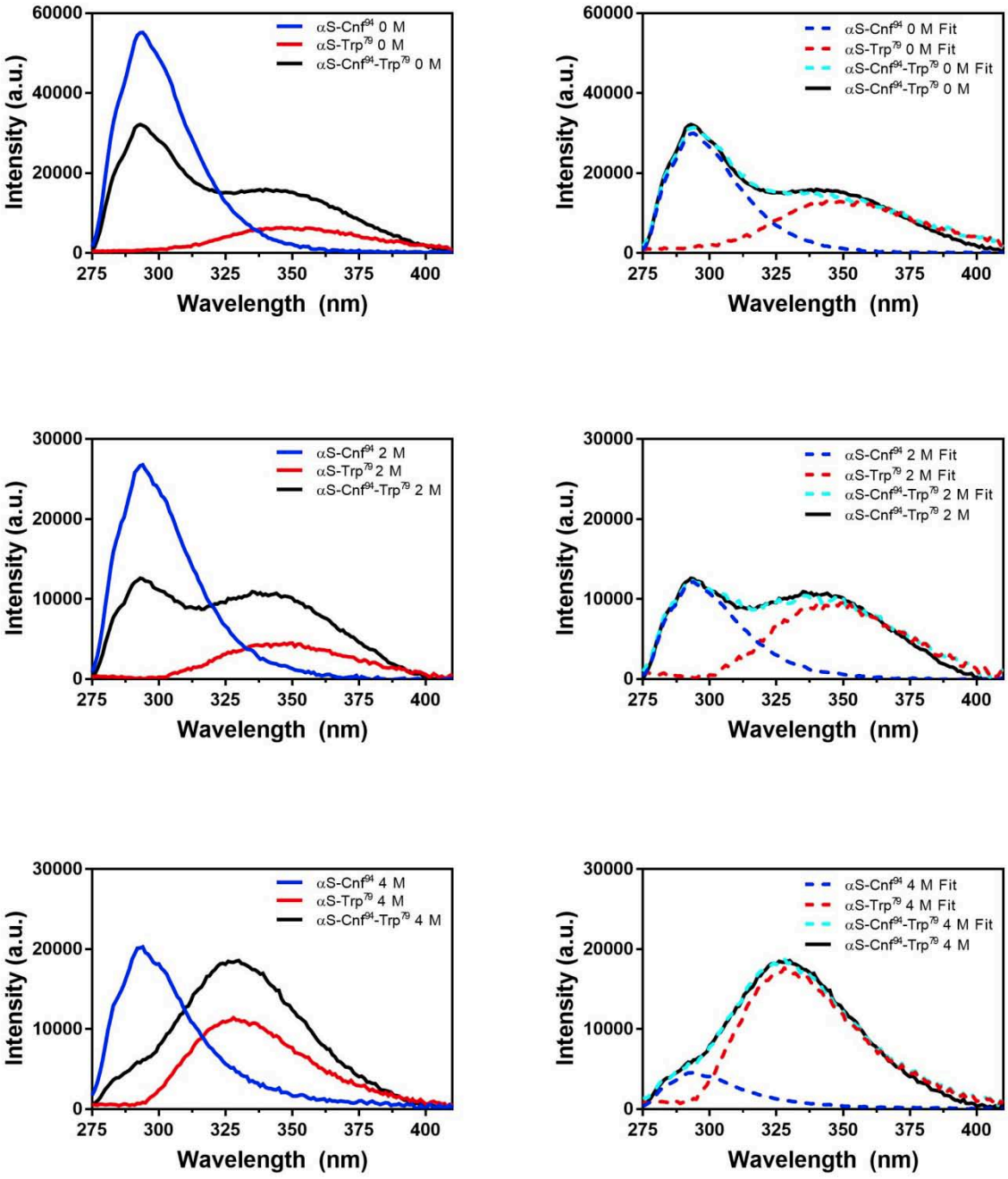

Figure S24: Steady-state Fluorescence Spectra of αS-Cnf^{94}-Trp^{79}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

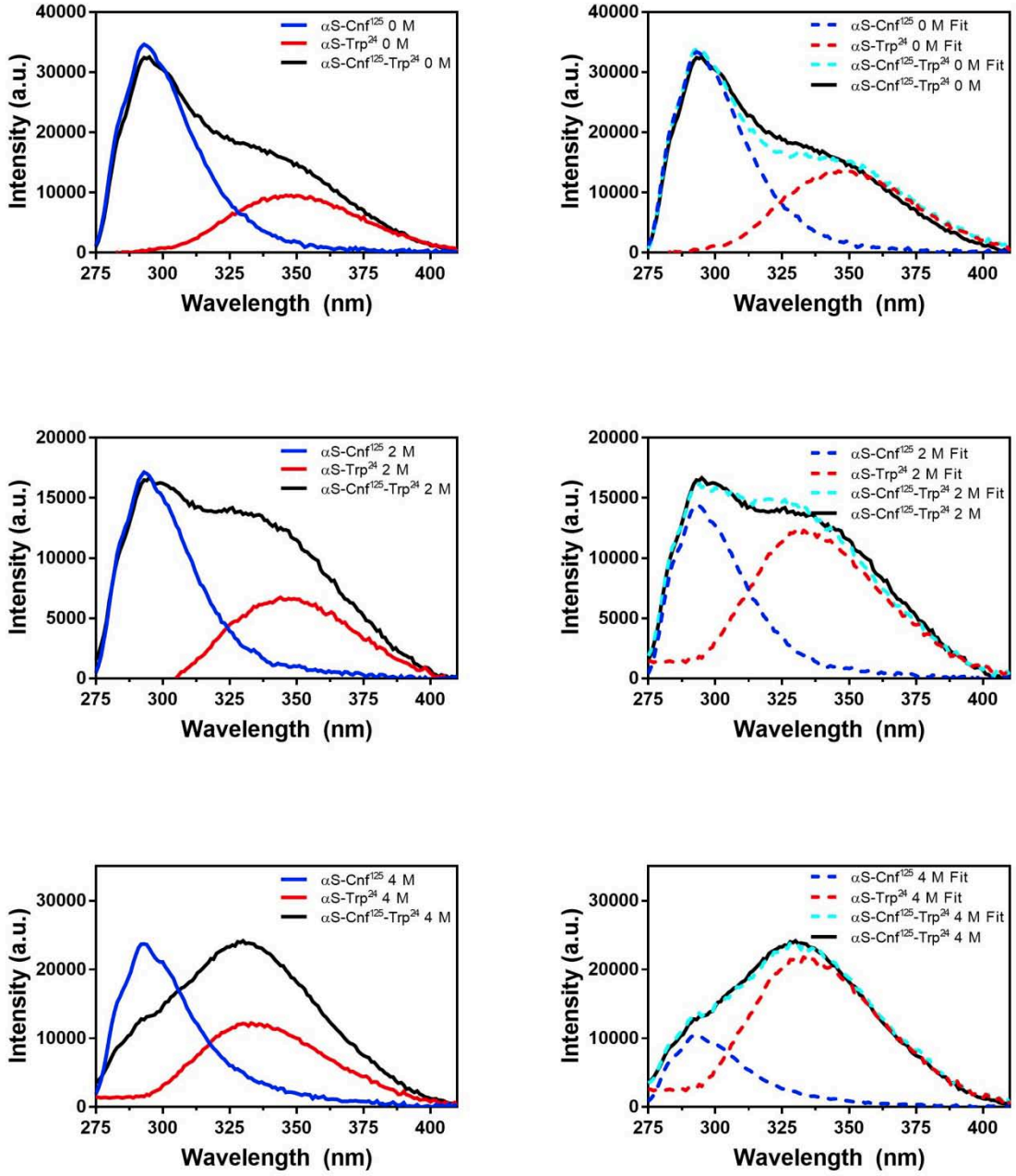

Figure S25: Steady-state Fluorescence Spectra of αS-Cnf^{125}-Trp^{24}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

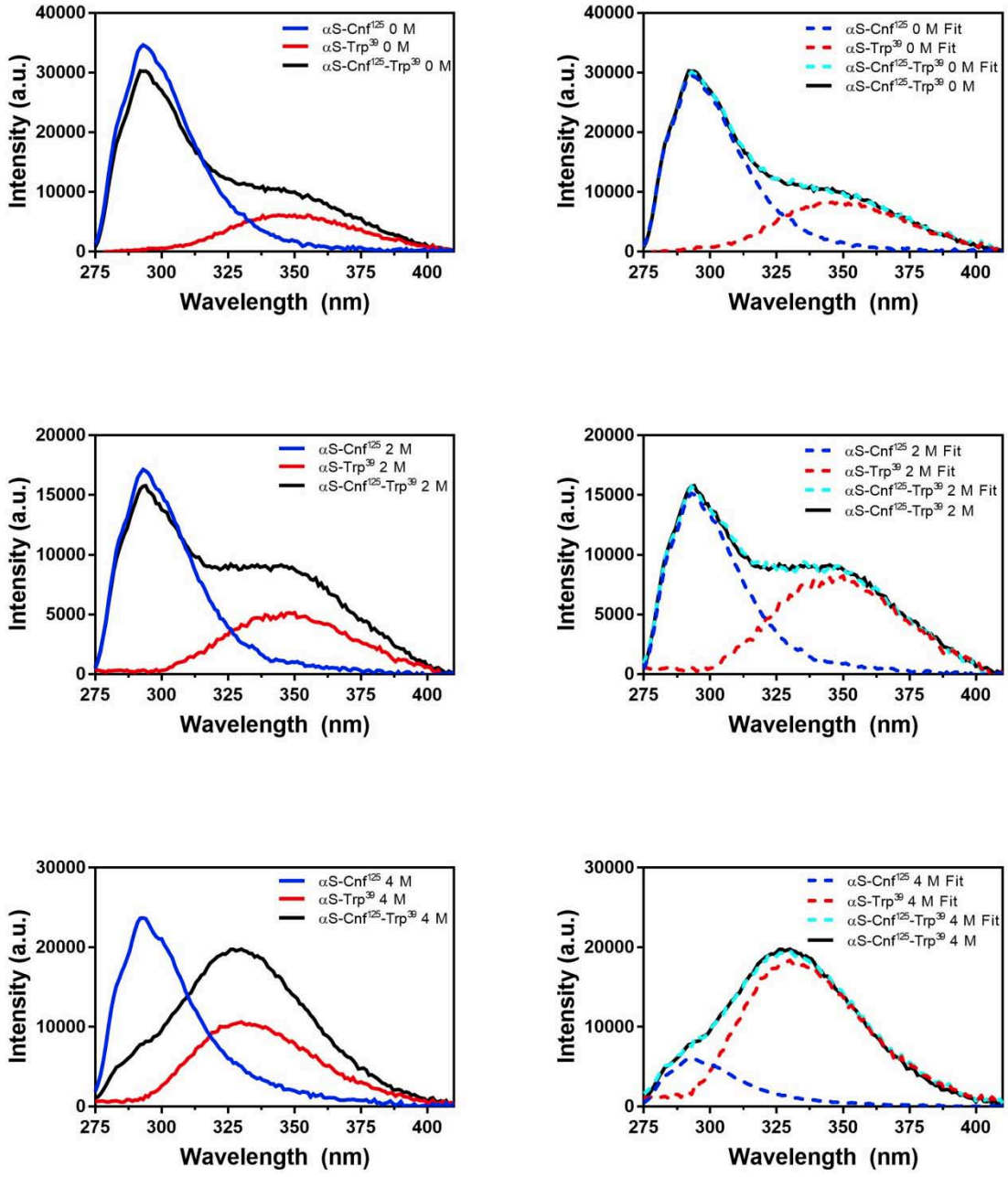

Figure S26: Steady-state Fluorescence Spectra of αS-Cnf^{125}-Trp^{39}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

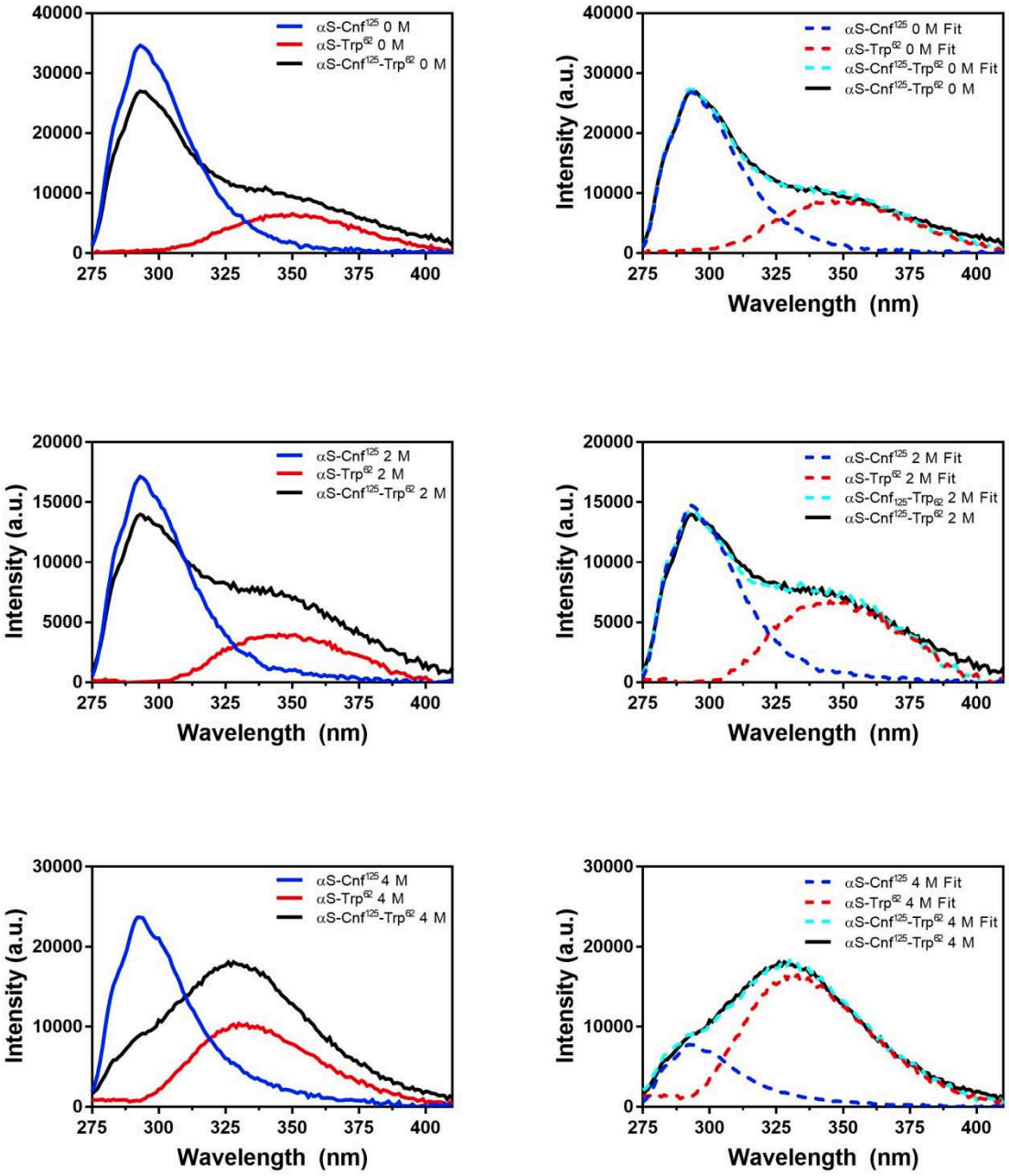

Figure S27: Steady-state Fluorescence of αS-Cnf^{125}-Trp^{62}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

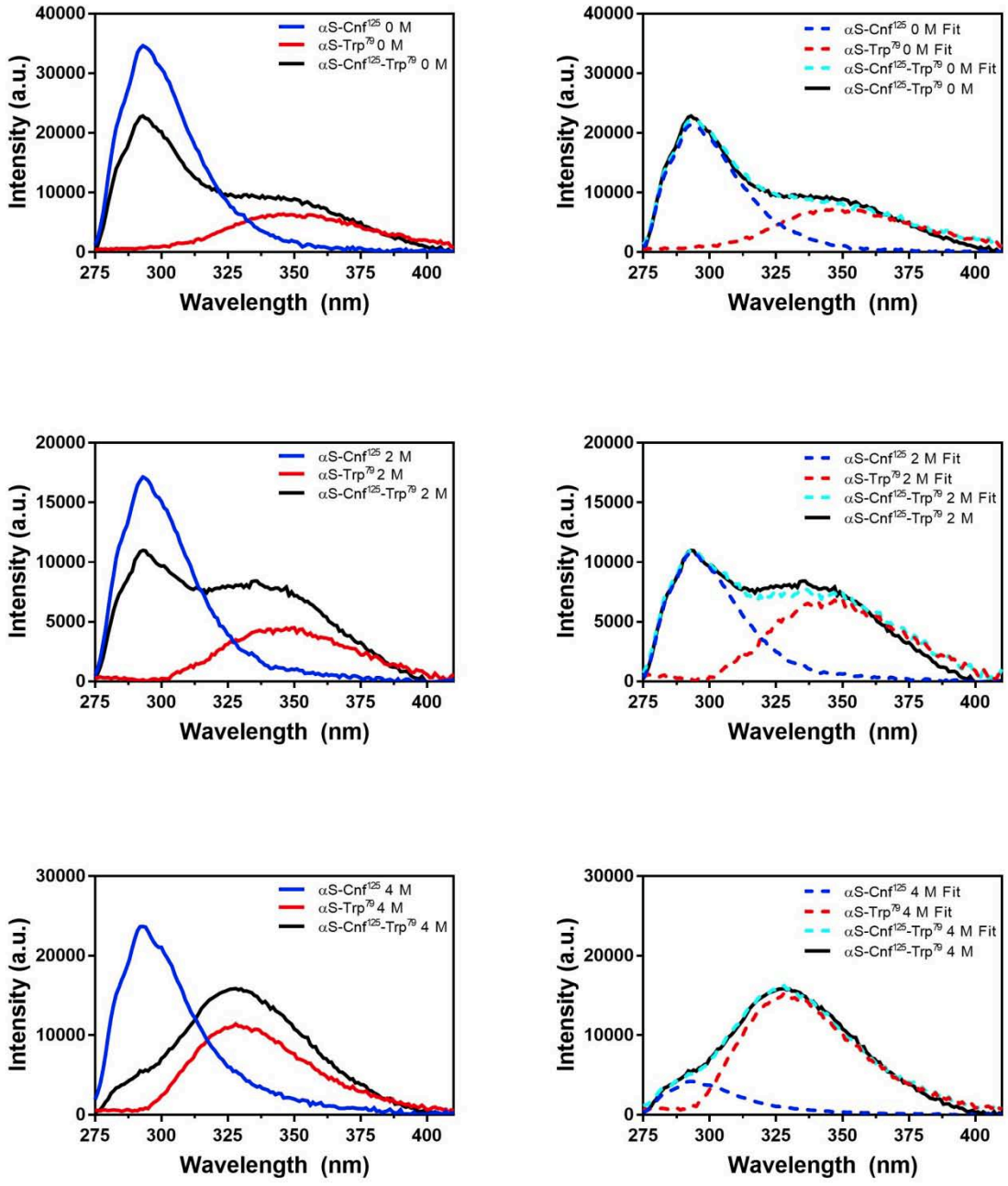

Figure S28: Steady-state Fluorescence Spectra of αS-Cnf^{125}-Trp^{79}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

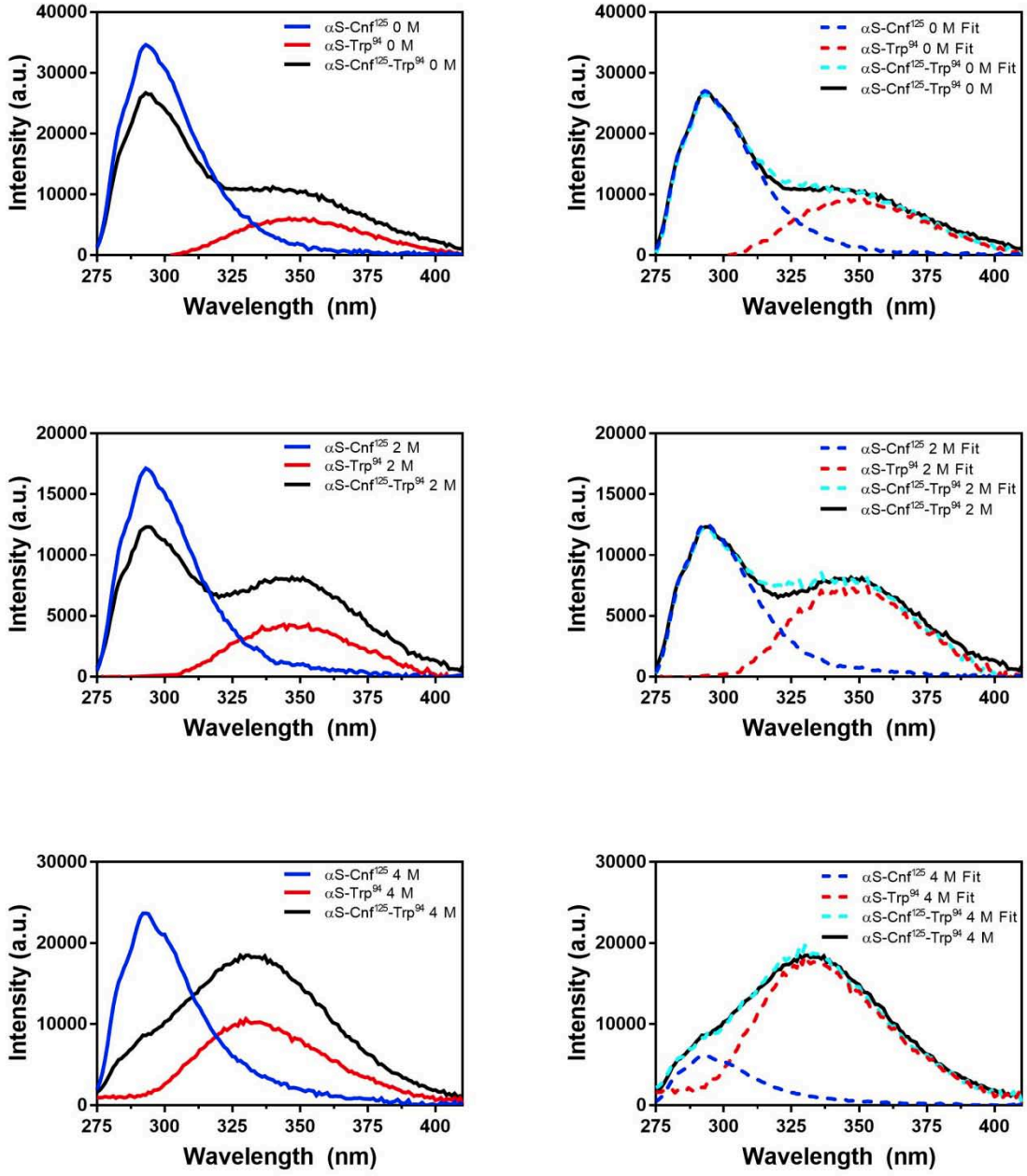

Figure S29: Steady-state Fluorescence Spectra of αS-Cnf^{125}-Trp^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

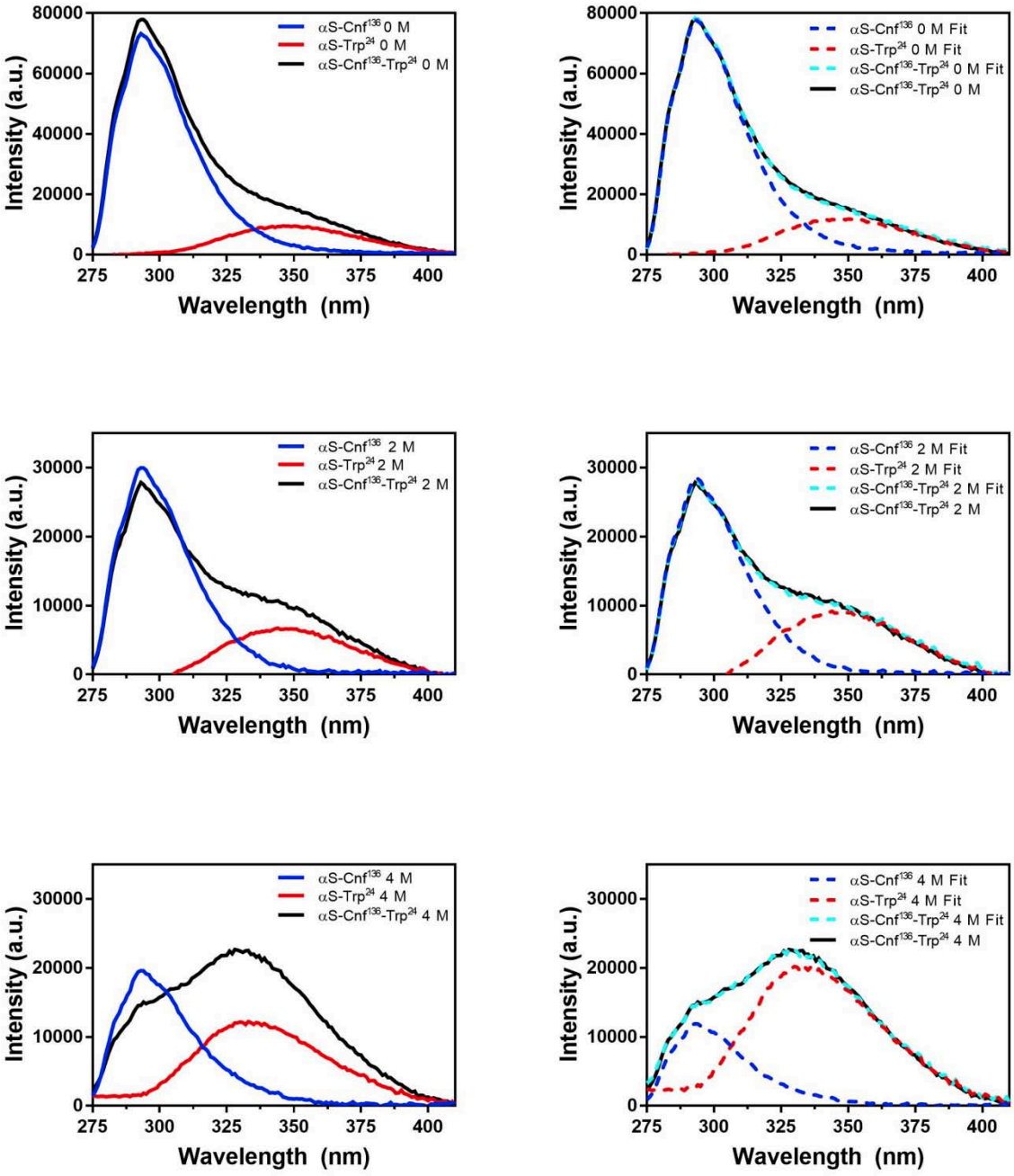

Figure S30: Steady-state Fluorescence Spectra of αS-Cnf^{136}-Trp^{24}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

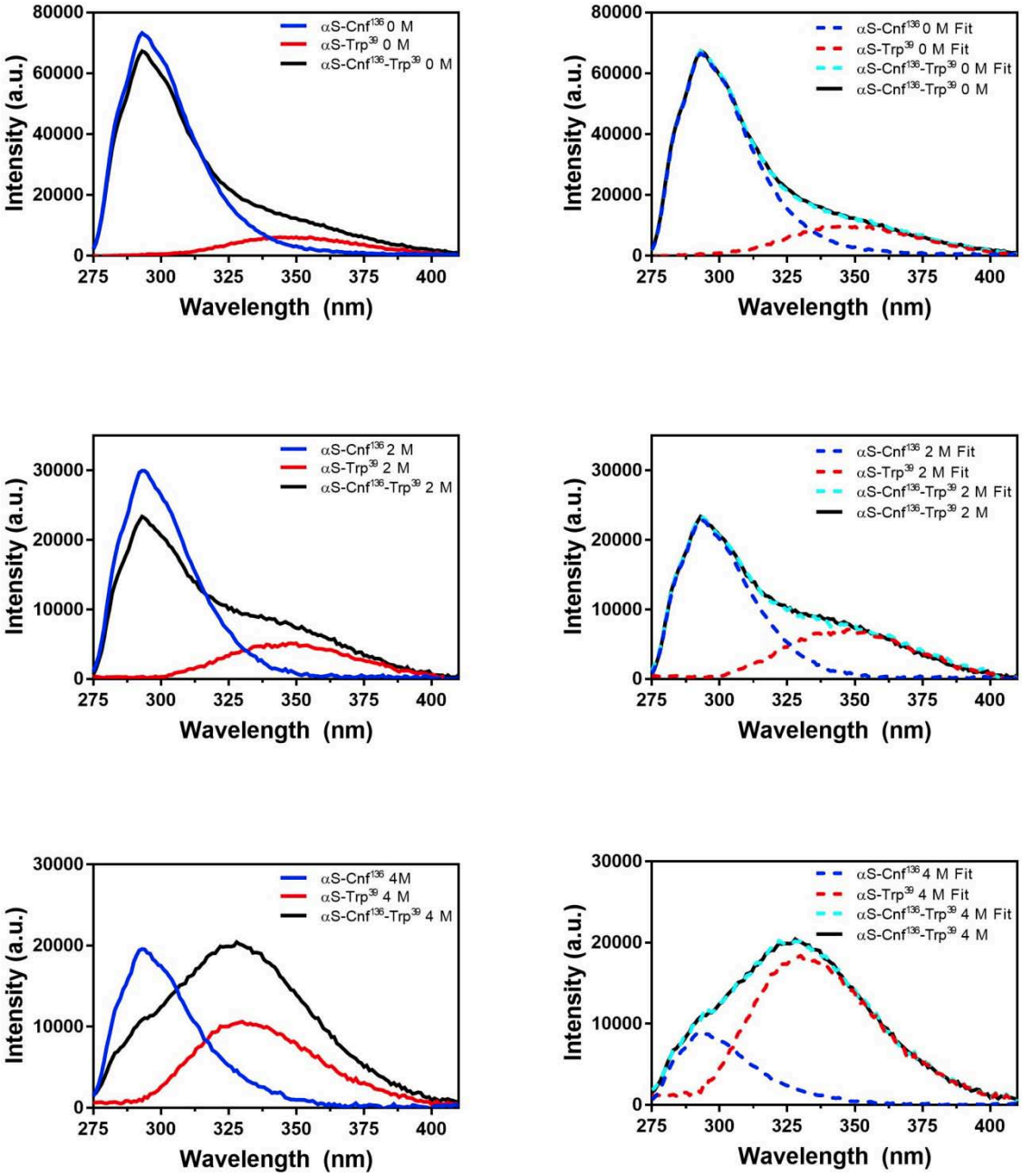

Figure S31: Steady-state Fluorescence Spectra of αS-Cnf^{136}-Trp^{39}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

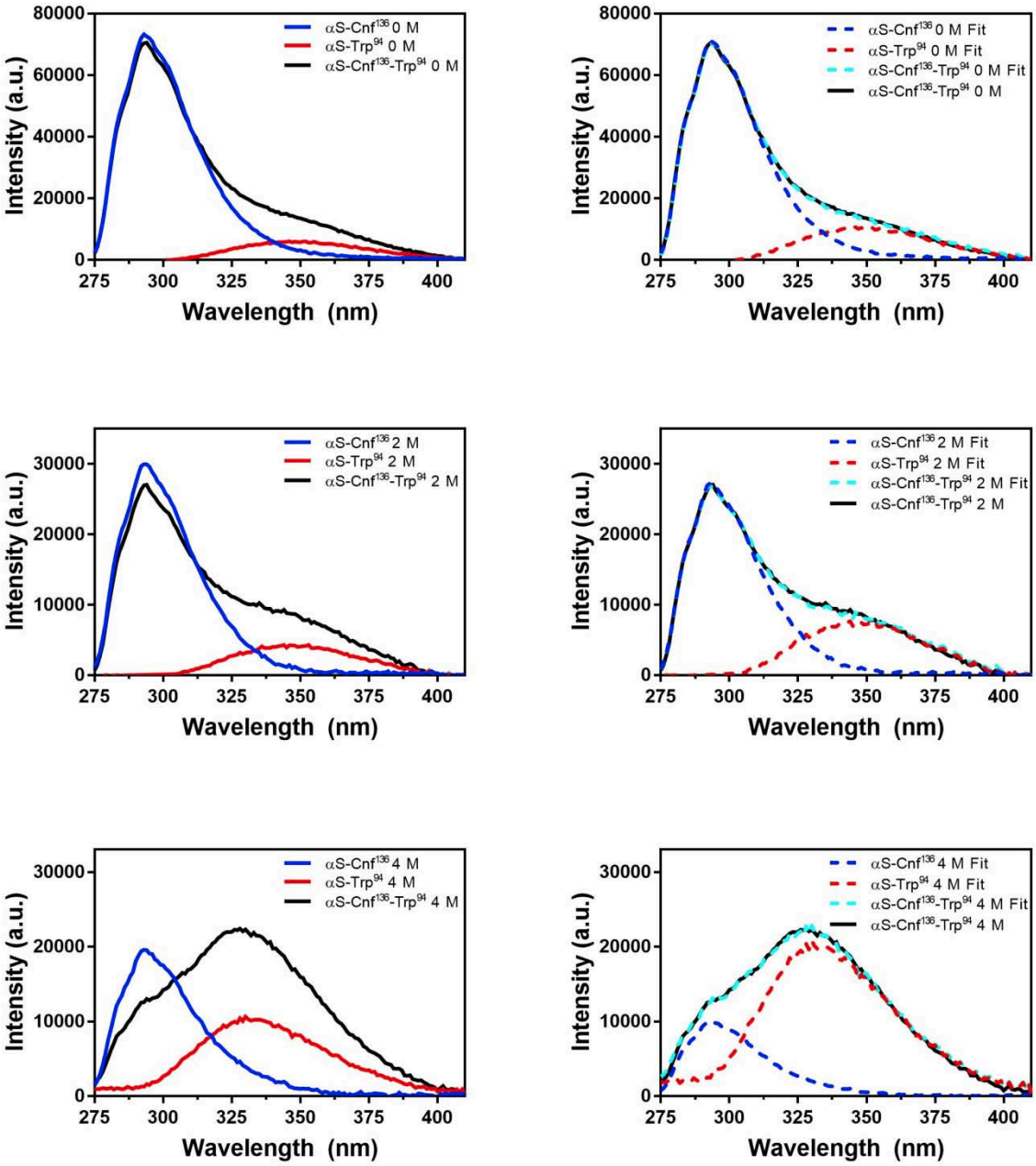

Figure S32: Steady-state Fluorescence Spectra of αS-Cnf^{136}-Trp^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

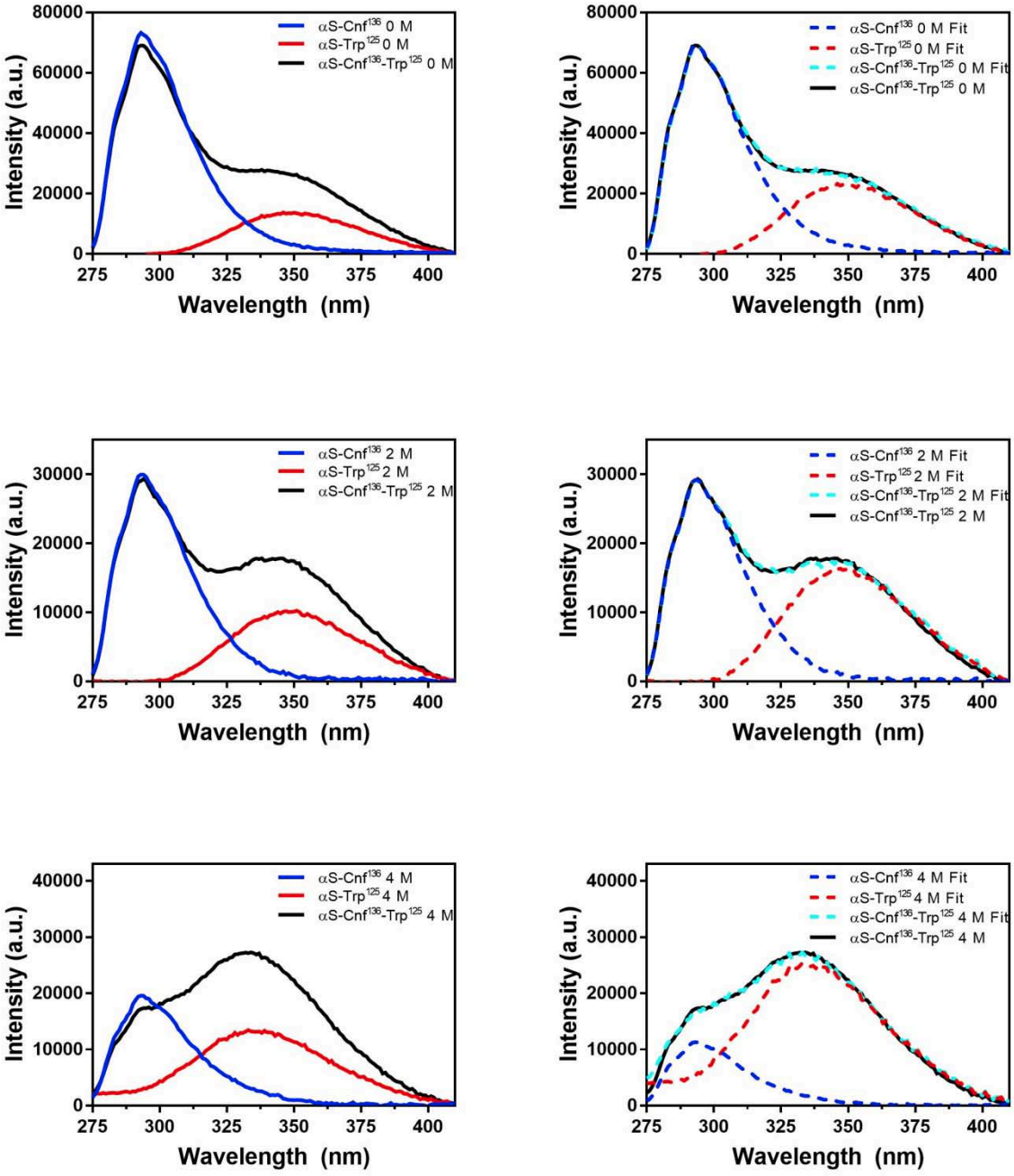

Figure S33: Steady-state Fluorescence Spectra of αS-Cnf^{136}-Trp^{125}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Cnf (blue) and Trp (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Cnf (red dashed) and Trp (blue dashed) to the fit.

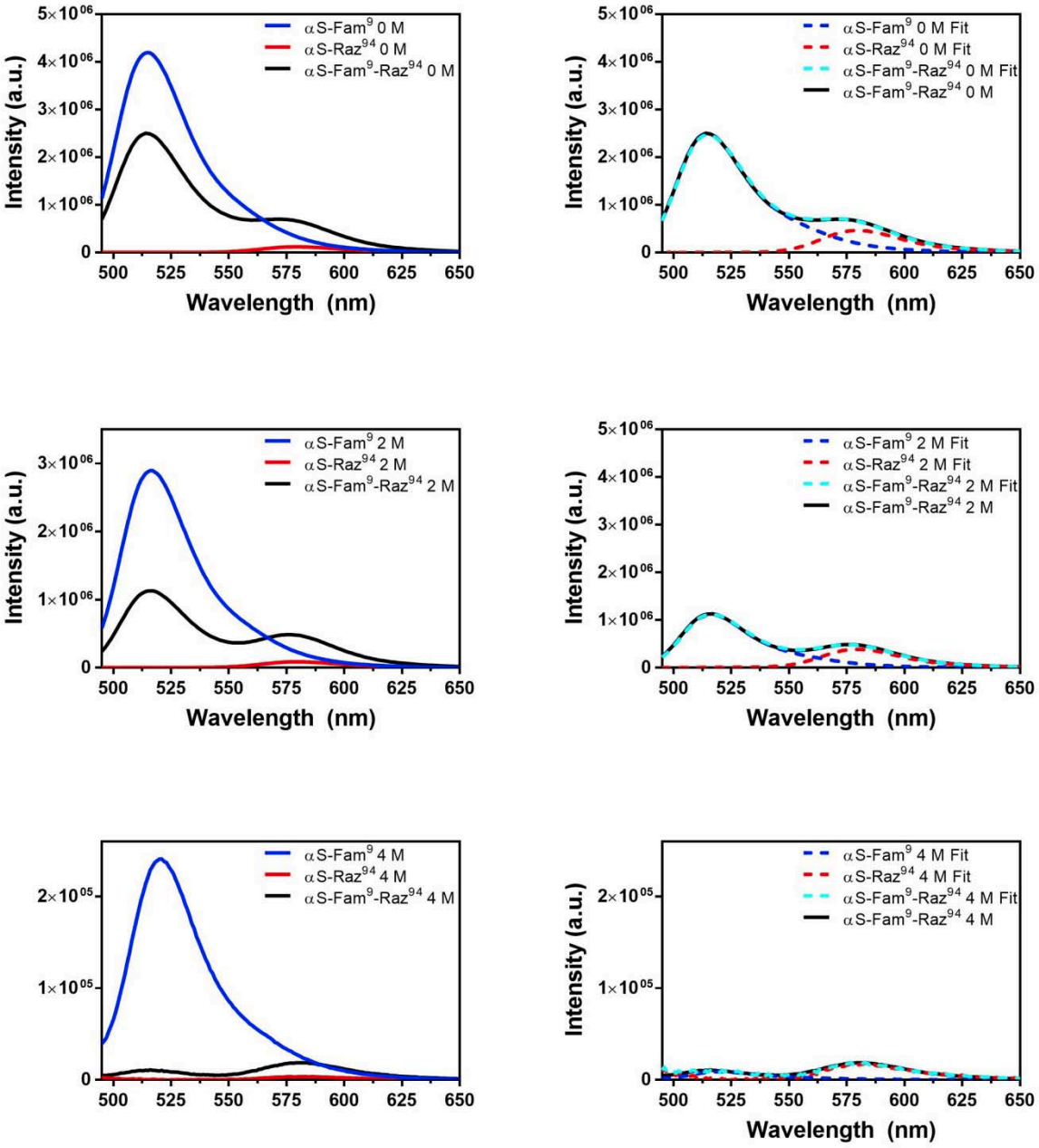

Figure S34: Steady-state Fluorescence Spectra of αS-Fam^{9}-Raz^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Fam (blue) and Raz (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Fam (red dashed) and Raz (blue dashed) to the fit.

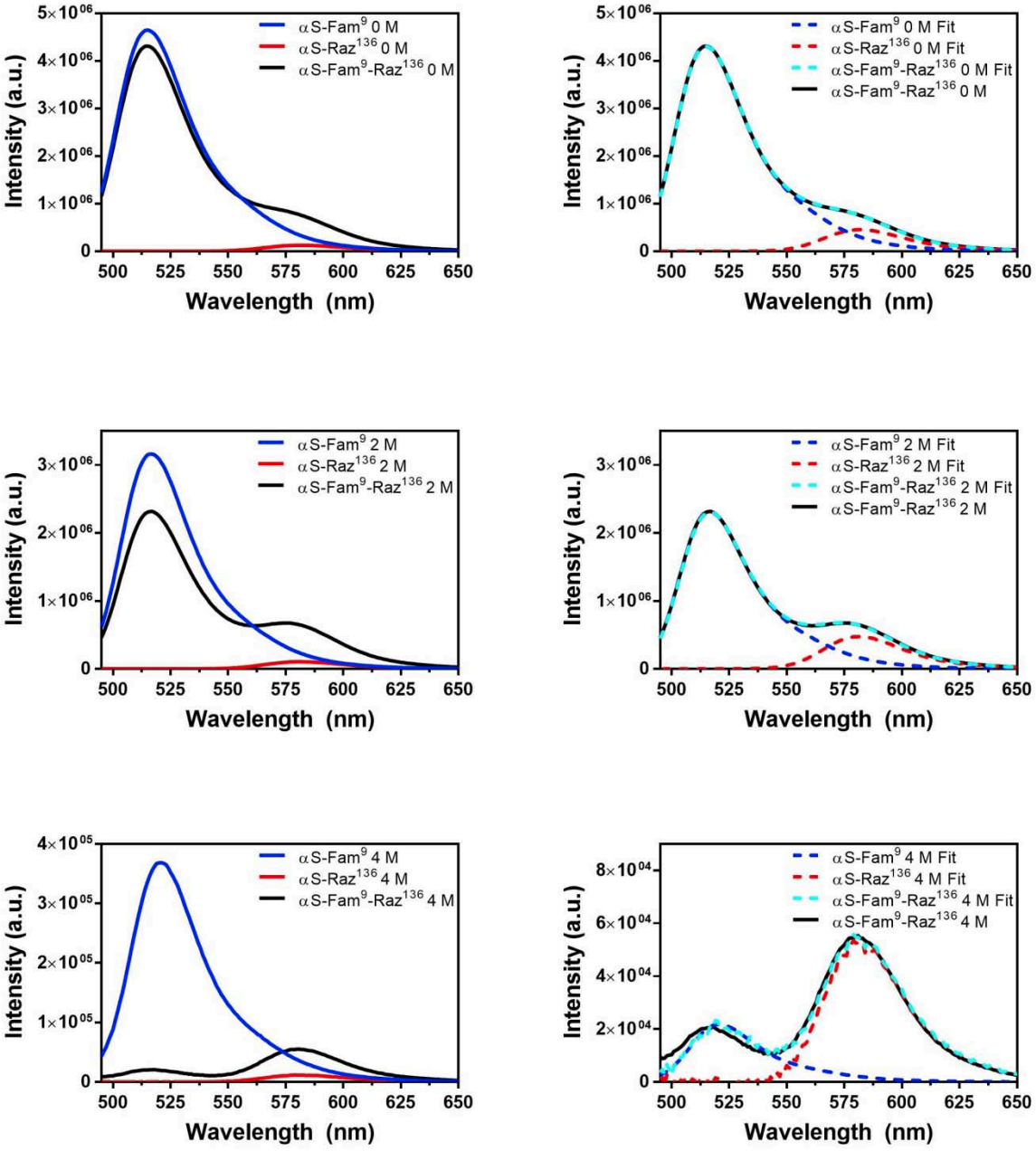

Figure S35: Steady-state Fluorescence Spectra of αS-Fam^{9}-Raz^{136}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Fam (blue) and Raz (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Fam (red dashed) and Raz (blue dashed) to the fit.

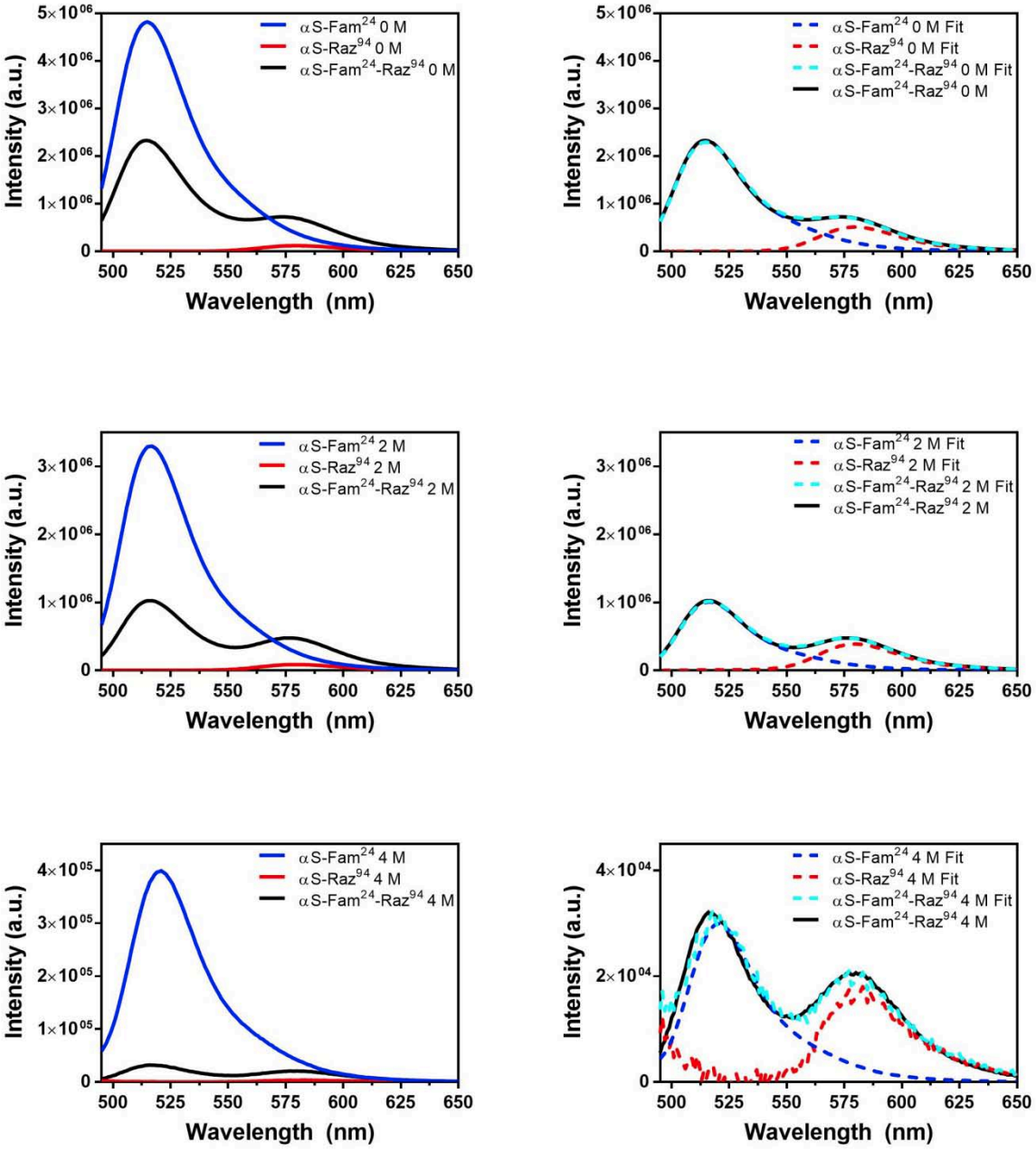

Figure S36: Steady-state Fluorescence Spectra of αS-Fam^{24}-Raz^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Fam (blue) and Raz (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Fam (red dashed) and Raz (blue dashed) to the fit.

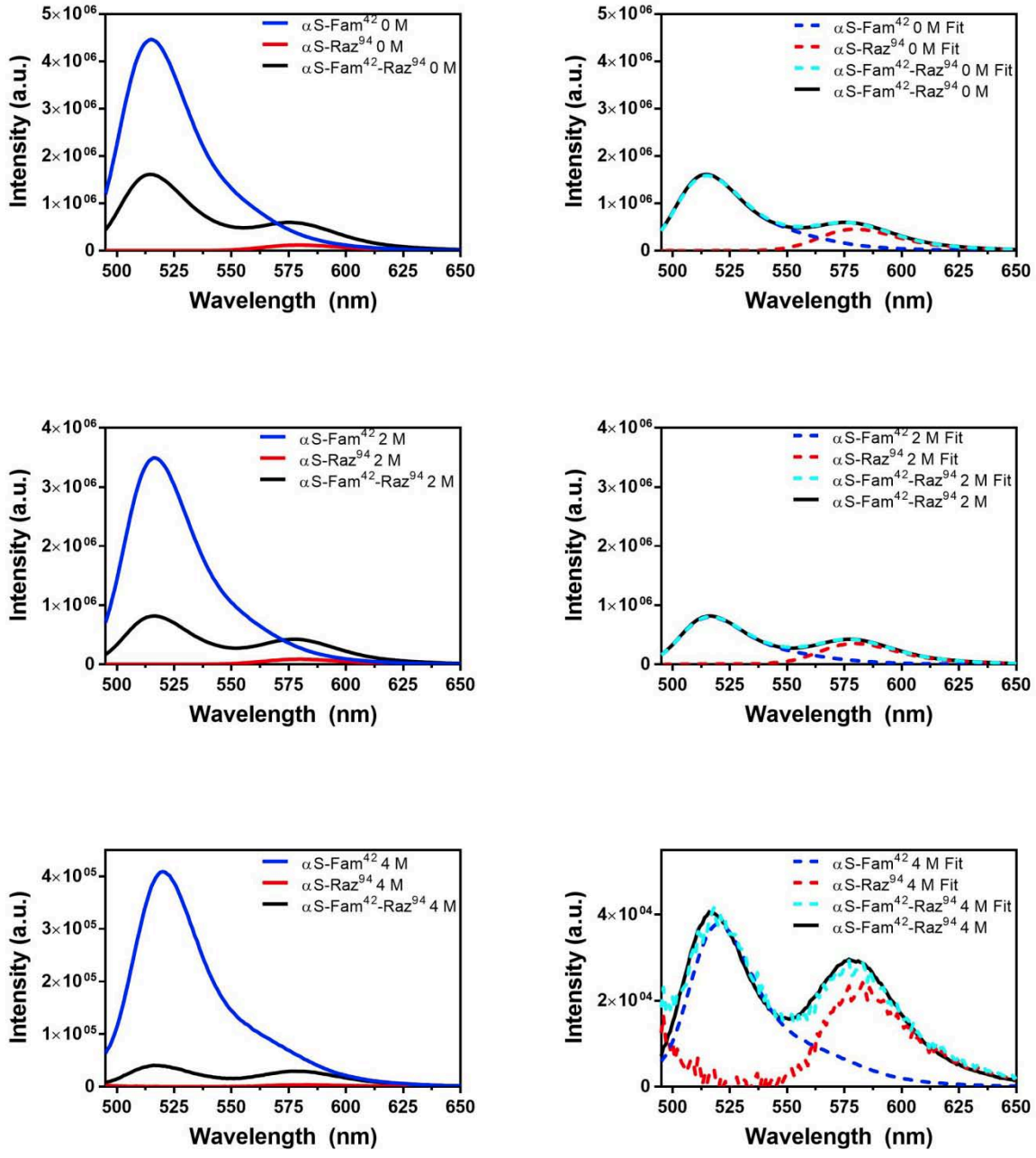

Figure S37: Steady-state Fluorescence Spectra of αS-Fam^{42}-Raz^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Fam (blue) and Raz (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Fam (red dashed) and Raz (blue dashed) to the fit.

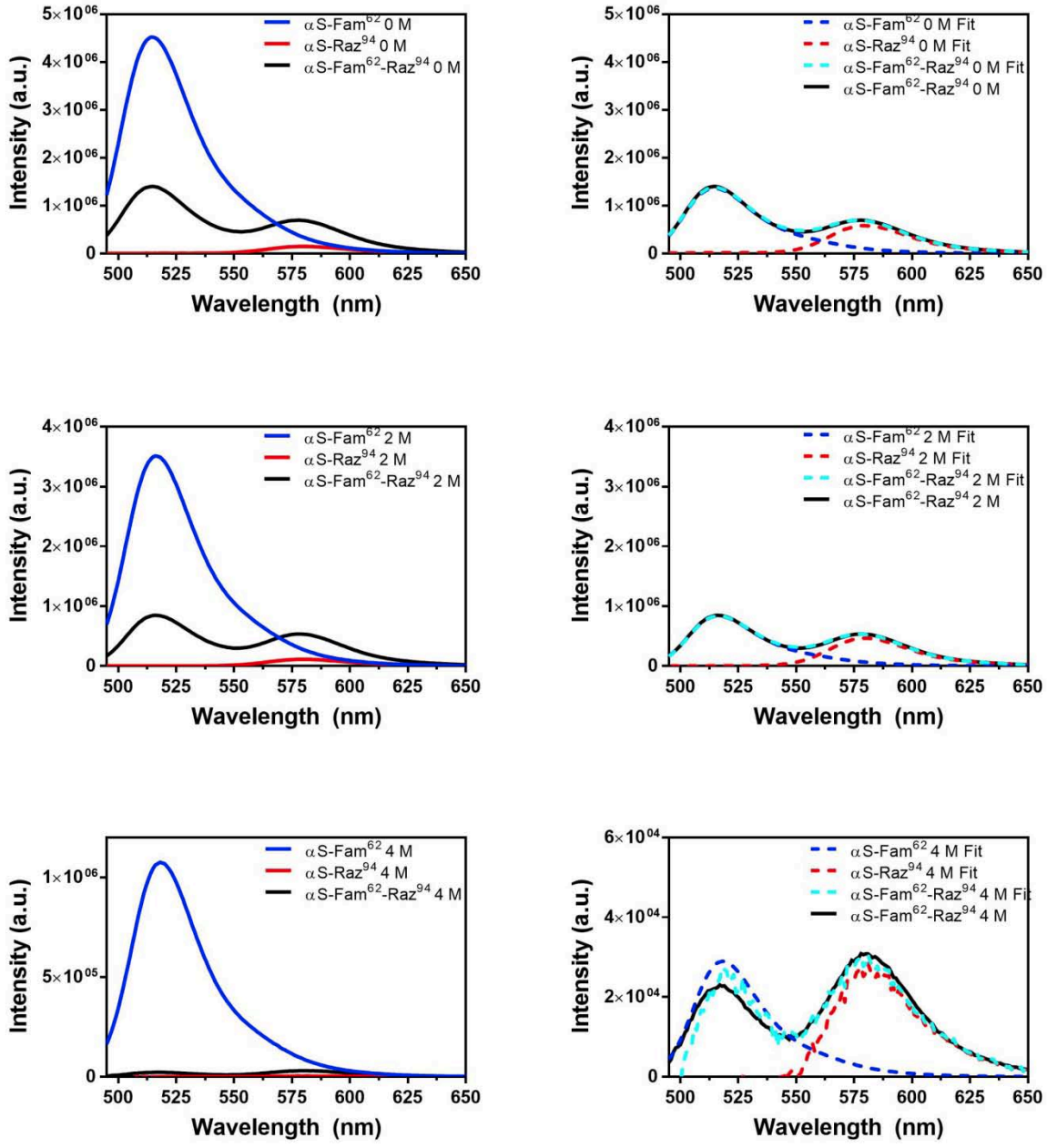

Figure S38: Steady-state Fluorescence Spectra of αS-Fam^{62}-Raz^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Fam (blue) and Raz (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Fam (red dashed) and Raz (blue dashed) to the fit.

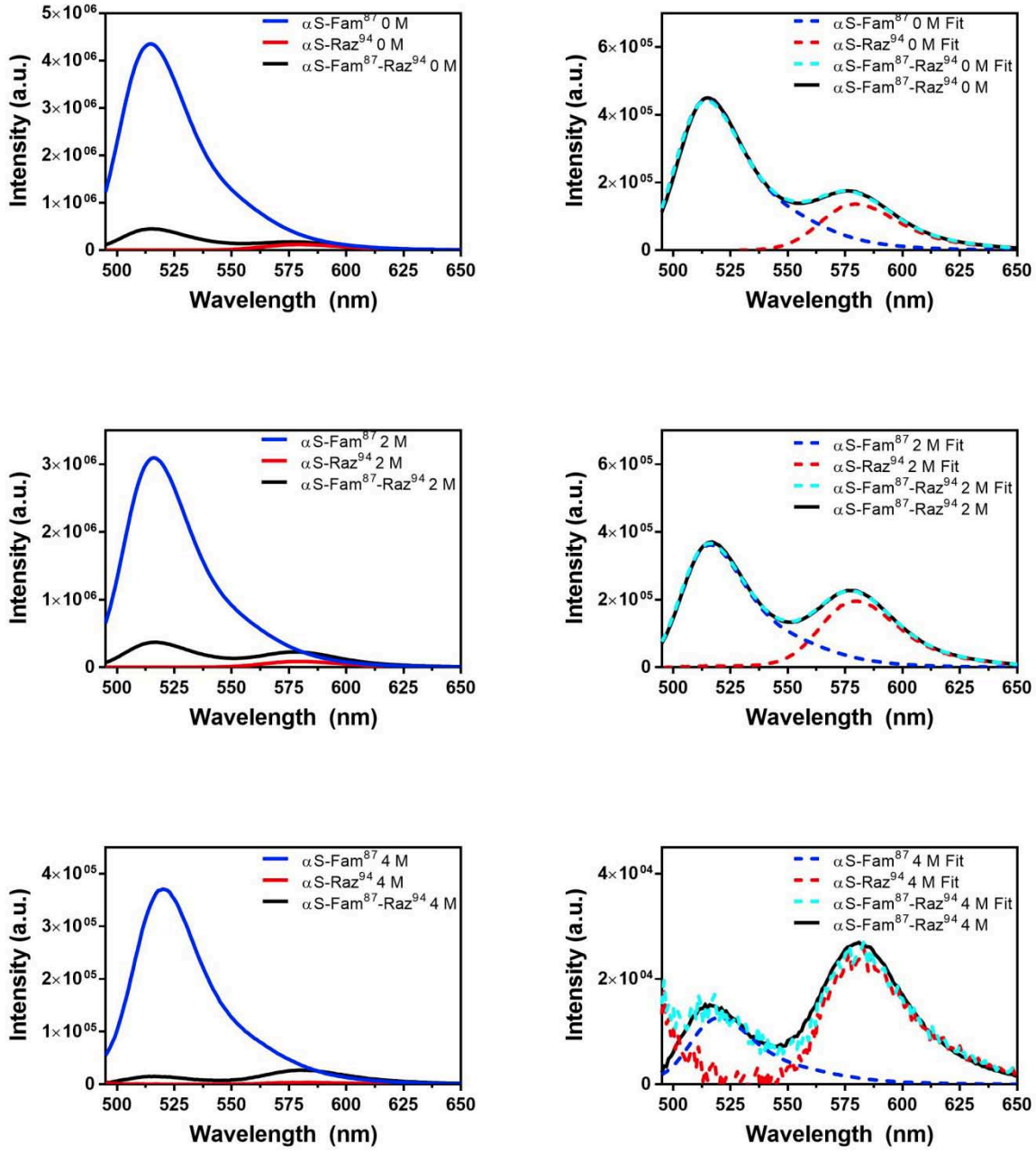

Figure S39: Steady-state Fluorescence Spectra of αS-Fam^{87}-Raz^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Fam (blue) and Raz (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Fam (red dashed) and Raz (blue dashed) to the fit.

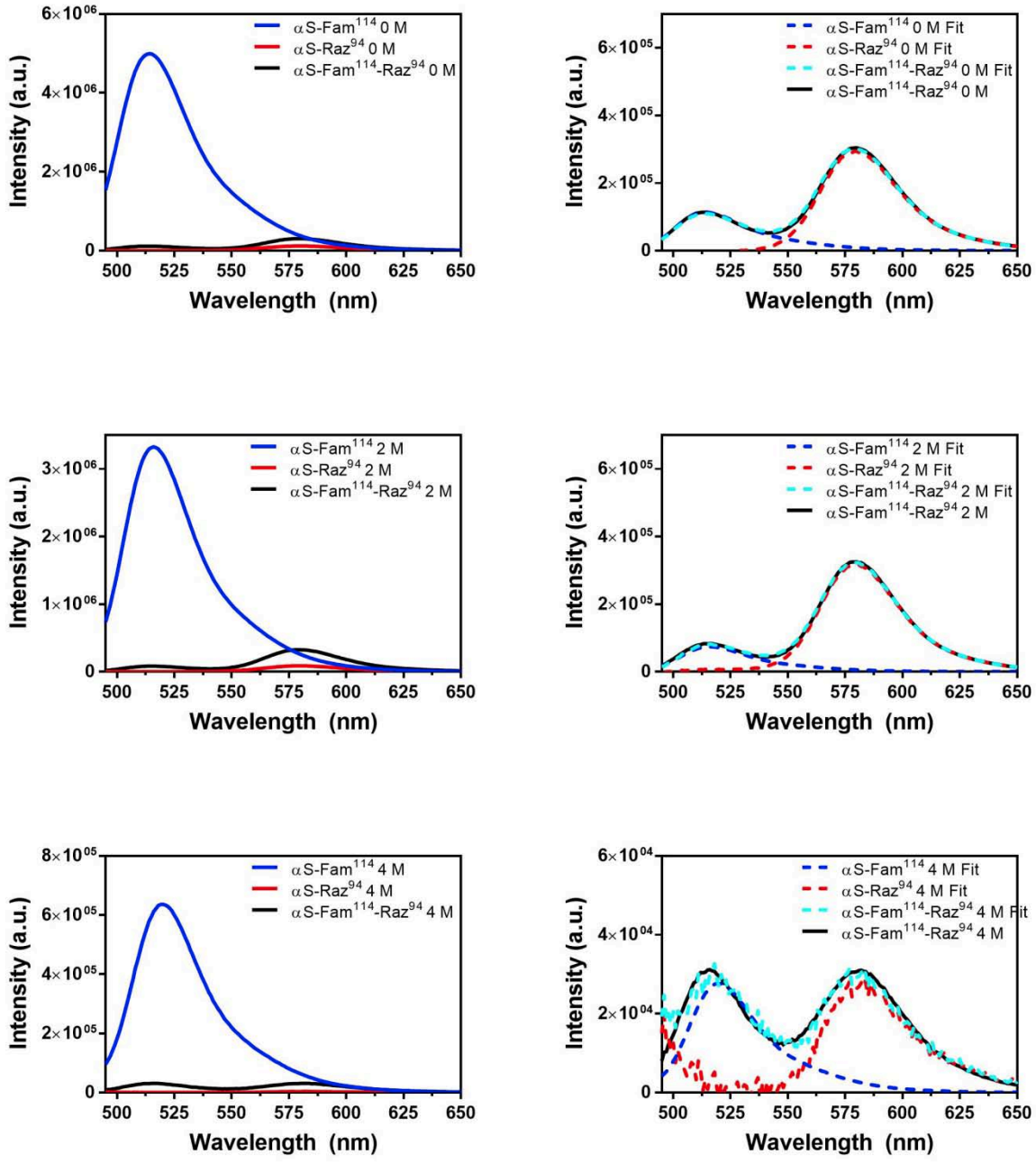

Figure S40: Steady-state Fluorescence Spectra of αS-Fam^{114}-Raz^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Fam (blue) and Raz (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Fam (red dashed) and Raz (blue dashed) to the fit.

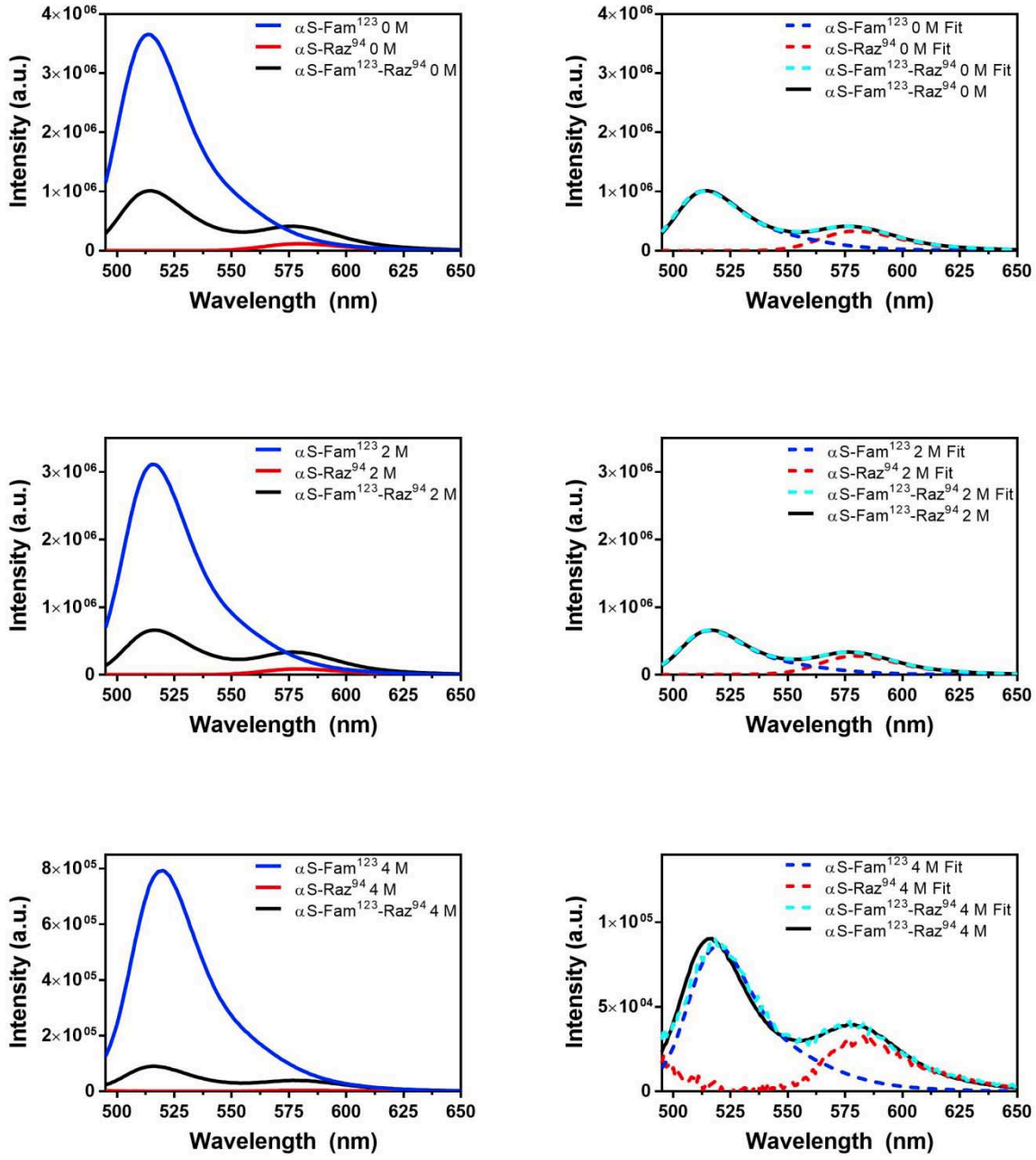

Figure S41: Steady-state Fluorescence Spectra of αS-Fam^{123}-Raz^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Fam (blue) and Raz (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Fam (red dashed) and Raz (blue dashed) to the fit.

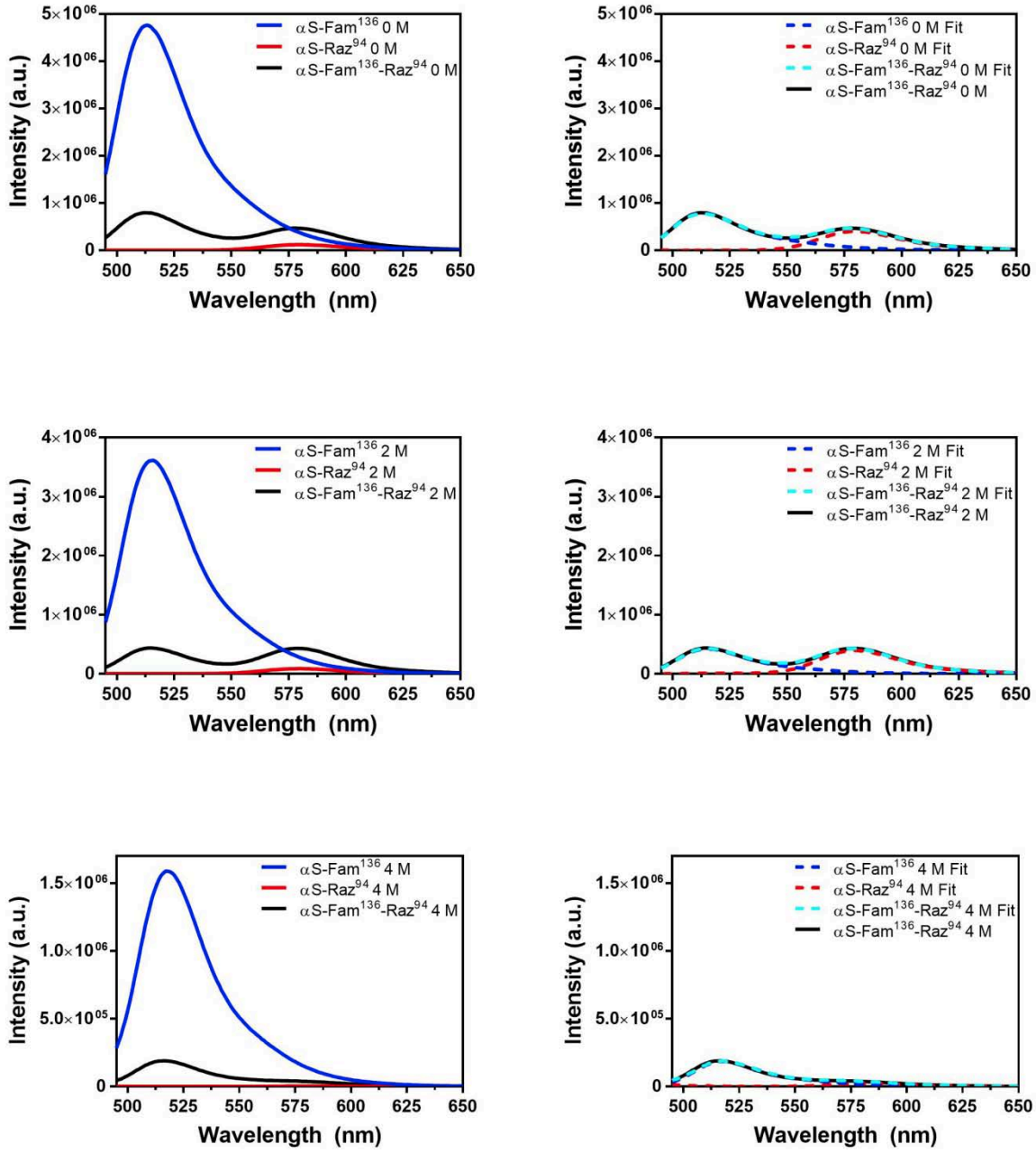

Figure S42: Steady-state Fluorescence Spectra of αS-Fam^{136}-Raz^{94}. Plots show background subtracted data (Left) and fit data (Right) for 0, 2 and 4 M TMAO (Top-Bottom). Background substracted plots contain cooresponding single-labeled Fam (blue) and Raz (red) labed αS along with the double-labeled potein (black). Plots showing the fits contain the background subtacted double-labeled potein (black) spectrum along with the fit (cyan dashed) and contributions of Fam (red dashed) and Raz (blue dashed) to the fit.

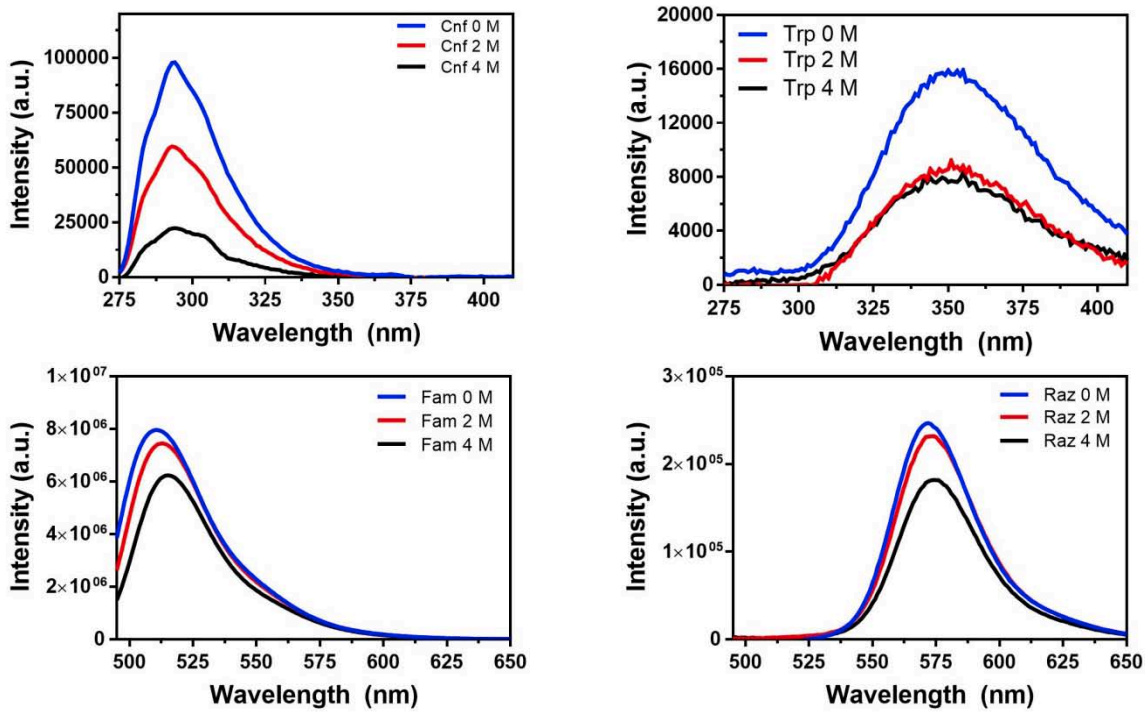

Figure S43: Steady-state Fluorescence Spectra of Free Fluorophores. Plots show background subtracted spectra of Cnf (top left), Trp (top right), Fam (bottom left) and Raz (bottom right) for 0 (blue), 2 (red) and 4 M TMAO (black).

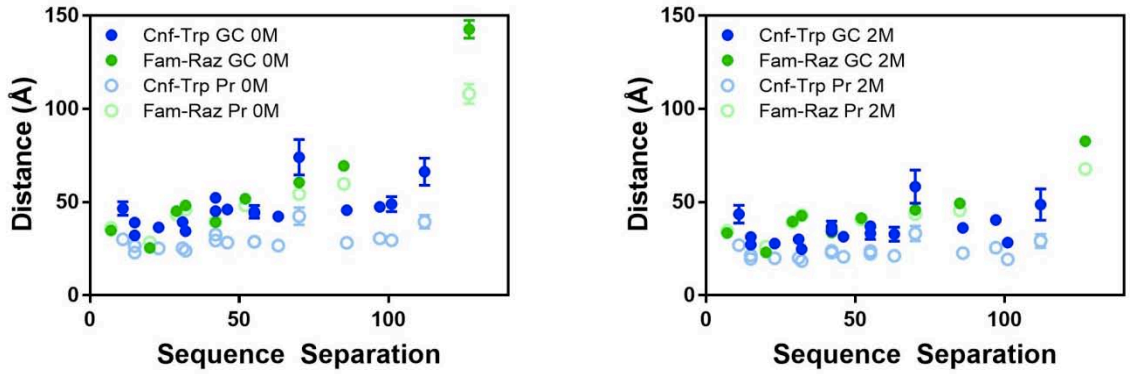

Figure S44: Distance from Polymer-Scaled Förster Equation. Plot show distances obtained from steady-state derived E_{FRET} values through the Gaussian chain (GC) or semi-epirical (Pr) forms of the polymer-scaled Förster equation. Plots are shown for data obtained in 0 M (left) and 2 M (right) TMAO.

Table S3a. Steady-State FRET Fitting and Distance Determination

														Gaussian Chain		Semi-Empirical	
Pair	PSS	TMAO	Q_D	R_0	$\varepsilon_A/\varepsilon_D$	A	E_D	S_A	B	E_A	S_B	E_{FRET}	S_{EFRET}	r (Å)	S_{Dist}	r (Å)	S_{Dist}
Cnf^{39}-Trp^{24}	15	0M	0.10	16.7	0.13	0.587	0.413	0.094	1.619	0.080	0.005	0.096	0.006	39.1	0.8	26.4	0.4
	15	2M	0.04	14.6	0.13	0.613	0.387	0.140	1.782	0.102	0.008	0.116	0.010	31.3	0.9	21.4	0.4
	15	4M	0.04	14.2	0.13	0.233	0.767	0.146	1.390	0.051	0.010	0.094	0.013	30.0	1.0	20.7	0.5
Cnf^{39}-Trp^{62}	23	0M	0.10	16.7	0.13	0.599	0.401	0.138	1.841	0.109	0.006	0.121	0.008	36.5	0.7	25.2	0.3
	23	2M	0.04	14.6	0.13	0.535	0.465	0.218	2.173	0.153	0.007	0.163	0.010	27.8	0.7	19.9	0.3
	23	4M	0.04	14.2	0.13	0.155	0.845	0.209	1.386	0.050	0.010	0.085	0.013	31.7	1.3	21.4	0.6
Cnf^{39}-Trp^{79}	42	0M	0.10	16.7	0.13	0.675	0.325	0.078	1.201	0.026	0.005	0.044	0.007	52.3	2.2	32.6	1.0
	42	2M	0.04	14.6	0.13	0.589	0.411	0.198	1.564	0.073	0.011	0.090	0.014	34.6	1.5	22.9	0.7
	42	4M	0.04	14.2	0.13	0.111	0.889	0.177	1.382	0.050	0.009	0.089	0.012	31.0	0.5	21.1	0.2
Cnf^{39}-Trp^{94}	55	0M	0.10	16.7	0.13	0.920	0.080	0.097	1.697	0.091	0.010	0.090	0.012	44.1	2.3	28.8	1.1
	55	2M	0.04	14.6	0.13	0.950	0.050	0.206	2.194	0.155	0.025	0.144	0.031	33.2	3.1	22.3	1.4
	55	4M	0.04	14.2	0.13	0.250	0.750	0.168	1.938	0.122	0.006	0.145	0.009	25.5	0.5	18.6	0.2
Cnf^{94}-Trp^{24}	70	0M	0.08	16.2	0.13	1.104	-0.104	0.033	1.156	0.020	0.007	0.020	0.007	74.1	9.5	42.5	4.7
	70	2M	0.04	14.4	0.13	0.916	0.084	0.059	1.122	0.016	0.010	0.026	0.012	58.4	8.9	33.2	4.0
	70	4M	0.03	13.8	0.13	0.350	0.650	0.080	1.496	0.064	0.009	0.123	0.011	27.4	1.0	19.2	0.4
Cnf^{94}-Trp^{39}	55	0M	0.08	16.2	0.13	0.945	0.055	0.097	1.655	0.085	0.013	0.082	0.016	44.8	3.3	28.7	1.5
	55	2M	0.04	14.4	0.13	0.741	0.259	0.088	1.496	0.064	0.008	0.081	0.010	37.0	1.7	23.9	0.7
	55	4M	0.03	13.8	0.13	0.265	0.735	0.094	2.002	0.130	0.006	0.166	0.008	23.3	0.5	17.4	0.2
Cnf^{94}-Trp^{62}	32	0M	0.08	16.2	0.13	0.950	0.050	0.080	2.401	0.182	0.009	0.169	0.011	34.4	1.1	23.9	0.5
	32	2M	0.04	14.4	0.13	0.853	0.147	0.135	3.059	0.268	0.012	0.258	0.015	24.8	0.7	18.4	0.3
	32	4M	0.03	13.8	0.13	0.305	0.695	0.092	1.920	0.120	0.006	0.153	0.008	24.2	0.5	17.8	0.2
Cnf^{94}-Trp^{79}	15	0M	0.08	16.2	0.13	0.543	0.457	0.134	2.045	0.136	0.006	0.149	0.008	32.2	0.5	22.9	0.2
	15	2M	0.04	14.4	0.13	0.453	0.547	0.182	2.114	0.145	0.008	0.162	0.011	27.2	0.6	19.5	0.3
	15	4M	0.03	13.8	0.13	0.226	0.774	0.103	1.543	0.071	0.008	0.119	0.010	26.4	0.8	18.8	0.4
Cnf^{125}-Trp^{24}	101	0M	0.05	15.0	0.13	0.965	0.035	0.057	1.423	0.055	0.009	0.052	0.011	48.9	4.0	29.6	1.7
	101	2M	0.02	13.3	0.13	0.842	0.158	0.105	2.116	0.145	0.024	0.148	0.028	28.3	2.0	19.3	0.8
	101	4M	0.03	14.1	0.13	0.440	0.560	0.121	1.785	0.102	0.013	0.146	0.017	27.3	0.8	19.4	0.4
Cnf^{125}-Trp^{39}	86	0M	0.05	15.0	0.13	0.853	0.147	0.035	1.361	0.047	0.003	0.055	0.004	45.7	1.4	28.2	0.6
	86	2M	0.02	13.3	0.13	0.884	0.116	0.032	1.598	0.078	0.005	0.083	0.006	36.2	0.9	22.7	0.4
	86	4M	0.03	14.1	0.13	0.254	0.746	0.037	1.729	0.095	0.003	0.141	0.004	27.8	0.5	19.6	0.2
Cnf^{125}-Trp^{62}	63	0M	0.05	15.0	0.13	0.777	0.223	0.080	1.358	0.047	0.009	0.064	0.011	42.3	2.5	26.7	1.1
	63	2M	0.02	13.3	0.13	0.860	0.140	0.163	1.691	0.090	0.029	0.097	0.034	32.8	3.8	21.2	1.6
	63	4M	0.03	14.1	0.13	0.328	0.672	0.146	1.598	0.078	0.008	0.110	0.011	28.4	0.8	19.9	0.4
Cnf^{125}-Trp^{79}	46	0M	0.05	15.0	0.13	0.619	0.381	0.009	1.116	0.015	0.001	0.045	0.001	46.0	0.3	28.3	0.1
	46	2M	0.02	13.3	0.13	0.633	0.367	0.015	1.527	0.068	0.002	0.096	0.002	31.4	0.2	20.7	0.1
	46	4M	0.03	14.1	0.13	0.177	0.823	0.015	1.336	0.044	0.001	0.088	0.001	30.5	0.1	20.8	0.0
Cnf^{125}-Trp^{94}	31	0M	0.05	15.0	0.13	0.782	0.218	0.098	1.523	0.068	0.010	0.082	0.013	39.3	2.3	25.3	1.0
	31	2M	0.02	13.3	0.13	0.737	0.263	0.166	1.728	0.095	0.021	0.113	0.026	30.1	2.4	20.1	1.0
	31	4M	0.03	14.1	0.13	0.258	0.742	0.140	1.729	0.095	0.008	0.128	0.010	26.3	0.7	18.9	0.3

Table S3b. Steady-State FRET Fitting and Distance Determination

Pair	PSS	TMAO	Q_D	R_0	$\varepsilon_A/\varepsilon_D$	A	E_D	S_A	B	E_A	S_B	E_{FRET}	S_{EFRET}	Gaussian Chain r (Å)	Gaussian Chain S_{Dist}	Semi-Empirical r (Å)	Semi-Empirical S_{Dist}
Cnf^{136}-Trp^{24}	112	0M	0.10	17.0	0.13	1.065	-0.065	0.050	1.243	0.032	0.009	0.032	0.009	66.3	7.3	39.5	3.5
	112	2M	0.04	14.6	0.13	0.947	0.053	0.114	1.362	0.047	0.020	0.048	0.024	48.7	8.5	29.2	3.7
	112	4M	0.03	13.7	0.13	0.607	0.393	0.160	1.663	0.086	0.014	0.111	0.018	27.7	1.6	19.3	0.7
Cnf^{136}-Trp^{39}	97	0M	0.10	17.0	0.13	0.909	0.091	0.072	1.571	0.074	0.007	0.076	0.009	47.5	2.0	30.5	0.9
	97	2M	0.04	14.6	0.13	0.761	0.239	0.129	1.391	0.051	0.011	0.065	0.014	40.4	2.8	25.5	1.3
	97	4M	0.03	13.7	0.13	0.458	0.542	0.201	1.730	0.095	0.011	0.118	0.015	26.2	1.1	18.6	0.5
Cnf^{136}-Trp^{94}	42	0M	0.10	17.0	0.13	0.973	0.027	0.077	1.752	0.098	0.008	0.091	0.010	45.2	2.0	29.5	0.9
	42	2M	0.04	14.6	0.13	0.909	0.091	0.163	1.786	0.102	0.017	0.101	0.022	36.8	3.1	23.9	1.4
	42	4M	0.03	13.7	0.13	0.511	0.489	0.199	1.954	0.124	0.009	0.140	0.012	24.9	1.0	18.1	0.5
Cnf^{136}-Trp^{125}	11	0M	0.10	17.0	0.13	0.946	0.054	0.085	1.693	0.090	0.014	0.085	0.017	46.6	3.6	30.1	1.7
	11	2M	0.04	14.6	0.13	0.973	0.027	0.089	1.607	0.079	0.018	0.070	0.022	43.5	4.8	26.9	2.1
	11	4M	0.03	13.7	0.13	0.577	0.423	0.147	1.887	0.115	0.013	0.140	0.017	25.2	1.0	18.2	0.4
Fam^{9}-Raz^{94}	85	0M	0.49	54.3	0.24	0.590	0.410	0.009	3.932	0.717	0.075	0.444	0.012	69.5	1.2	59.8	0.8
	85	2M	0.33	50.9	0.24	0.384	0.616	0.005	4.439	0.841	0.033	0.643	0.006	49.3	0.4	45.6	0.3
	85	4M	0.03	33.9	0.24	0.037	0.963	0.010	4.961	0.968	0.123	0.964	0.013	16.4	1.2	18.5	1.0
Fam^{24}-Raz^{94}	70	0M	0.57	55.6	0.24	0.477	0.523	0.023	4.325	0.813	0.219	0.551	0.030	60.5	2.5	54.3	1.7
	70	2M	0.38	52.1	0.24	0.305	0.695	0.005	4.454	0.844	0.043	0.711	0.007	45.8	0.5	43.5	0.3
	70	4M	0.05	36.8	0.24	0.075	0.925	0.007	4.938	0.963	0.138	0.927	0.009	20.9	0.6	22.8	0.5
Fam^{42}-Raz^{94}	52	0M	0.52	54.9	0.24	0.357	0.643	0.008	3.864	0.700	0.068	0.649	0.010	51.8	0.7	48.2	0.5
	52	2M	0.40	52.6	0.24	0.229	0.771	0.004	4.051	0.746	0.038	0.769	0.005	41.4	0.4	40.7	0.3
	52	4M	0.05	37.1	0.24	0.093	0.907	0.008	6.634	1.378	0.174	0.907	0.011	22.3	0.7	23.9	0.5
Fam^{62}-Raz^{94}	32	0M	0.54	55.2	0.24	0.301	0.699	0.003	3.955	0.722	0.019	0.702	0.004	48.2	0.2	45.9	0.2
	32	2M	0.45	53.5	0.24	0.238	0.762	0.003	4.116	0.762	0.021	0.762	0.004	42.7	0.3	43.0	0.2
	32	4M	0.16	45.3	0.24	0.027	0.973	0.003	6.541	1.355	0.220	0.973	0.003	20.6	0.5	23.3	0.4
Fam^{87}-Raz^{94}	7	0M	0.51	54.7	0.24	0.102	0.898	0.001	1.154	0.038	0.010	0.807	0.002	34.8	0.1	36.1	0.1
	7	2M	0.35	51.5	0.24	0.117	0.883	0.002	2.244	0.304	0.015	0.817	0.002	33.4	0.2	34.6	0.2
	7	4M	0.04	36.3	0.24	0.034	0.966	0.009	7.093	1.490	0.176	0.966	0.009	17.3	1.0	19.6	0.8
Fam^{114}-Raz^{94}	20	0M	0.59	56.0	0.24	0.023	0.977	0.001	2.486	0.363	0.012	0.920	0.002	25.4	0.3	28.4	0.3
	20	2M	0.38	52.1	0.24	0.022	0.978	0.001	3.635	0.644	0.013	0.942	0.002	23.1	0.4	26.0	0.4
	20	4M	0.08	39.7	0.24	0.043	0.957	0.006	8.040	1.721	0.199	0.957	0.006	20.0	0.6	22.2	0.6
Fam^{123}-Raz^{94}	29	0M	0.42	53.0	0.24	0.276	0.724	0.006	2.821	0.445	0.040	0.689	0.007	45.2	0.5	43.4	0.4
	29	2M	0.36	51.5	0.24	0.210	0.790	0.003	3.205	0.539	0.023	0.762	0.004	39.6	0.3	39.2	0.2
	29	4M	0.09	41.2	0.24	0.108	0.892	0.005	8.941	1.941	0.220	0.892	0.005	25.8	0.3	27.2	0.3
Fam^{136}-Raz^{94}	42	0M	0.56	55.5	0.24	0.164	0.836	0.002	3.393	0.585	0.015	0.811	0.002	39.4	0.2	39.7	0.1
	42	2M	0.42	52.9	0.24	0.116	0.884	0.004	4.536	0.864	0.032	0.882	0.004	33.7	0.4	35.1	0.3
	42	4M	0.19	46.2	0.24	0.116	0.884	0.002	6.291	1.293	0.195	0.884	0.002	29.5	0.2	30.8	0.1
Fam^{9}-Raz^{136}	127	0M	0.55	55.5	0.24	0.925	0.075	0.006	3.558	0.625	0.048	0.134	0.007	142.7	4.8	108.1	5.3
	127	2M	0.40	52.6	0.24	0.729	0.271	0.013	4.500	0.856	0.090	0.348	0.017	82.7	2.7	67.9	1.7
	127	4M	0.06	37.9	0.24	0.059	0.941	0.005	4.517	0.860	0.036	0.931	0.006	20.4	0.6	22.5	0.5

Fluorescence Lifetime Data and Fitting

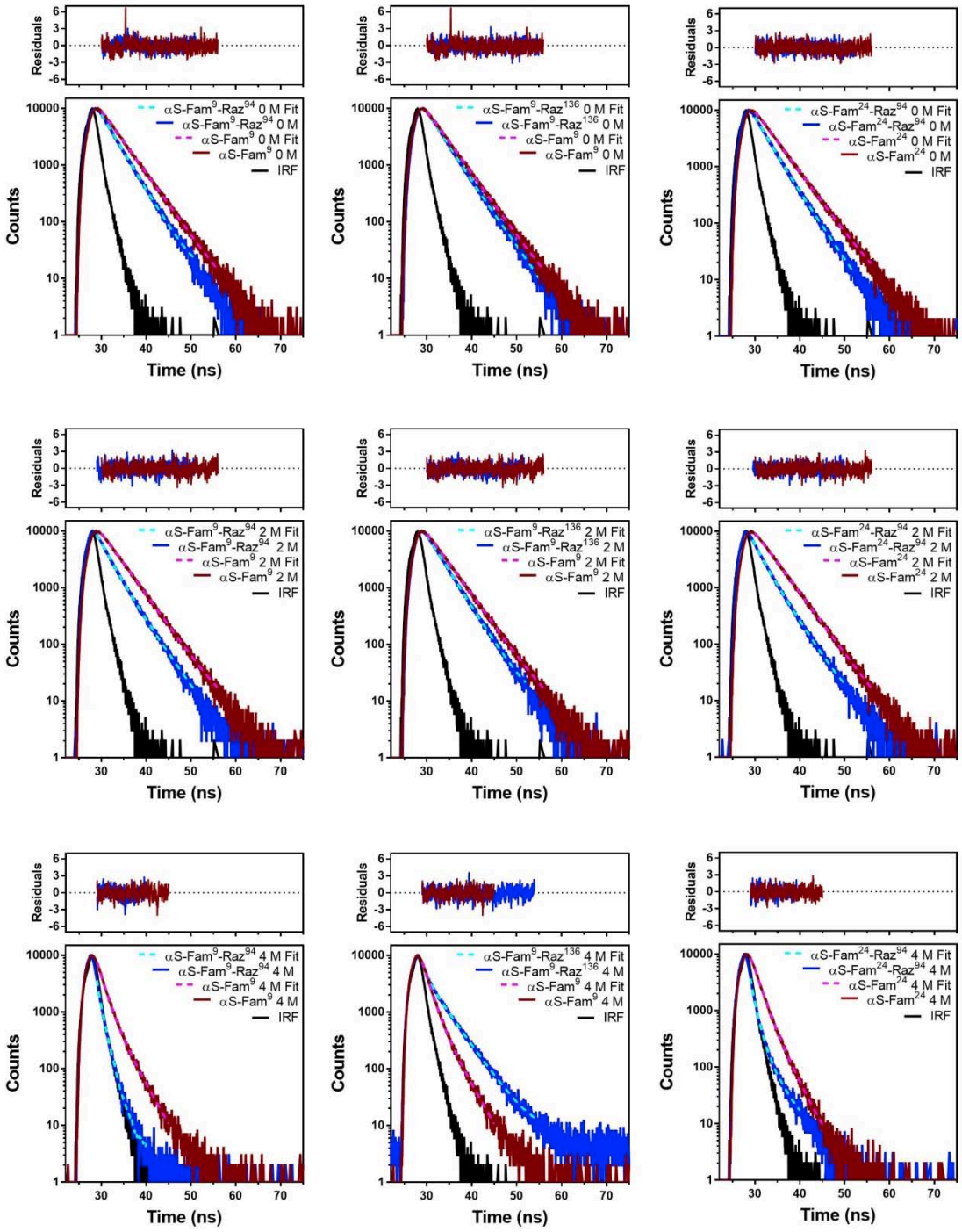

Figure S45: Intramolecular FRET Lifetime Decays and Fits. Plots show lifetime decays and fits of (Left) αS-Fam^{9}-Raz^{94}, (Middle) αS-Fam^{9}-Raz^{136} and (Right) αS-Fam^{24}-Raz^{94} in (Top-Bottom) 0, 2 and 4 M TMAO. Double-labeled constuct decays are shown in solid blue with associated single-labeled Fam constructs shown in red with fits shown as dashed lines.

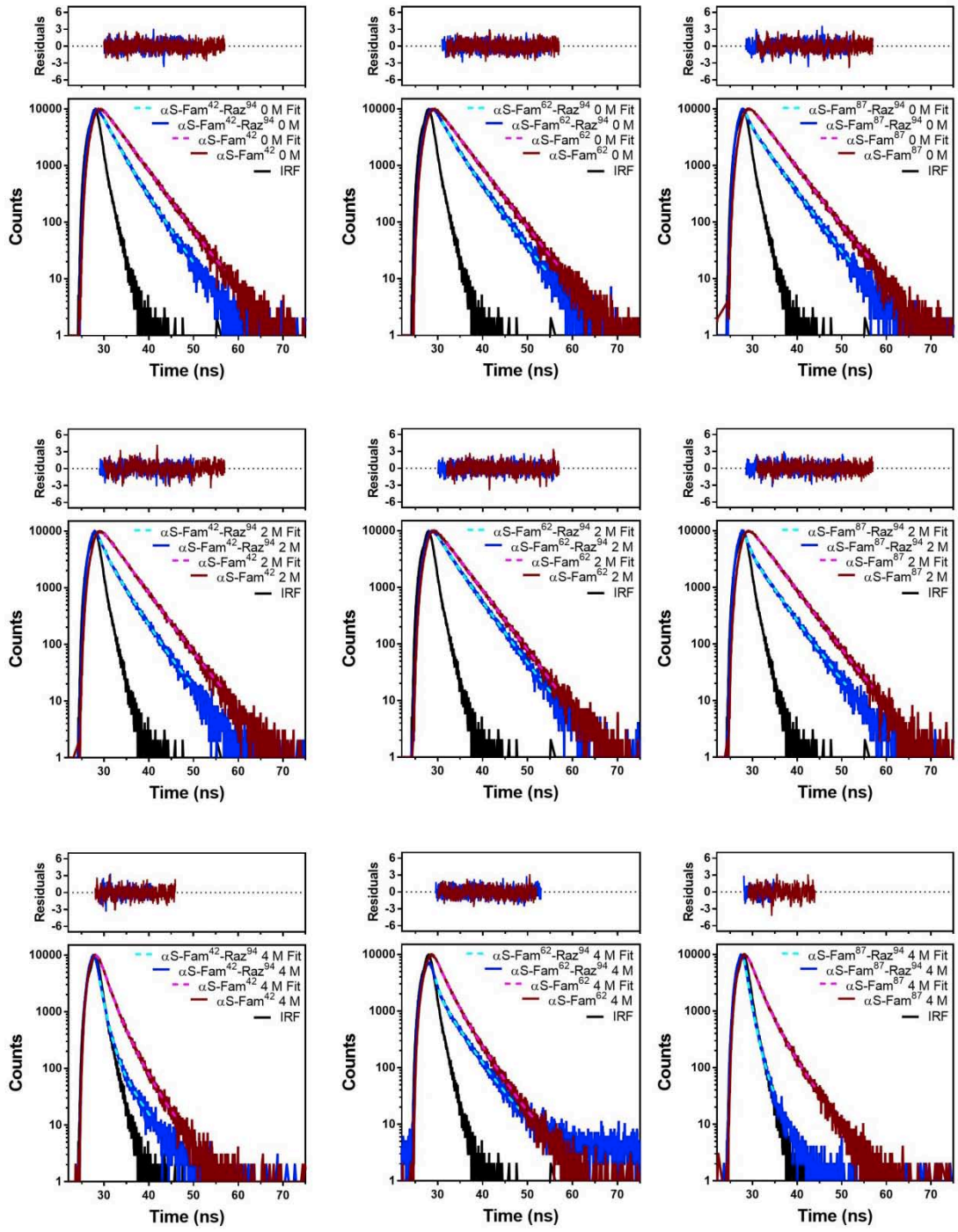

Figure S46: Intramolecular FRET Lifetime Decays and Fits. Plots show lifetime decays and fits of (Left) αS-Fam^{42}-Raz^{94}, (Middle) αS-Fam^{62}-Raz^{136} and (Right) αS-Fam^{87}-Raz^{94} in (Top-Bottom) 0, 2 and 4 M TMAO. Double-labeled constuct decays are shown in solid blue with associated single-labeled Fam constructs shown in red with fits shown as dashed lines.

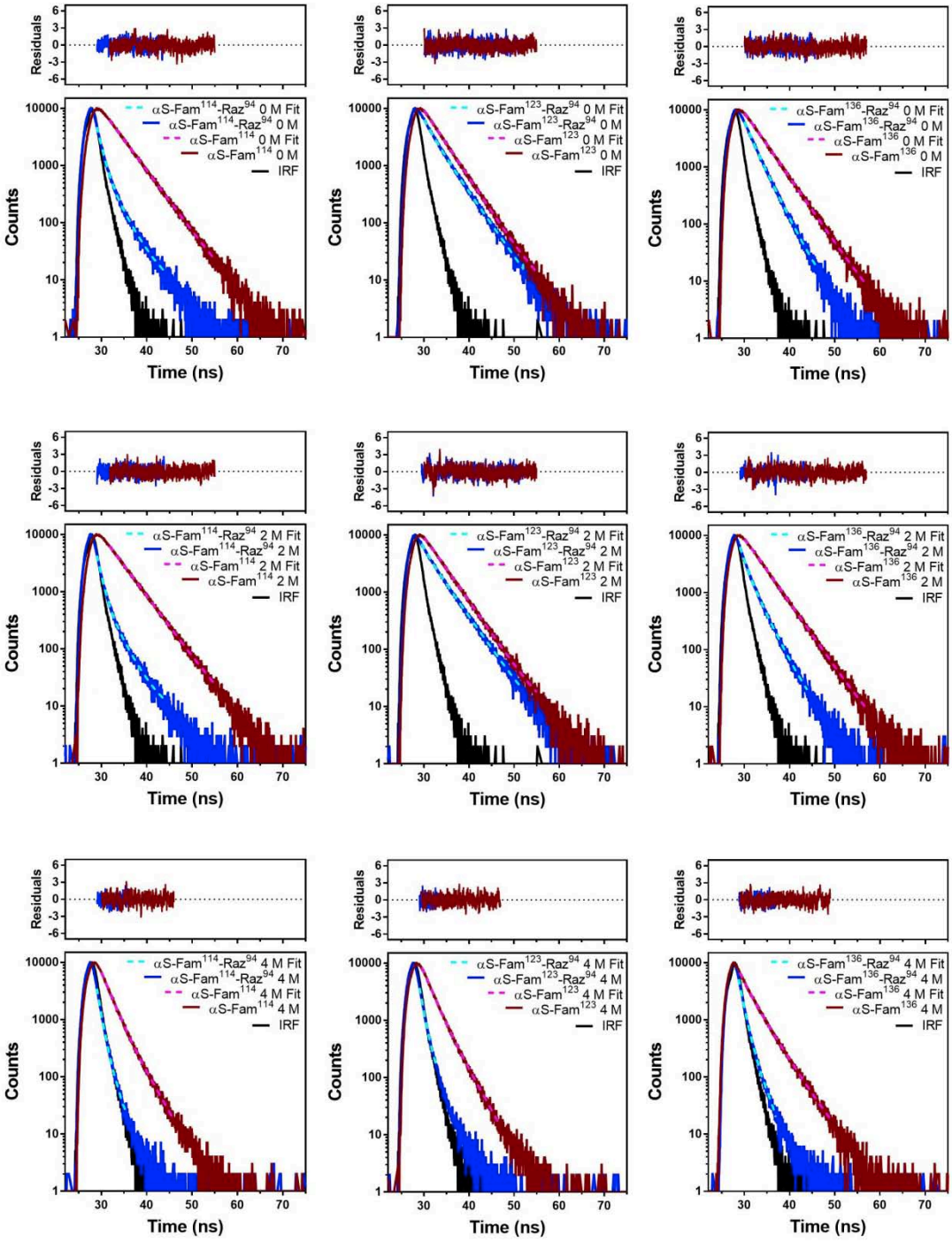

Figure S47: Intramolecular FRET Lifetime Decays and Fits. Plots show lifetime decays and fits of (Left) αS-Fam^{114}-Raz^{94}, (Middle) αS-Fam^{123}-Raz^{136} and (Right) αS-Fam^{136}-Raz^{94} in (Top-Bottom) 0, 2 and 4 M TMAO. Double-labeled constuct decays are shown in solid blue with associated single-labeled Fam constructs shown in red with fits shown as dashed lines.

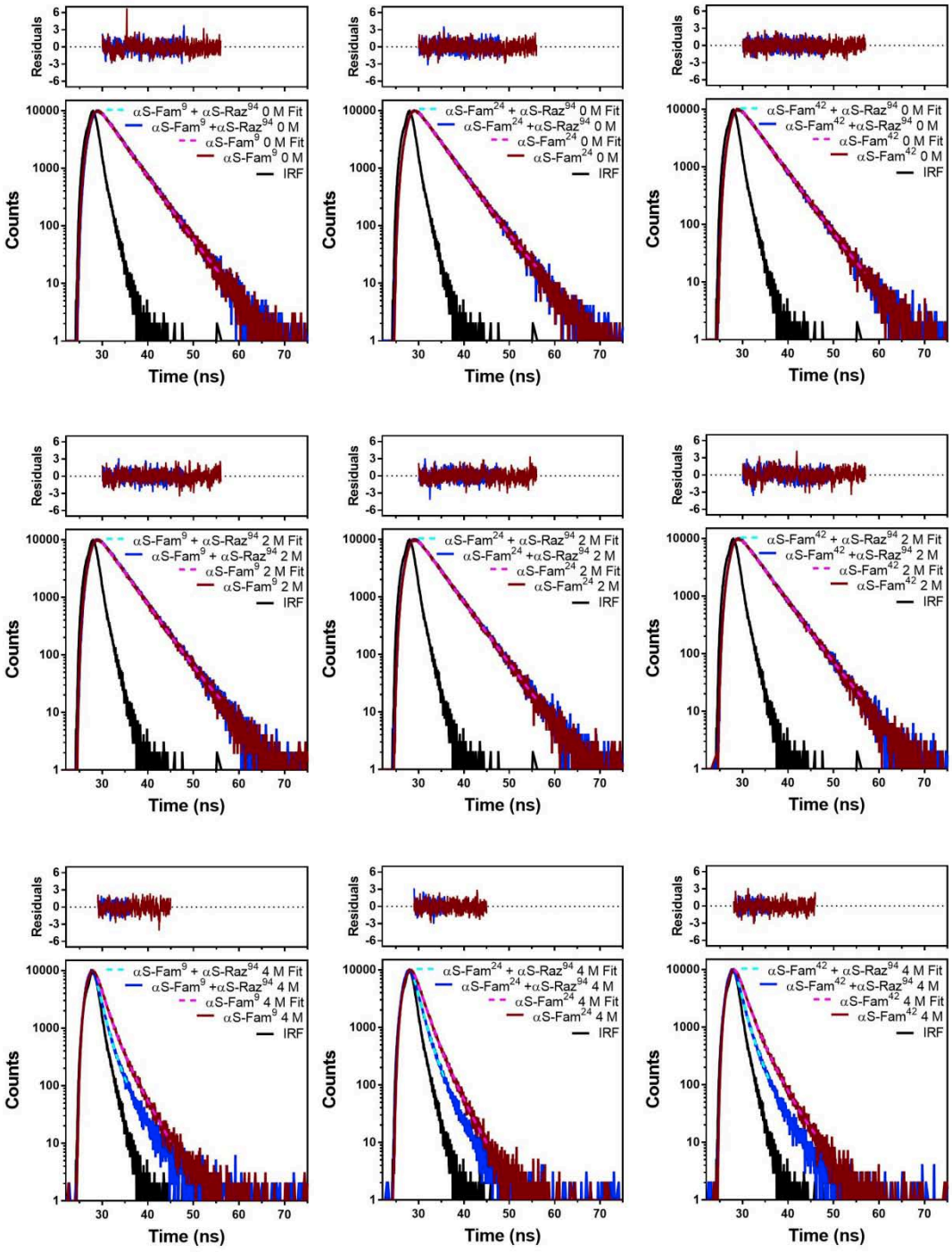

Figure S48: Intermolecular FRET Lifetime Decays and Fits. Plots show lifetime decays and fits of (Left) αS-Fam^{9} + αS-Raz^{94}, (Middle) αS-Fam^{9} + αS-Raz^{136} and (Right) αS-Fam^{24} + αS-Raz^{94} in (Top-Bottom) 0, 2 and 4 M TMAO. Decay of single-labeled Fam constuct decays are shown before (red) and after (blue) the addition of the Raz containing construct with fits shown as dashed lines.

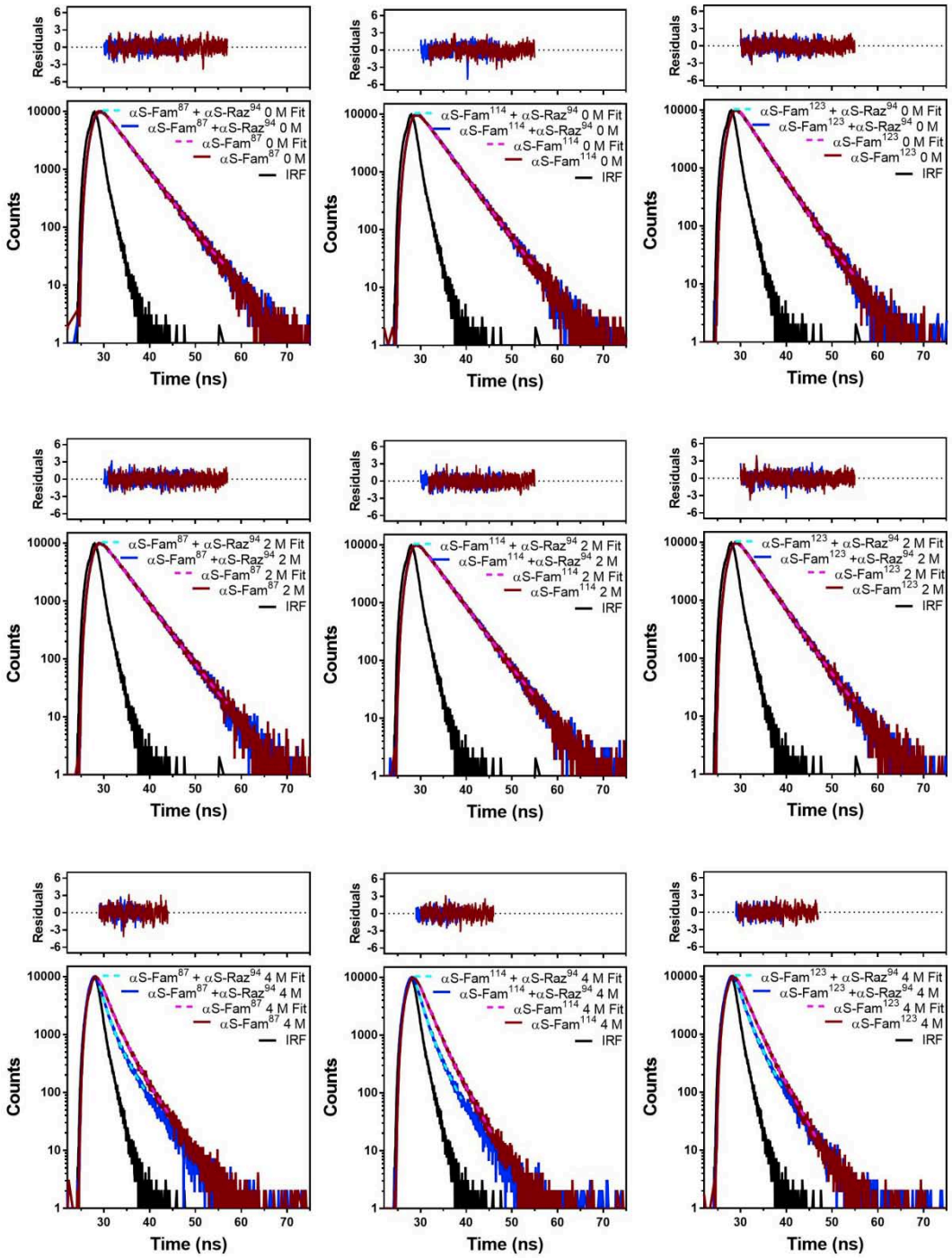

Figure S49: Intermolecular FRET Lifetime Decays and Fits. Plots show lifetime decays and fits of (Left) αS-Fam^{87} + αS-Raz^{94}, (Middle) αS-Fam^{114} + αS-Raz^{136} and (Right) αS-Fam^{123} + αS-Raz^{94} in (Top-Bottom) 0, 2 and 4 M TMAO. Decay of single-labeled Fam constuct decays are shown before (red) and after (blue) the addition of the Raz containing construct with fits shown as dashed lines.

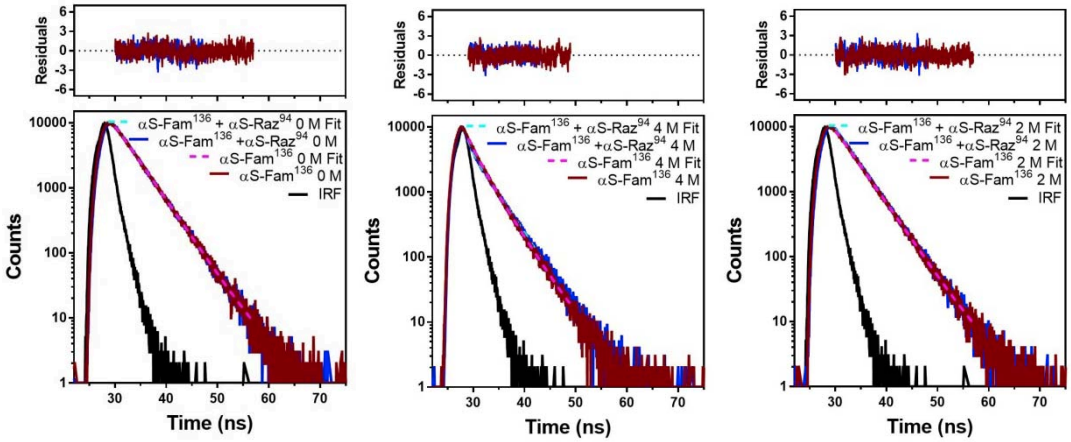

Figure S50: Intermolecular FRET Lifetime Decays and Fits. Plots show lifetime decays and fits of αS-Fam^{136} + αS-Raz^{94} in (Left-Right) 0, 2 and 4 M TMAO. Decay of single-labeled Fam constuct decays are shown before (red) and after (blue) the addition of the Raz containing construct with fits shown as dashed lines.

Table S4a. TCSPC Data and Fitting

Construct	TMAO	τ_1	τ_2	Amplitude$_1$	Amplitude$_2$	Intensity$_1$	Intensity$_2$	$\tau_{Amp.\,Avg.}$	$\tau_{Int.\,Avg.}$	χ^2
αS-Fam^{9}	0 M	3.97E+00 ± 5.02E-03	-	1.15E+00 ± 8.16E-03	-	1.00E+00	-	3.97E+00	3.97E+00	9.57E-01
	2 M	3.98E+00 ± 7.48E-03	-	7.66E-01 ± 3.70E-02	-	1.00E+00	-	3.98E+00	3.98E+00	9.40E-01
	4 M	3.03E+00 ± 6.06E-02	1.27E+00 ± 5.36E-02	2.79E-01 ± 2.29E-02	1.10E+00 ± 1.05E-01	2.01E-01	7.99E-01	1.63E+00	1.94E+00	8.64E-01
αS-Fam^{24}	0 M	3.99202 ± 5.72E-03	-	7.93E-01 ± 1.34E-02	-	1.00E+00	-	3.99202	3.99202	9.86E-01
	2 M	4.02857 ± 7.13E-03	-	7.77E-01 ± 2.35E-02	-	1.00E+00	-	4.02857	4.02857	1.02E+00
	4 M	2.68377 ± 8.74E-02	1.16144 ± 4.39E-02	1.77E-01 ± 1.88E-02	9.12E-01 ± 2.45E-02	3.10E-01	6.90E-01	1.40886	1.63276	9.64E-01
αS-Fam^{42}	0 M	4.03269 ± 7.48E-03	-	7.83E-01 ± 2.63E-02	-	1.00E+00	-	4.03269	4.03269	9.49E-01
	2 M	4.02E+00 ± 6.07E-03	-	8.06E-01 ± 2.58E-02	-	1.00E+00	-	4.02E+00	4.02E+00	1.1847
	4 M	2.88E+00 ± 5.03E-02	1.12E+00 ± 2.36E-02	1.88E-01 ± 1.01E-02	8.98E-01 ± 5.93E-03	3.50E-01	6.50E-01	1.42E+00	1.73E+00	1.02E+00
αS-Fam^{62}	0 M	4.14E+00 ± 6.45E-03	-	9.65E-01 ± 6.35E-02	-	1.00E+00	-	4.14E+00	4.14E+00	9.82E-01
	2 M	4.13E+00 ± 5.64E-03	-	9.77E-01 ± 7.47E-03	-	1.00E+00	-	4.13E+00	4.13E+00	1.10E+00
	4 M	3.54E+00 ± 5.82E-03	1.46E+00 ± 5.23E-03	3.33E-01 ± 1.50E-03	1.02E+00 ± 6.77E-03	4.43E-01	5.57E-01	1.97E+00	2.38E+00	9.18E-01
αS-Fam^{87}	0 M	4.16E+00 ± 8.86E-03	-	7.72E-01 ± 5.40E-02	-	1.00E+00	-	4.16E+00	4.16E+00	1.08E+00
	2 M	4.12E+00 ± 8.53E-03	-	7.21E-01 ± 3.60E-02	-	1.00E+00	-	4.12E+00	4.12E+00	8.96E-01
	4 M	3.07E+00 ± 7.69E-02	1.15E+00 ± 4.94E-02	2.36E-01 ± 1.47E-02	7.52E-01 ± 9.02E-03	4.57E-01	5.43E-01	1.61E+00	2.03E+00	1.30E+00
αS-Fam^{114}	0 M	4.04E+00 ± 6.09E-03	-	1.35E+00 ± 1.14E-02	-	1.00E+00	-	4.04E+00	4.04E+00	1.01E+00
	2 M	4.07E+00 ± 5.90E-03	-	1.52E+00 ± 3.87E-02	-	1.00E+00	-	4.07E+00	4.07E+00	9.03E-01
	4 M	3.95E+00 ± 5.31E-03	1.92E+00 ± 1.19E-03	1.21E-01 ± 4.47E-04	1.61E+00 ± 2.65E-03	1.34E-01	8.66E-01	2.06E+00	2.19E+00	1.05E+00
αS-Fam^{123}	0 M	3.64E+00 ± 4.64E-03	-	1.21E+00 ± 8.97E-03	-	1.00E+00	-	3.64E+00	3.64E+00	9.56E-01
	2 M	3.81E+00 ± 1.52E-03	-	8.70E-01 ± 6.82E-04	-	1.00E+00	-	3.81E+00	3.81E+00	1.02E+00
	4 M	3.15E+00 ± 1.26E-01	1.56E+00 ± 9.23E-02	2.39E-01 ± 3.34E-02	7.03E-01 ± 5.11E-02	4.06E-01	5.94E-01	1.96E+00	2.20E+00	9.78E-01
αS-Fam^{136}	0 M	3.77E+00 ± 7.07E-03	-	9.01E-01 ± 4.82E-02	-	1.00E+00	-	3.77E+00	3.77E+00	9.87E-01
	2 M	3.82E+00 ± 6.62E-03	-	8.28E-01 ± 3.28E-02	-	1.00E+00	-	3.82E+00	3.82E+00	1.08E+00
	4 M	3.51E+00 ± 4.72E-02	1.47E+00 ± 8.81E-02	4.82E-01 ± 4.91E-02	9.22E-01 ± 1.48E-01	5.55E-01	4.45E-01	2.17E+00	2.60E+00	9.44E-01
αS-Fam^{9}-Raz^{94}	0 M	3.25E+00 ± 5.53E-03	- -	2.62E+00 ± 2.91E-01	- ± -	1.00E+00	-	3.25E+00	3.25E+00	9.41E-01
	2 M	3.67E+00 ± 3.10E-02	2.03E+00 ± 4.24E-02	4.60E-01 ± 1.80E-02	7.19E-01 ± 2.91E-02	5.37E-01	4.63E-01	2.67E+00	2.91E+00	1.10E+00
	4 M	3.88E-01 ± 1.79E-02	- -	5.97E+00 ± 4.08E-01	- ± -	1.00E+00	-	3.88E-01	3.88E-01	1.19E+00
αS-Fam^{9}-Raz^{136}	0 M	4.57E+00 ± 1.44E-02	3.01E+00 ± 7.35E-03	1.58E-01 ± 1.55E-03	6.29E-01 ± 3.18E-03	2.75E-01	7.25E-01	3.32E+00	3.44E+00	9.50E-01
	2 M	3.41E+00 ± 2.47E-03	- -	3.40E+00 ± 9.08E-03	- ± -	1.00E+00	-	3.41E+00	3.41E+00	1.04E+00
	4 M	3.87E+00 ± 1.23E-02	5.41E-01 ± 4.71E-02	2.93E-01 ± 5.70E-03	3.06E+00 ± 5.15E-01	4.07E-01	5.93E-01	8.32E-01	1.90E+00	1.11E+00
αS-Fam^{24}-Raz^{94}	0 M	3.94E+00 ± 1.63E-02	2.51E+00 ± 1.82E-02	4.34E-01 ± 7.96E-03	9.94E-01 ± 1.65E-02	4.07E-01	5.93E-01	2.94E+00	3.09E+00	9.46E-01
	2 M	3.47E+00 ± 3.68E-02	1.34E+00 ± 8.49E-02	5.65E-01 ± 5.40E-02	9.89E-01 ± 1.58E-01	5.96E-01	4.04E-01	2.12E+00	2.61E+00	8.80E-01
	4 M	3.48E+00 ± 2.52E-01	4.27E-01 ± 3.18E-02	1.89E-02 ± 2.02E-03	3.46E+00 ± 4.48E-01	4.28E-02	9.57E-01	4.43E-01	5.57E-01	1.13E+00
αS-Fam^{42}-Raz^{94}	0 M	3.77E+00 ± 1.48E-02	2.09E+00 ± 1.84E-02	3.89E-01 ± 5.91E-03	7.53E-01 ± 1.31E-02	4.81E-01	5.19E-01	2.66E+00	2.90E+00	1.04E+00
	2 M	3.53E+00 ± 2.01E-02	9.99E-01 ± 4.10E-02	3.49E-01 ± 2.10E-02	9.01E-01 ± 1.07E-01	5.78E-01	4.22E-01	1.70E+00	2.46E+00	1.03E+00
	4 M	3.82E+00 ± 2.90E-01	5.50E-01 ± 1.54E-02	1.51E-02 ± 1.58E-03	2.17E+00 ± 1.39E-01	4.62E-02	9.54E-01	5.73E-01	7.02E-01	9.94E-01

Table S4b. TCSPC Data and Fitting

Construct	TMAO	τ_1	τ_2	Amplitude$_1$	Amplitude$_2$	Intensity$_1$	Intensity$_2$	$\tau_{Amp.\,Avg.}$	$\tau_{Int.\,Avg.}$	χ^2
αS-Fam^{62}-Raz^{94}	0 M	4.02E+00 ± 3.27E-02	1.71E+00 ± 6.02E-02	3.90E-01 ± 1.25E-02	6.57E-01 ± 3.40E-02	5.82E-01	4.18E-01	2.57E+00	3.06E+00	8.68E-01
	2 M	3.98E+00 ± 2.25E-02	1.57E+00 ± 8.34E-02	5.09E-01 ± 1.33E-02	4.21E-01 ± 1.95E-02	7.54E-01	2.46E-01	2.89E+00	3.39E+00	9.24E-01
	4 M	3.77E+00 ± 2.18E-02	2.35E-01 ± 3.29E-02	1.12E-01 ± 1.57E-03	3.47E+00 ± 5.95E-01	3.40E-01	6.60E-01	3.46E-01	1.44E+00	9.16E-01
αS-Fam^{87}-Raz^{94}	0 M	4.04E+00 ± 1.42E-02	7.74E-01 ± 2.87E-02	3.98E-01 ± 1.64E-02	1.40E+00 ± 1.90E-01	5.98E-01	4.02E-01	1.50E+00	2.73E+00	1.09E+00
	2 M	3.97E+00 ± 1.57E-02	4.83E-01 ± 3.72E-02	2.83E-01 ± 6.14E-03	2.51E+00 ± 3.68E-01	4.81E-01	5.19E-01	8.37E-01	2.16E+00	1.01E+00
	4 M	4.86E-01 ± 6.49E-03	-	3.04E+00 ± 5.30E-02	-	1.00E+00	-	4.86E-01	4.86E-01	1.06E+00
αS-Fam^{114}-Raz^{94}	0 M	3.25E+00 ± 1.88E-02	7.56E-01 ± 3.34E-03	6.22E-02 ± 7.98E-04	1.99E+00 ± 1.81E-02	1.19E-01	8.81E-01	8.31E-01	1.05E+00	1.01E+00
	2 M	3.14E+00 ± 7.13E-02	6.01E-01 ± 1.70E-02	4.98E-02 ± 2.27E-03	1.57E+00 ± 9.54E-02	1.43E-01	8.57E-01	6.79E-01	9.63E-01	9.38E-01
	4 M	4.95E-01 ± 1.90E-02	-	2.63E+00 ± 2.12E-01	-	1.00E+00	-	4.95E-01	4.95E-01	8.61E-01
αS-Fam^{123}-Raz^{94}	0 M	3.86E+00 ± 5.96E-02	1.81E+00 ± 1.04E-01	5.11E-01 ± 3.72E-02	7.49E-01 ± 7.82E-02	5.93E-01	4.07E-01	2.65E+00	3.03E+00	1.00E+00
	2 M	3.87E+00 ± 1.08E-02	1.32E+00 ± 3.04E-02	5.84E-01 ± 4.95E-03	8.56E-01 ± 3.90E-02	6.67E-01	3.33E-01	2.35E+00	3.02E+00	1.09E+00
	4 M	5.25E-01 ± 1.06E-02	-	2.18E+00 ± 1.13E-01	-	1.00E+00	-	5.25E-01	5.25E-01	9.64E-01
αS-Fam^{136}-Raz^{94}	0 M	2.78E+00 ± 1.00E-02	1.39E+00 ± 1.99E-02	5.25E-01 ± 6.82E-03	7.57E-01 ± 1.93E-02	5.81E-01	4.19E-01	1.96E+00	2.20E+00	1.02E+00
	2 M	2.75E+00 ± 4.57E-02	9.53E-01 ± 4.36E-02	2.63E-01 ± 1.92E-02	1.03E+00 ± 1.04E-01	4.24E-01	5.76E-01	1.32E+00	1.71E+00	1.05E+00
	4 M	3.67E+00 ± 9.63E-01	5.80E-01 ± 2.70E-02	1.52E-02 ± 4.99E-03	5.65E+00 ± 5.38E-01	1.68E-02	9.83E-01	5.89E-01	6.32E-01	8.18E-01
αS-Fam^{9}+	0 M	3.84E+00 ± 1.12E-02	-	7.97E-01 ± 3.15E-02	-	1.00E+00	-	3.84E+00	3.84E+00	1.06E+00
αS-Raz^{94}	2 M	3.89E+00 ± 1.08E-02	-	7.06E-01 ± 1.43E-02	-	1.00E+00	-	3.89E+00	3.89E+00	9.95E-01
	4 M	1.87E+00 ± 2.43E-01	6.12E-01 ± 5.06E-02	1.46E-01 ± 4.68E-02	1.75E+00 ± 1.05E-01	2.02E-01	7.98E-01	7.08E-01	8.66E-01	8.82E-01
αS-Fam^{24}+	0 M	3.96E+00 ± 1.12E-02	-	7.40E-01 ± 1.73E-02	-	1.00E+00	-	3.96E+00	3.96E+00	1.03E+00
αS-Raz^{94}	2 M	3.98E+00 ± 1.19E-02	-	6.94E-01 ± 1.78E-02	-	1.00E+00	-	3.98E+00	3.98E+00	1.05E+00
	4 M	1.09E+01 ± 3.81E+00	1.09E+00 ± 1.75E-02	2.78E-02 ± 2.99E-03	2.21E+00 ± 1.39E-01	1.12E-01	8.88E-01	1.21E+00	2.20E+00	1.06E+00
αS-Fam^{42}+	0 M	3.99E+00 ± 1.30E-02	-	7.16E-01 ± 2.35E-02	-	1.00E+00	-	3.99E+00	3.99E+00	8.63E-01
αS-Raz^{94}	2 M	3.96E+00 ± 8.80E-03	-	6.91E-01 ± 1.12E-02	-	1.00E+00	-	3.96E+00	3.96E+00	1.15E+00
	4 M	3.01E+00 ± 7.68E-03	7.91E-01 ± 9.01E-04	7.68E-02 ± 3.12E-04	1.37E+00 ± 2.59E-03	1.76E-01	8.24E-01	9.08E-01	1.18E+00	9.02E-01
αS-Fam^{62}+	0 M	4.15E+00 ± 1.32E-02	-	6.91E-01 ± 1.50E-02	-	1.00E+00	-	4.15E+00	4.15E+00	9.06E-01
αS-Raz$_{94}$	2 M	4.09E+00 ± 1.01E-02	-	6.70E-01 ± 9.15E-03	-	1.00E+00	-	4.09E+00	4.09E+00	9.62E-01
	4 M	2.84E+00 ± 4.19E-03	7.47E-01 ± 1.18E-03	1.55E-01 ± 3.86E-04	1.22E+00 ± 3.23E-03	3.26E-01	6.74E-01	9.84E-01	1.43E+00	1.06E+00
αS-Fam^{114}+	0 M	4.00E+00 ± 1.26E-02	-	7.22E-01 ± 1.72E-02	-	1.00E+00	-	4.00E+00	4.00E+00	1.11E+00
αS-Raz$_{94}$	2 M	4.03E+00 ± 1.43E-02	-	7.52E-01 ± 2.54E-02	-	1.00E+00	-	4.03E+00	4.03E+00	9.12E-01
	4 M	2.70E+00 ± 2.64E-02	8.19E-01 ± 9.15E-03	1.46E-01 ± 2.92E-03	1.12E+00 ± 1.48E-02	3.01E-01	6.99E-01	1.04E+00	1.39E+00	8.68E-01
αS-Fam^{123}+	0 M	3.58E+00 ± 1.07E-02	-	8.99E-01 ± 3.99E-02	-	1.00E+00	-	3.58E+00	3.58E+00	1.02E+00
αS-Raz^{94}	2 M	3.78E+00 ± 1.13E-02	-	8.35E-01 ± 3.10E-02	-	1.00E+00	-	3.78E+00	3.78E+00	1.07E+00
	4 M	2.60E+00 ± 1.18E-01	7.56E-01 ± 6.47E-02	2.23E-01 ± 1.88E-02	9.39E-01 ± 2.30E-02	4.50E-01	5.50E-01	1.11E+00	1.59E+00	9.34E-01
αS-Fam^{136}+	0 M	3.72E+00 ± 9.63E-03	-	8.29E-01 ± 2.93E-02	-	1.00E+00	-	3.72E+00	3.72E+00	9.76E-01
αS-Raz^{94}	2 M	3.79E+00 ± 1.18E-02	-	9.81E-01 ± 6.79E-02	-	1.00E+00	-	3.79E+00	3.79E+00	1.05E+00
	4 M	3.79E+00 ± 1.00E-03	7.55E-01 ± 5.69E-04	2.70E-01 ± 1.01E-04	8.85E-01 ± 1.25E-03	6.05E-01	3.95E-01	1.46E+00	2.59E+00	9.35E-01

Table S5. Calculated E_{FRET} values from TCSPC

Intramolecular				Intermolecular			
Construct	TMAO	$EFRET_{Amp}$	$EFRET_{Int}$	Construct	TMAO	$EFRET_{Amp}$	$EFRET_{Int}$
αS-Fam^{9}-Raz^{94}	0 M	0.18	0.18	αS-Fam^{9} +	0 M	0.03	0.03
	2 M	0.33	0.27	αS-Raz^{94}	2 M	0.02	0.02
	4 M	0.45	0.55		4 M	0.57	0.55
αS-Fam^{9}-Raz^{136}	0 M	0.16	0.13	αS-Fam^{24} +	0 M	0.01	0.01
	2 M	0.14	0.14	αS-Raz^{94}	2 M	0.01	0.01
	4 M	0.49	0.02		4 M	0.14	-0.34
αS-Fam^{24}-Raz^{94}	0 M	0.26	0.23	αS-Fam$_{42}$ +	0 M	0.01	0.01
	2 M	0.47	0.35	αS-Raz^{94}	2 M	0.01	0.01
	4 M	0.69	0.66		4 M	0.36	0.32
αS-Fam^{42}-Raz^{94}	0 M	0.34	0.28	αS-Fam$_{62}$ +	0 M	0.00	0.00
	2 M	0.58	0.39	αS-Raz^{94}	2 M	0.01	0.01
	4 M	0.60	0.60		4 M	0.39	0.29
αS-Fam^{62}-Raz^{94}	0 M	0.38	0.26	αS-Fam$_{114}$ +	0 M	0.01	0.01
	2 M	0.30	0.18	αS-Raz^{94}	2 M	0.01	0.01
	4 M	0.82	0.40		4 M	0.50	0.37
αS-Fam^{87}-Raz^{94}	0 M	0.64	0.34	αS-Fam$_{123}$ +	0 M	0.02	0.02
	2 M	0.80	0.48	αS-Raz^{94}	2 M	0.01	0.01
	4 M	0.70	0.76		4 M	0.43	0.28
αS-Fam^{114}-Raz^{94}	0 M	0.79	0.74	αS-Fam$_{136}$ +	0 M	0.01	0.01
	2 M	0.83	0.76	αS-Raz^{94}	2 M	0.01	0.01
	4 M	0.76	0.77		4 M	0.32	0.00
αS-Fam^{123}-Raz^{94}	0 M	0.27	0.17				
	2 M	0.38	0.21				
	4 M	0.73	0.76				
αS-Fam^{136}-Raz^{94}	0 M	0.48	0.42				
	2 M	0.65	0.55				
	4 M	0.73	0.76				

FCS and AFM Characterization

FCS Measurements: FCS measurements were done at 20°C on a lab-built instrument based on an Olympus IX71 microscope. A continuous emission 488 nm DPSS 50 mW laser (Spectra-Physics, Santa Clara, CA) was adjusted to 4.5 μW power just prior to entering the microscope. Fluorescence was collected through the objective and separated from the excitation laser using a Z488rdc long pass dichroic and an HQ600/200m bandpass filter (Chroma, Bellows Falls, VT). Fluorescence was focused onto the aperture of a 50 μm optical fiber coupled to an avalanche photodiode (Perkin Elmer, Waltham, MA). 10 autocorrelation curves of 10 seconds each were taken using a digital correlator (Flex03Q-12, correlator.com, Bridgewater, NJ). Fitting was done using MATLAB (The MathWorks, Natick, MA).

Eight-well chambered coverglasses (Nunc, Rochester, NY) were cleaned by plasma treatment followed by incubation with polylysine-conjugated polyethylene glycol (PEG-PLL). PEG-PLL was prepared from a modified Pierce PEGylation protocol (Pierce, Rockford, IL). After overnight incubation with PEG-PLL, chambers were rinsed with Millipore water and stored until use for measurements. Measurements were done by first replacing the water with solutions of ~400 μM wild type α-synuclein in each concentration of TMAO to prevent adsorption of labeled protein. 20 nM of Alexa Fluor 488 or labeled α-synuclein were added into a chamber, and each sample was mixed by pipetting immediately before measurement. The free dye measurements were used to obtain reference diffusion times.

The observed fluorescence fluctuations were autocorrelated and the raw autocorrelation data was fit to the equation:

$$G(\tau) = \frac{1}{N}\left(\frac{1}{1+\frac{\tau}{\tau_{\alpha S}}} \times \left(\frac{1}{1+\frac{s^2\tau}{\tau_{\alpha S}}} \right)^{1/2} \right) \qquad \text{(Eq. S15)}$$

N is the number of molecules in the focal volume, τ is the delay time, $\tau_{\alpha S}$ is the time spent by the sample in the focal volume, and s represents the eccentricity of the confocal volume, and is fixed to 0.2. The diffusion time of the protein, $\tau_{\alpha S}$, is extracted from fit.

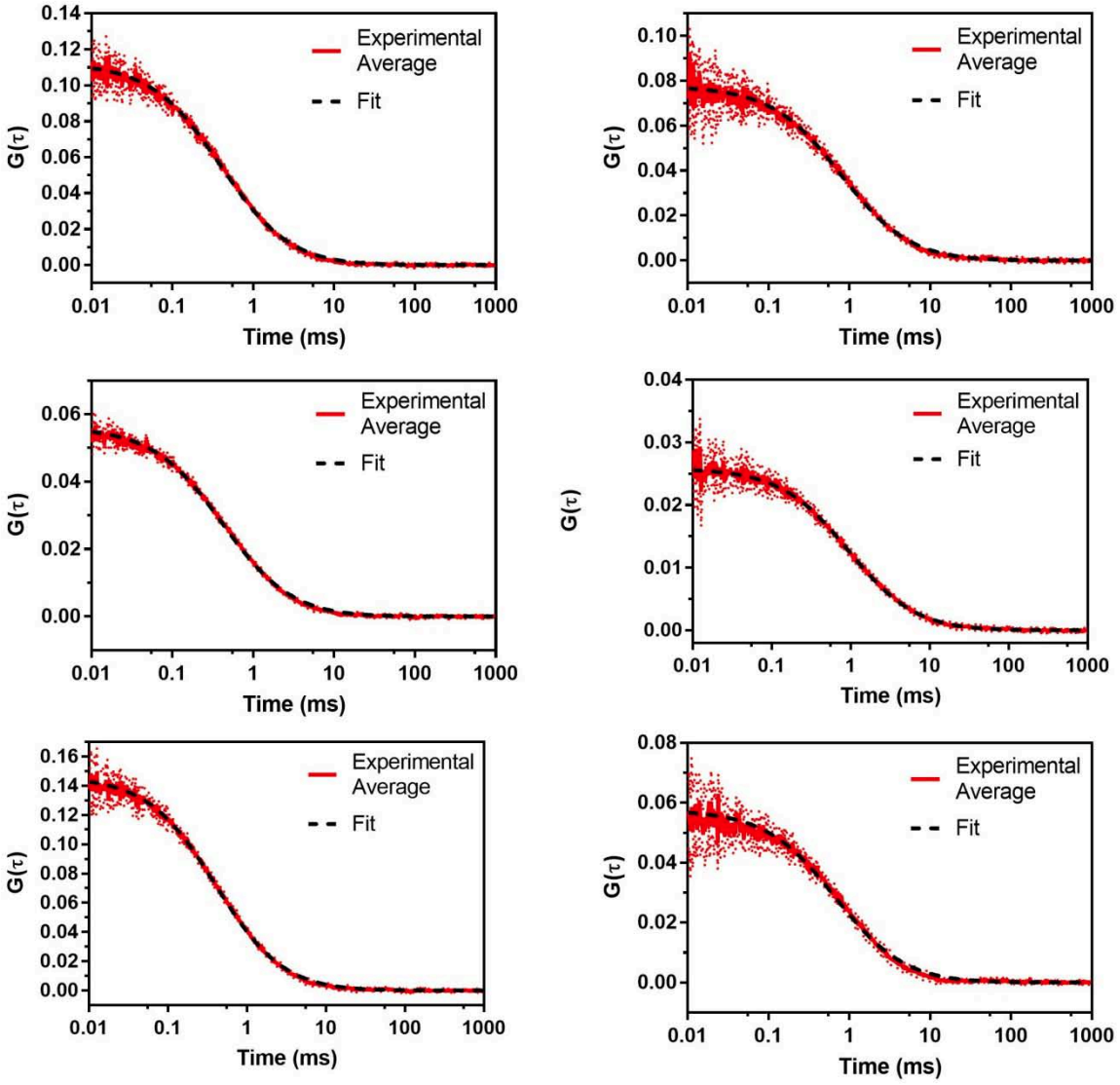

Figure S51: Autocorrelation plots from FCS Data. Plots show the average autocorrelation decays from 10 measurements (red) with fits (black-dashed) for αS-AF_{488}^{9} (Top), αS-AF_{488}^{114} (Middle) and αS-AF_{488}^{130} (Left) in 0 M (Left) and 2 M (Right) TMAO.

Table S6. FCS Data

Position (TMAO)	**9 (0 M)**	**114 (0 M)**	**130 (0 M)**
N	8.9006	17.809	6.8298
Delay Time	0.401	0.42034	0.40662
Intensity	59007	193000	51376
ταS	0.43±0.027	0.418±0.010	0.411±0.010
Position (TMAO)	9 (2 M)	114 (2 M)	130 (2 M)
N	12.889	38.693	17.384
Delay Time	0.79827	0.94794	0.68422
Intensity	44824	163510	54824
ταS	0.822±±0.007	0.946±0.035	0.779±0.037

AFM Measurements: Three droplets of αS solution were incubated on the mica substrate for 5 minutes. To remove the excess solution, the samples were blotted by Kimwipe, gently rinsed with 2 mL Milli-Q water, and then dried using a weak nitrogen stream. During the sample preparation, the substrate was tilted 30-45° to prevent the solution from flowing backwards. The samples were imaged in tapping mode using a Keysight 5500 AFM instrument (Keysight Technologies) equipped with a closed-loop scanner. Rotated silicon probes with aluminum reflex coating (BudgetSensors, Tap-300G, resonance frequency ~300 kHz, tip radius <10 nm, force constant 40 N/m) were used to record topographic, amplitude, and phase images with 512 x 512 pixel resolution. The AFM images were analyzed by Gwyddion package. A third-order polynomial was used to flatten the background for topographic images. The volumetric analyses of globular structures were performed using Gwyddion software.

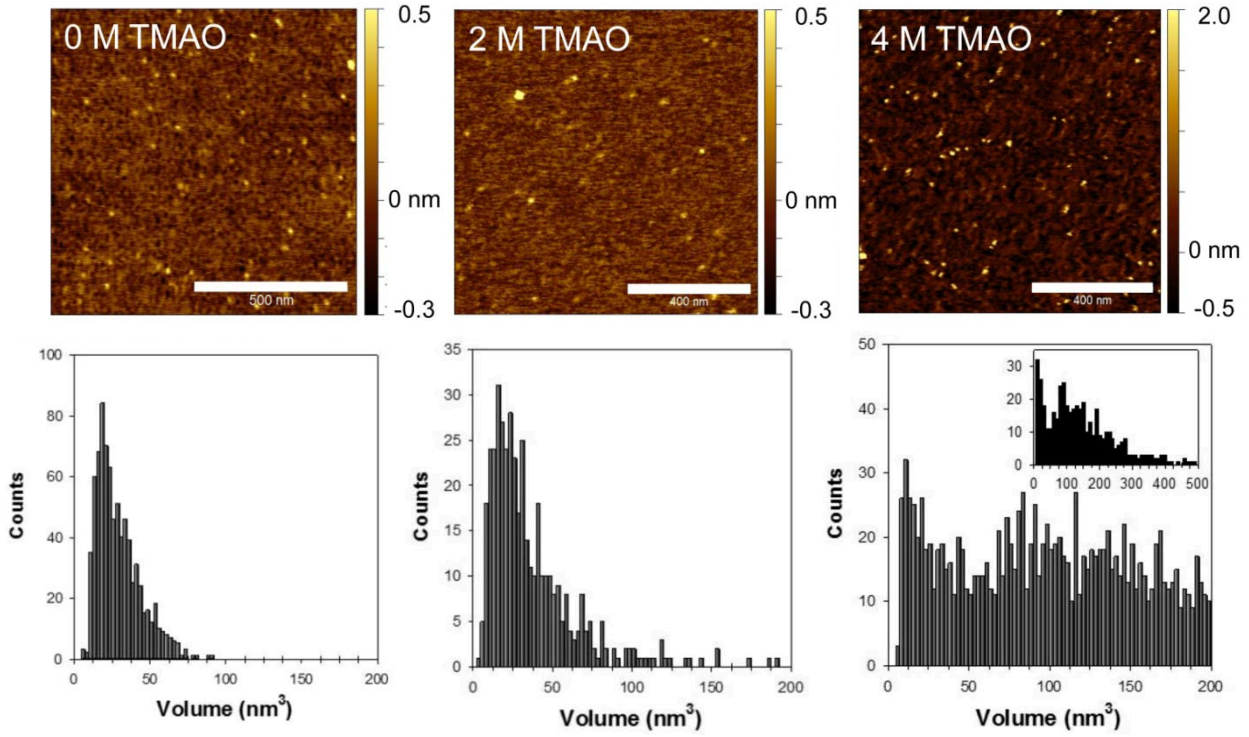

Figure S52: The effect of TMAO on the aggregation of wild type αS. Top: AFM images of the dried samples. Bottom: corresponding volume histograms of the αS aggregates. The αS concentration used for these samples are (left) 10 μM, (middle) 0.1 μM, and (right) 0.5 μM. Each volume histogram was obtained with accumulated data from more than five AFM images screened at various area on the sample. No αS fibrillization is observed, which is consistent with the fact that αS fibrillization in solution can't occur in low concentrations and under static conditions, and no agitation or shacking of the solution were performed here.

Computational Modeling Procedures

Semi-Empirical Probability Distribution: This simulation was run utilizing the same format as the 0 M TMAO unconstrained simulation where the score function at each step was replaced simply with a single repulsive van der Waals term. For the combined set of outputs, all segments which are spaced by the same number of residues in the primary sequence were fit to a normal distribution. From this an empirical relationship was constructed between the distance spanning any two parts of the protein and the average breadth, σ.

PyRosetta Simulations: All simulations were performed in PyRosetta on the University of Pennsylvania School of Arts and Science General Purpose Cluster. The simulation format followed a basic simulated annealing procedure detailed in the general script attached. In short, the initial protein structure was randomized by sampling random backbone ϕ/ψ angles against a score function consisting solely of repulsive van der Waals energy terms while the protein was represented in the course-grained centroid model. Following randomization, constraints were applied with continued sampling of ϕ/ψ angles in centroid model while increasing the score function complexity from score0-score3 ending with the score3 function as well as long and short range hydrogen bonding terms (hbond_srbb and hbond_lrbb). Lastly, the protein representation was switched into full-atom and backbone torsion angles were sampled along with side-chain χ angles while the 'beta' score function was applied with the addition of constraints. Acceptance for a set of moves was determined as usual in a Monte Carlo simulation by comparing the sum of the score energy and constraint energy of the previous structure to the newly generated structure. The lowest energy structure generated from each of these steps was retained and used as the starting structure for the next simulation step. The final output structure as determined from the sum of the total 'beta' score function energy and the constraint energy.

Several different methods were employed in an attempt to capture the influence of TMAO on the protein structure in addition to applying constraints. Simulations were performed where the solvation term (fa_sol) was removed from the score function since solvation effects from TMAO introduction could not be accounted for directly. Moreover, fragment insertions was employed alongside ϕ/ψ angle sampling in an effort to incorporate the possible formation of significant secondary structure. Fragment libraries were prepared from the primary sequence of wild-type αS using Robetta, a protein prediction server which uses Rosetta software and was developed and

supported by the Rosetta Commons and the Baker Lab (9). Modifications to the base method for each simulation for the 2 M TMAO data set are detailed with the reported results.

In order to produce a simulation which accurately represented the FRET data, each distance constraint was implemented with a "knowledge" weighting factor, γ. This knowledge weighting factor, γ, is inversely proportional to the standard deviation in the calculated distance, based on a given version of the polymer-scaled Förster equation.

$$\gamma\left(E_{FRET}\right)=\frac{N}{S_{DA}\times\delta x\left(E_{FRET}\right)/\delta E_{FRET}} \qquad \text{(Eq. S16)}$$

$$N=\left(\int_{E_{FRET}}\frac{1}{S_{DA}\times\delta x\left(E_{FRET}\right)/\delta E_{FRET}}d\left(E_{FRET}\right)\right)^{-1} \qquad \text{(Eq. S17)}$$

Constraints were applied using one of two functions based on the probability distribution functions used to interpret the FRET data. The Gaussian chain (Eq. 3) and normal distribution (Eq. 4) functions were transformed into Eq. S18 and S19 respectively to generate potential energy functions termed the Gaussian chain constraint and the harmonic constraint functions respectively.

$$U_1\left(r\right)=a\times\gamma\times\log\left(4\pi r\left(\frac{3}{2\pi\left\langle r^2\right\rangle}\right)^{3/2}\exp\left(-\frac{3}{2}\frac{r^2}{\left\langle r^2\right\rangle}\right)\right) \qquad \text{(Eq. S18)}$$

$$U_2\left(r\right)=a\times\gamma\times\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\frac{\left(r-\left\langle r^2\right\rangle\right)^2}{\sigma^2}\right)\right) \qquad \text{(Eq. S19)}$$

The additional term, a, in both equations represents the “thermodynamic” weighting factor of the constraints with respect to the Rosetta score function. The optimal thermodynamic weighting factor was determined empirically by running trial simulations with a set to values from 0.25-5. The breadth, σ for the harmonic restraints (Eq. S19) was determined from a relationship between the average distance and the normal distribution breadth derived from the simulated structures used to generate the semi-empirical constraints. The optimal weight (a) was determined to be 1,

since at that value constraints were satisfied but the resulting ensembles were not overconstrained so as to yield physically unreasonable conformations. For each interpretation of the data, both constraint functions were applied to determine the efficacy of the constraining method and the data interpretation. Simulations assessing the constraint weight and the method of constraining generated 120 outputs, where the lowest 100 structures were used in analysis due to the lack of differences in energy in the outputs. Final simulations of the Gaussian chain interpreted and Gaussian chain constrained data produced 1020 structures where the lowest 1000 were used for interpretation.

Simulation Output Analysis: HydroPro was used to calculate the diffusion coefficients and radii of gyration. For analysis of 2 M simulated structures in HydroPro, the input viscosity was multiplied by the ratio of the 0 M:2 M diffusion times of AlexaFluor488. This effectively accounted for the increase in viscosity in 2 M versus 0 M TMAO. Following determination in HydroPro, the diffusion coefficient was converted for comparison with the diffusion coefficient determined from FCS using a previously published conversion equation detailed in Eq. S20, while the diffusion coefficient was determined from the measured diffusion time via Eq. S21 (10).

$$D_{Calibrated} = \frac{D_{HydroPro} + 0.582}{1.08} \qquad \text{(Eq. S20)}$$

$$D_P = D_{AF488} \frac{\tau_{AF488}}{\tau_P} \qquad \text{(Eq. S21)}$$

In these equations, D_P, $D_{AF488,}$ $D_{HydroPro}$ and $D_{Calibrated}$ are the diffusion coefficients of the protein of interest, AlexaFluor488, as determined from literature, from HydroPro and the experimentally calibrated diffusion coefficient for comparison to D_P respectively. Additionally, τ_{AF488} and τ_P are the experimentally derived diffusion times of AlexaFluor488 and the protein of interest respectively.

All distances required for Flory scaling plots, heat maps and comparisons to experimental FRET data were extracted from C-alpha to C-alpha distances for each residue using Python or BioPython and all calculations were performed using NumPy. For E_{FRET} determination, distance values from each member of a given ensemble were transformed into E_{FRET} values using the classical Förster equation prior to averaging to capture the explicit distance probability distribution

of the ensemble. The same method was used to calculate comparisons to paramagnetic relaxation enhancement data, where distances were extracted from C-alpha to amide proton distance. Distances converted into intensity ratios as described by Piana *et. al.* and the intensity ratios were averaged to calculate the final values (11).

Impact of the Thermodynamic Constraint Weight: Assessment of the thermodynamic constraint weight effects were assessed prior to application of the knowledge constraint weight. This was done in order to assess the maximum impact of the constraint set. Knowledge constraints applied to the full simulation were scaled to a maximum value determined by the thermodynamic constraint as shown in Eq. S18 and Eq. S19, which is why determination of the maximum impact of the constraint set was crucial. Experimentally constrained simulations were run as described below with the omission of side chain rotamer packing using constraints where the distances were determined by using the Gaussian chain version of the polymer scaled Förster equation and were applied using the Gaussian chain derived constraining function.

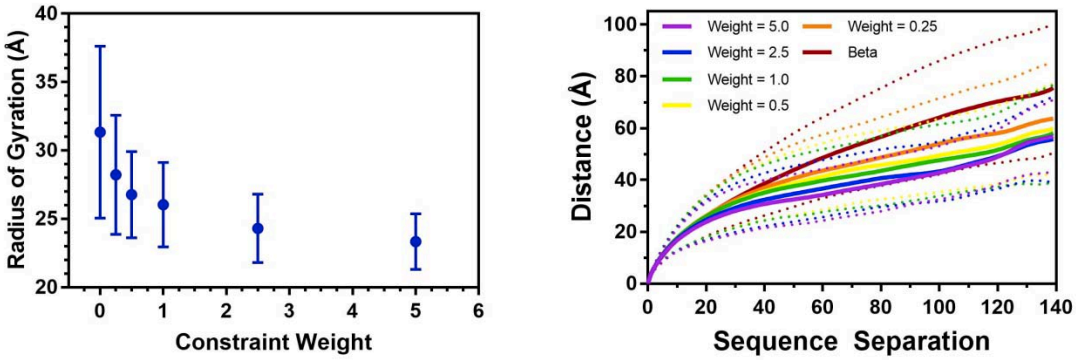

Figure S53: Determination of Optimal Thermodynamic Constraint Value. Radius of gyration (left) and average Flory scaling behavior (right) of simulated structures with varying values of the thermodynamic constraint weight, *a*.

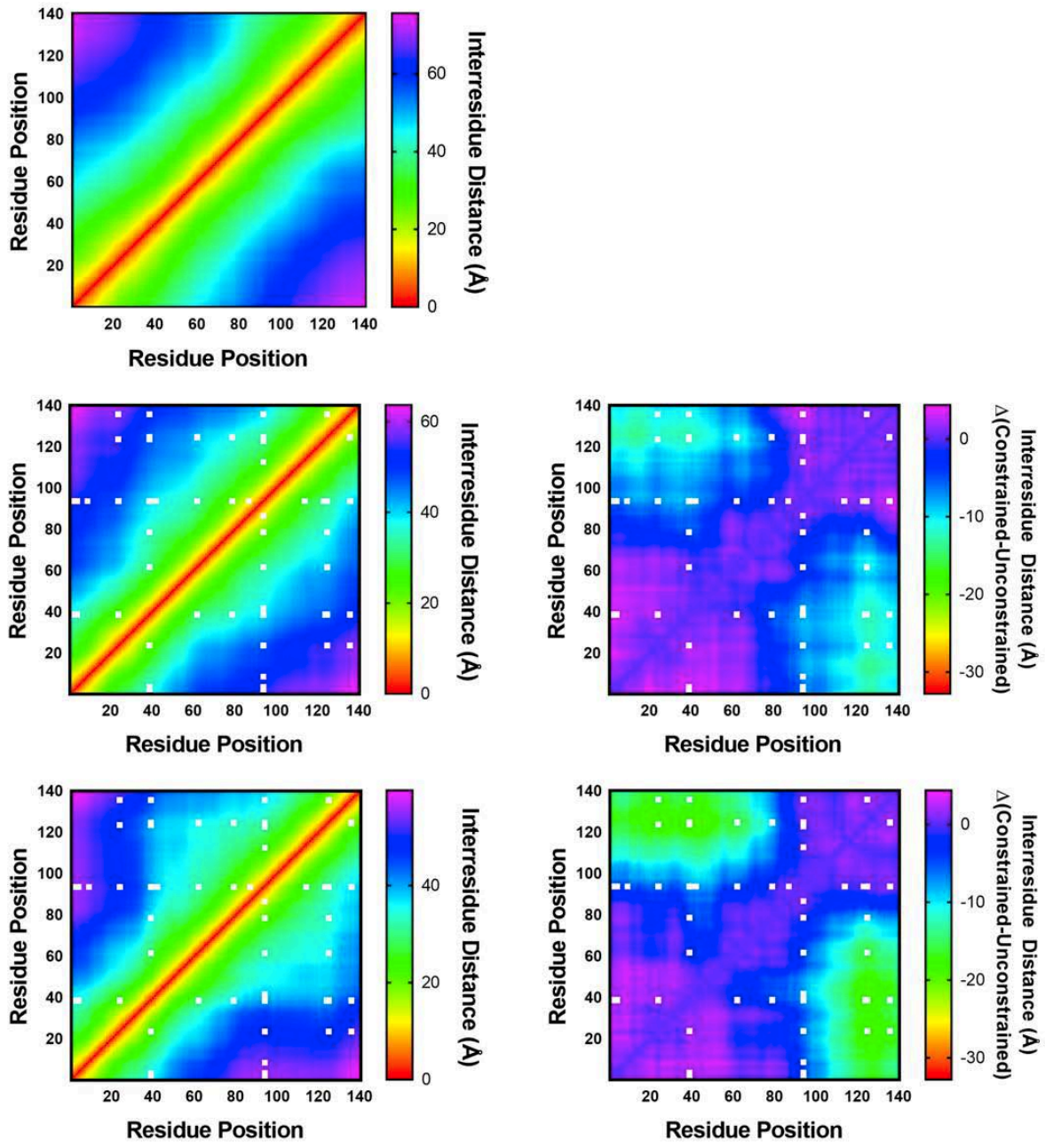

Figure S54: Heat maps depicting the average interresidue distance for each thermodynamic constraint weight, *a*, (left) and average difference in interresidue distance between constrained and unconstrained simulations (right). Heat maps depict averages from an unconstrained simulation (top row) and simulations with thermodynamic constraint weights set to 0.25 (middle row) and 0.5 (bottom row).

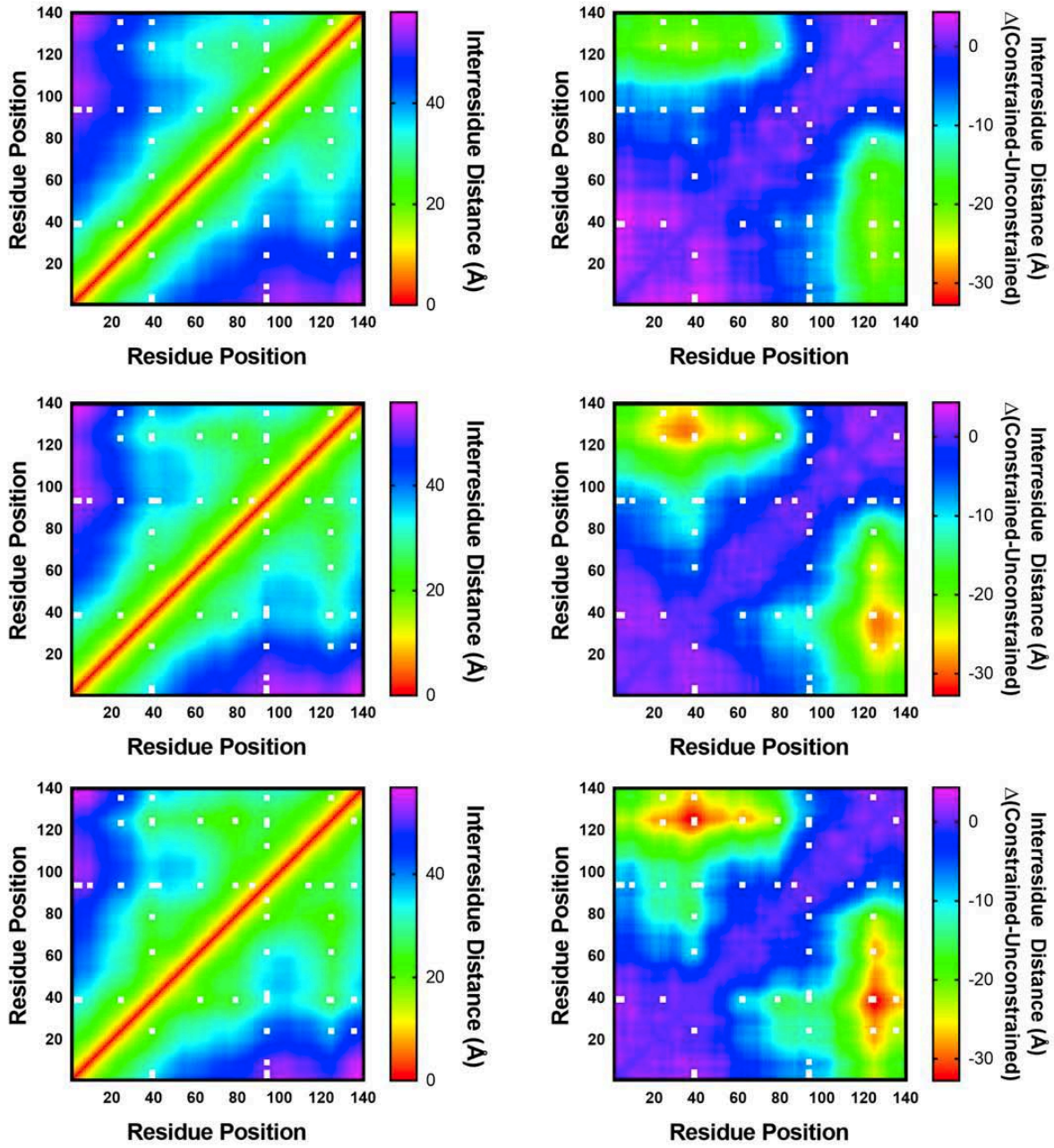

Figure S55: Heat maps depicting the average interresidue distance for each thermodynamic constraint weight, *a*, (left) and average difference in interresidue distance between constrained and unconstrained simulations (right). Heat maps depict averages from simulations with thermodynamic constraint weights set to 1.0 (top row), 2.5 (middle row) and 5.0 (bottom row).

Constraints for Utilized each Constraining Method:

0 M Gaussian-Chain Distance - Gaussian Chain Constraint

Res 1	Res 2	r	γ
39	24	39.061	0.25
39	62	36.48	0.308
39	79	52.301	0.097
39	94	44.112	0.171
94	24	74.119	0.025
94	39	44.823	0.146
94	62	34.425	0.332
94	79	32.21	0.401
125	24	48.945	0.084
125	39	45.722	0.106
125	62	42.349	0.137
125	79	46.02	0.104
125	94	39.257	0.175
136	24	66.266	0.045
136	39	47.473	0.141
136	94	45.153	0.167
136	125	46.607	0.15
9	94	69.478	1.105
24	94	60.531	1.393
42	94	51.785	1.573
62	94	48.25	1.619
87	94	34.824	1.335
114	94	25.412	0.588
123	94	45.228	1.625
136	94	39.352	1.512
9	136	142.742	0.184

0 M Semi-Empirical Distance - Harmonic Constraint

Res 1	Res 2	r	γ	σ
39	24	26.436	0.315	5.382
39	62	25.231	0.385	5.035
39	79	32.576	0.122	7.206
39	94	28.788	0.218	6.075
94	24	42.495	0.03	10.267
94	39	28.73	0.188	6.058
94	62	23.936	0.416	4.661
94	79	22.903	0.502	4.367
125	24	29.619	0.111	6.32
125	39	28.202	0.138	5.902
125	62	26.699	0.178	5.458
125	79	28.334	0.136	5.941
125	94	25.322	0.227	5.061
136	24	39.473	0.055	9.318
136	39	30.523	0.179	6.588
136	94	29.452	0.212	6.271
136	125	30.123	0.191	6.469
9	94	59.77	1.261	16.254
24	94	54.252	1.574	14.318
42	94	48.236	1.755	12.189
62	94	45.915	1.77	11.392
87	94	36.143	1.291	8.295
114	94	28.375	0.474	5.953
123	94	43.364	1.763	10.544
136	94	39.686	1.542	9.384
9	136	108.103	0.124	33.291

0 M Gaussian Chain Distance - Harmonic Constraint

Res 1	Res 2	r	γ	σ
39	24	39.061	0.25	9.19
39	62	36.48	0.308	8.399
39	79	52.301	0.097	13.627
39	94	44.112	0.171	10.791
94	24	74.119	0.025	20.898
94	39	44.823	0.146	11.028
94	62	34.425	0.332	7.768
94	79	32.21	0.401	7.096
125	24	48.945	0.084	12.438
125	39	45.722	0.106	11.327
125	62	42.349	0.137	10.221
125	79	46.02	0.104	11.427
125	94	39.257	0.175	9.251
136	24	66.266	0.045	18.37
136	39	47.473	0.141	11.924
136	94	45.153	0.167	11.138
136	125	46.607	0.15	11.626
9	94	69.478	1.105	19.401
24	94	60.531	1.393	16.51
42	94	51.785	1.573	13.443
62	94	48.25	1.619	12.194
87	94	34.824	1.335	7.89
114	94	25.412	0.588	5.087
123	94	45.228	1.625	11.163
136	94	39.352	1.512	9.28
9	136	142.742	0.184	30

0 M Semi-Empirical Distance - Gaussian Chain Constraint

Res 1	Res 2	r	γ
39	24	26.436	0.315
39	62	25.231	0.385
39	79	32.576	0.122
39	94	28.788	0.218
94	24	42.495	0.03
94	39	28.73	0.188
94	62	23.936	0.416
94	79	22.903	0.502
125	24	29.619	0.111
125	39	28.202	0.138
125	62	26.699	0.178
125	79	28.334	0.136
125	94	25.322	0.227
136	24	39.473	0.055
136	39	30.523	0.179
136	94	29.452	0.212
136	125	30.123	0.191
9	94	59.77	1.261
24	94	54.252	1.574
42	94	48.236	1.755
62	94	45.915	1.77
87	94	36.143	1.291
114	94	28.375	0.474
123	94	43.364	1.763
136	94	39.686	1.542
9	136	108.103	0.124

2 M Gaussian Chain Distance - Gaussian Chain Constraint

Res 1	Res 2	r	γ
39	24	31.254	0.323
39	62	27.808	0.45
39	79	34.562	0.239
39	94	33.201	0.27
94	24	58.374	0.039
94	39	37.018	0.184
94	62	24.762	0.587
94	79	27.224	0.458
125	24	28.271	0.334
125	39	36.222	0.155
125	62	32.75	0.214
125	79	31.391	0.244
125	94	30.104	0.277
136	24	48.748	0.078
136	39	40.401	0.147
136	94	36.77	0.199
136	125	43.522	0.115
9	94	49.257	1.551
24	94	45.782	1.618
42	94	41.443	1.611
62	94	42.748	1.616
87	94	33.392	1.37
114	94	23.095	0.541
123	94	39.607	1.596
136	94	33.74	1.338
9	136	82.725	0.729

2 M Semi-Empirical Distance - Harmonic Constraint

Res 1	Res 2	r	γ	σ
39	24	21.447	0.419	3.963
39	62	19.904	0.577	3.541
39	79	22.919	0.31	4.372
39	94	22.314	0.351	4.202
94	24	33.203	0.05	7.395
94	39	23.855	0.241	4.638
94	62	18.41	0.747	3.155
94	79	19.516	0.591	3.439
125	24	19.316	0.446	3.387
125	39	22.711	0.208	4.313
125	62	21.233	0.286	3.904
125	79	20.651	0.325	3.743
125	94	20.1	0.369	3.593
136	24	29.21	0.104	6.199
136	39	25.544	0.193	5.125
136	94	23.937	0.259	4.662
136	125	26.918	0.15	5.522
9	94	45.599	1.732	11.286
24	94	43.544	1.774	10.603
42	94	40.651	1.695	9.686
62	94	43.045	1.718	10.441
87	94	34.585	1.333	7.817
114	94	26	0.425	5.255
123	94	39.151	1.669	9.218
136	94	35.057	1.296	7.961
9	136	67.873	0.841	18.885

2 M Gaussian Chain Distance - Harmonic Constraint

Res 1	Res 2	r	γ	σ
39	24	31.254	0.323	6.808
39	62	27.808	0.45	5.785
39	79	34.562	0.239	7.81
39	94	33.201	0.27	7.394
94	24	58.374	0.039	15.78
94	39	37.018	0.184	8.564
94	62	24.762	0.587	4.899
94	79	27.224	0.458	5.613
125	24	28.271	0.334	5.922
125	39	36.222	0.155	8.319
125	62	32.75	0.214	7.258
125	79	31.391	0.244	6.849
125	94	30.104	0.277	6.463
136	24	48.748	0.078	12.369
136	39	40.401	0.147	9.608
136	94	36.77	0.199	8.488
136	125	43.522	0.115	10.596
9	94	49.257	1.551	12.549
24	94	45.782	1.618	11.348
42	94	41.443	1.611	9.935
62	94	42.748	1.616	10.347
87	94	33.392	1.37	7.452
114	94	23.095	0.541	4.422
123	94	39.607	1.596	9.359
136	94	33.74	1.338	7.558
9	136	82.725	0.729	23.729

2 M Semi-Empirical Distance - Gaussian Chain Constraint

Res 1	Res 2	r	γ
39	24	21.447	0.14
39	62	19.904	0.267
39	79	22.919	0.077
39	94	22.314	0.098
94	24	33.203	0.002
94	39	23.855	0.046
94	62	18.41	0.447
94	79	19.516	0.28
125	24	19.316	0.16
125	39	22.711	0.035
125	62	21.233	0.066
125	79	20.651	0.085
125	94	20.1	0.11
136	24	29.21	0.009
136	39	25.544	0.03
136	94	23.937	0.054
136	125	26.918	0.018
9	94	45.599	2.105
24	94	43.544	2.209
42	94	40.651	2.015
62	94	43.045	2.071
87	94	34.585	1.248
114	94	26	0.127
123	94	39.151	1.956
136	94	35.057	1.179
9	136	67.873	0.497

Assessment of the Constraint Functional Form and Sampling Method: Simulations were performed utilizing distances from the Gaussian chain (Eq. 3) and semi-empirical (Eq. 4) FRET distribution analyses in conjunction with either the Gaussian chain (Eq. S18) or harmonic (Eq. S19) constraints. This was done to assure that no bias was introduced by selecting a particular constraining method. No significant differences in the resulting ensembles were observed between simulations run with different constraint data or constraining functions when comparing the results in Flory-scaling plots (Fig. S54), or plots of computed E_{FRET} values (Fig. S55-S56), diffusion coefficients (Fig. S57), or "heat maps" of global structural analyses (Fig. S58-S59).

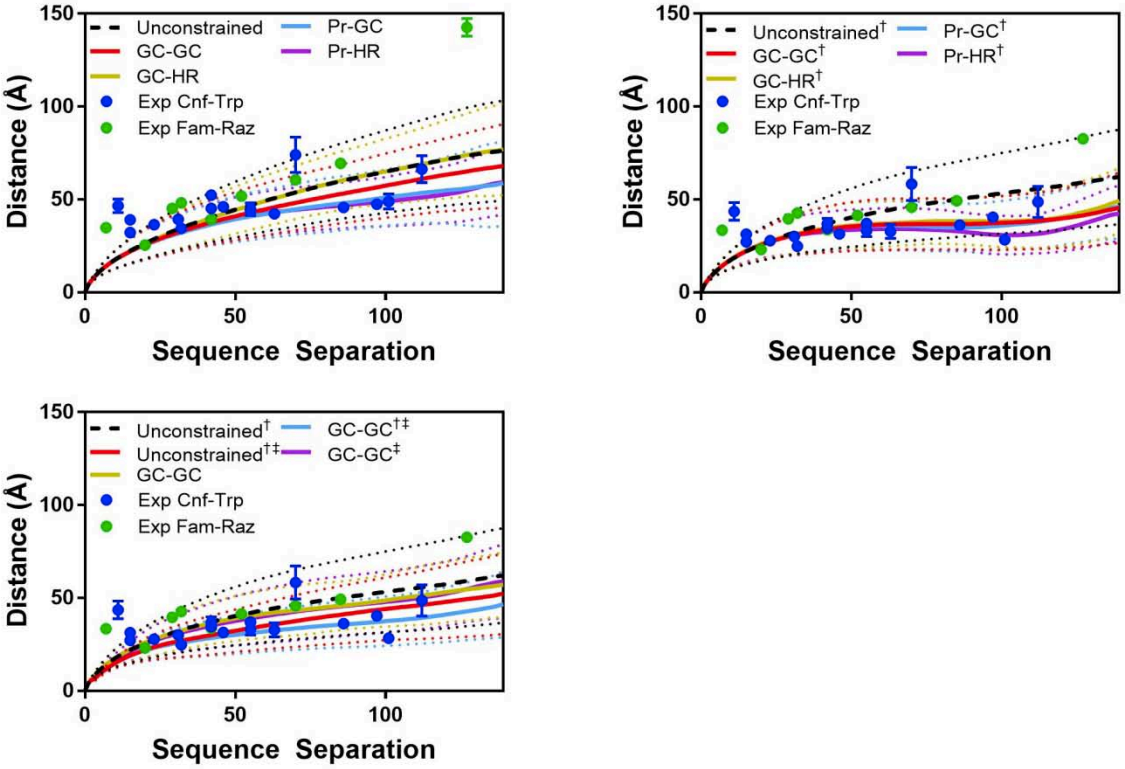

Figure S56: Flory Scaling Plots. Plots of Flory scaling determined from simulated ensembles constrained with distance constraints from measurements in 0 M (top left) and 2 M (top right/bottom left) TMAO. Names in the legends represent the probability distribution used in the polymer-scaled Förster equation for obtaining distances (Gaussian-chain = GC or semi-empirical = Pr) followed by the constraining function employed (Gaussian-chain = GC or harmonic potential = HR). Simulations marked with † indicate that the weight of the solvation term (fa_sol) was set to 0 while simulations marked with ‡ indicated that fragment insertion was incorporated within the search.

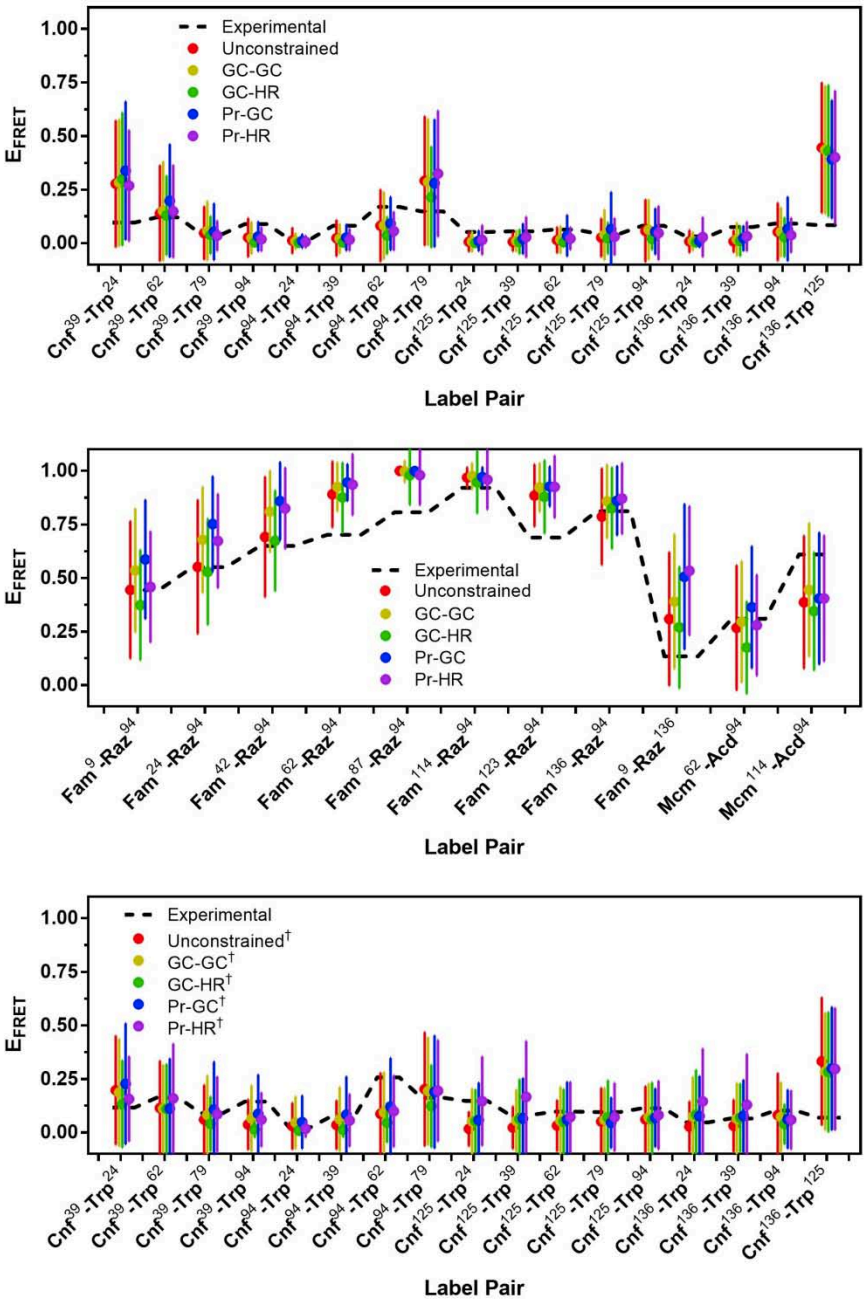

Figure S57: E_{FRET} Plots. Plots show the average and standard deviation of the E_{FRET} values determined from simulated ensembles constrained with distance constraints from measurements in 0 M (top/middle) and 2 M (bottom) TMAO. Names in the legends represent the probability distribution used in the polymer-scaled Förster equation for obtaining distances (Gaussian-chain = GC or semi-empirical = Pr) followed by the constraining function employed (Gaussian-chain = GC or harmonic potential = HR). Simulations marked with † indicate that the weight of the solvation term (fa_sol) was set to 0 while simulations marked with ‡ indicated that fragment insertion was incorporated within the search.

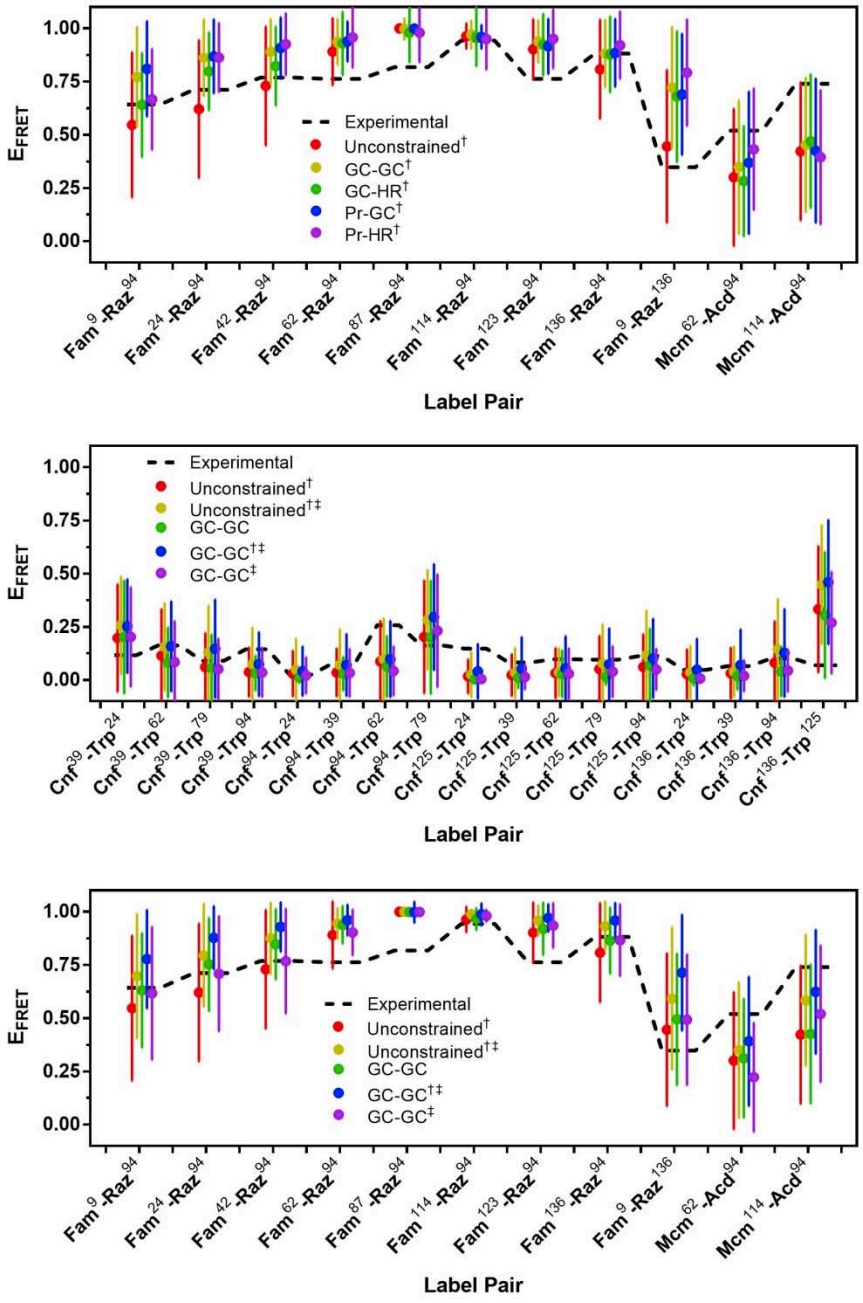

Figure S58: E_{FRET} Plots. Plots show the average and standard deviation of the E_{FRET} values determined from simulated ensembles constrained with distance constraints from measurements in 2 M TMAO. Names in the legends represent the probability distribution used in the polymer-scaled Förster equation for obtaining distances (Gaussian-chain = GC or semi-empirical = Pr) followed by the constraining function employed (Gaussian-chain = GC or harmonic potential = HR). Simulations marked with † indicate that the weight of the solvation term (fa_sol) was set to 0 while simulations marked with ‡ indicated that fragment insertion was incorporated within the search.

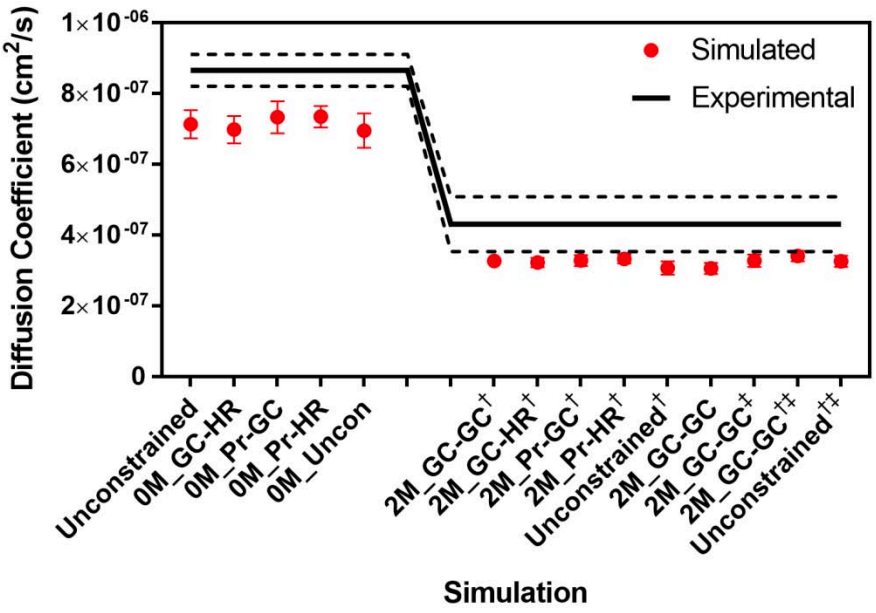

Figure S59: Diffusion Coefficients. Plots show the average diffusion coefficient determined from FCS data (black line) and from various simulations (red points). Simulation names represent the concentration of TMAO in which measurements were taken, the probability distribution used in the polymer-scaled Förster equation for obtaining distances (Gaussian-chain = GC or semi-empirical = Pr) and the constraining function employed (Gaussian-chain = GC or harmonic potential = HR), respectively. Simulations marked with † indicate that the weight of the solvation term (fa_sol) was set to 0 while simulations marked with ‡ indicated that fragment insertion was incorporated within the search.

Analysis of Simulated Structure Ensembles

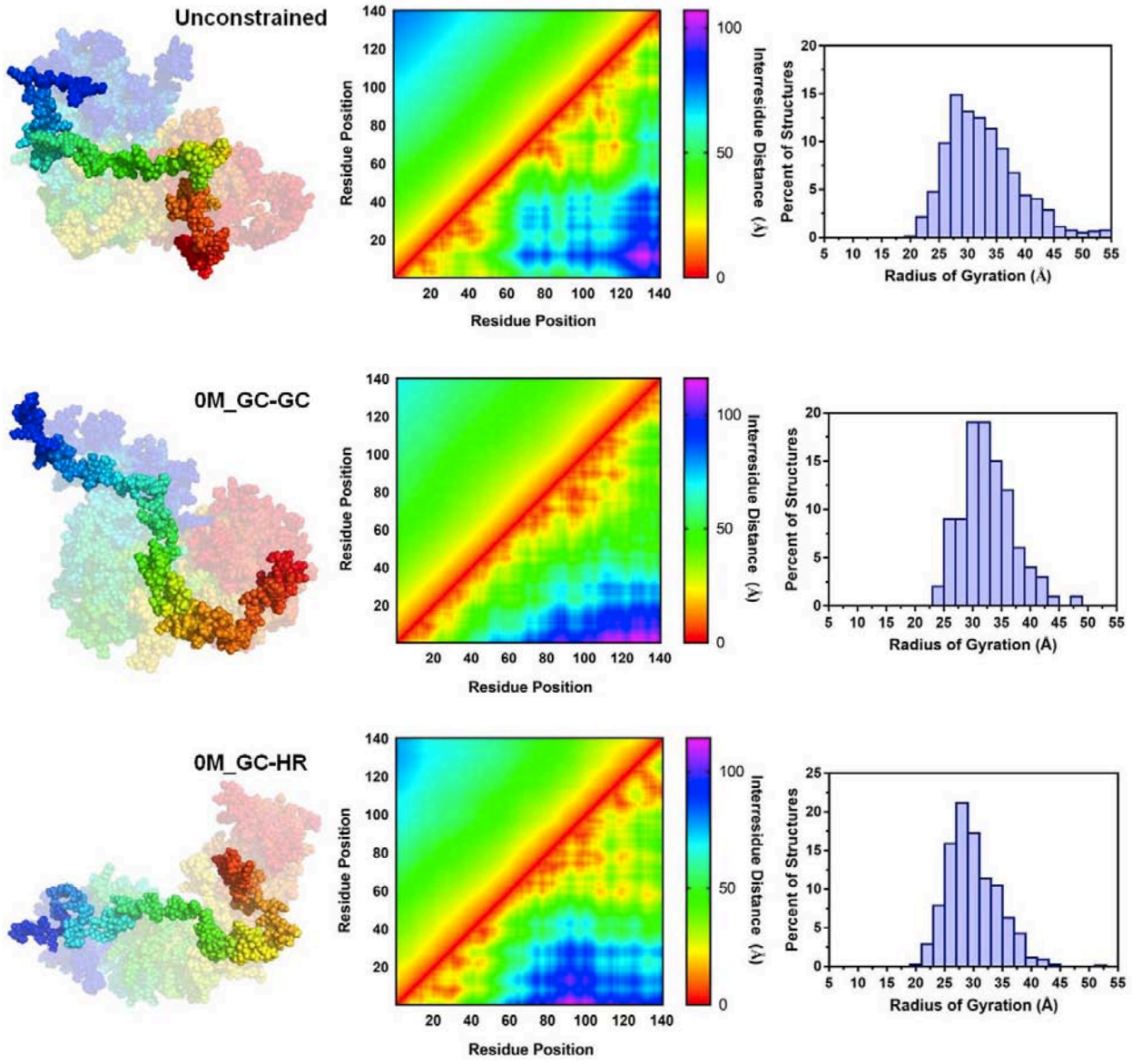

Figure S60: Global structural summary. Plots are derived from unconstrained (top set), 0M_GC-GC (middle set) and 0M_GC-HR (bottom set) simulations. Simulation names represent the concentration of TMAO in which measurements were taken, the probability distribution used in the polymer-scaled Förster equation for obtaining distances (Gaussian-chain = GC or semi-empirical = Pr) and the constraining function employed (Gaussian-chain = GC or harmonic potential = HR), respectively. Structures (left) show the 10 lowest energy structures faded with the single lowest energy structure darkened (blue - red, N-term - C-term). Heat map (middle) shows the average inter-residue distances from all output structures (above diagonal) and the inter-residue distances from the single lowest energy structure (below diagonal) as a function of residue pair. Histogram (right) of the percent of output structures for a given radius of gyration with bin widths of 2 Å.

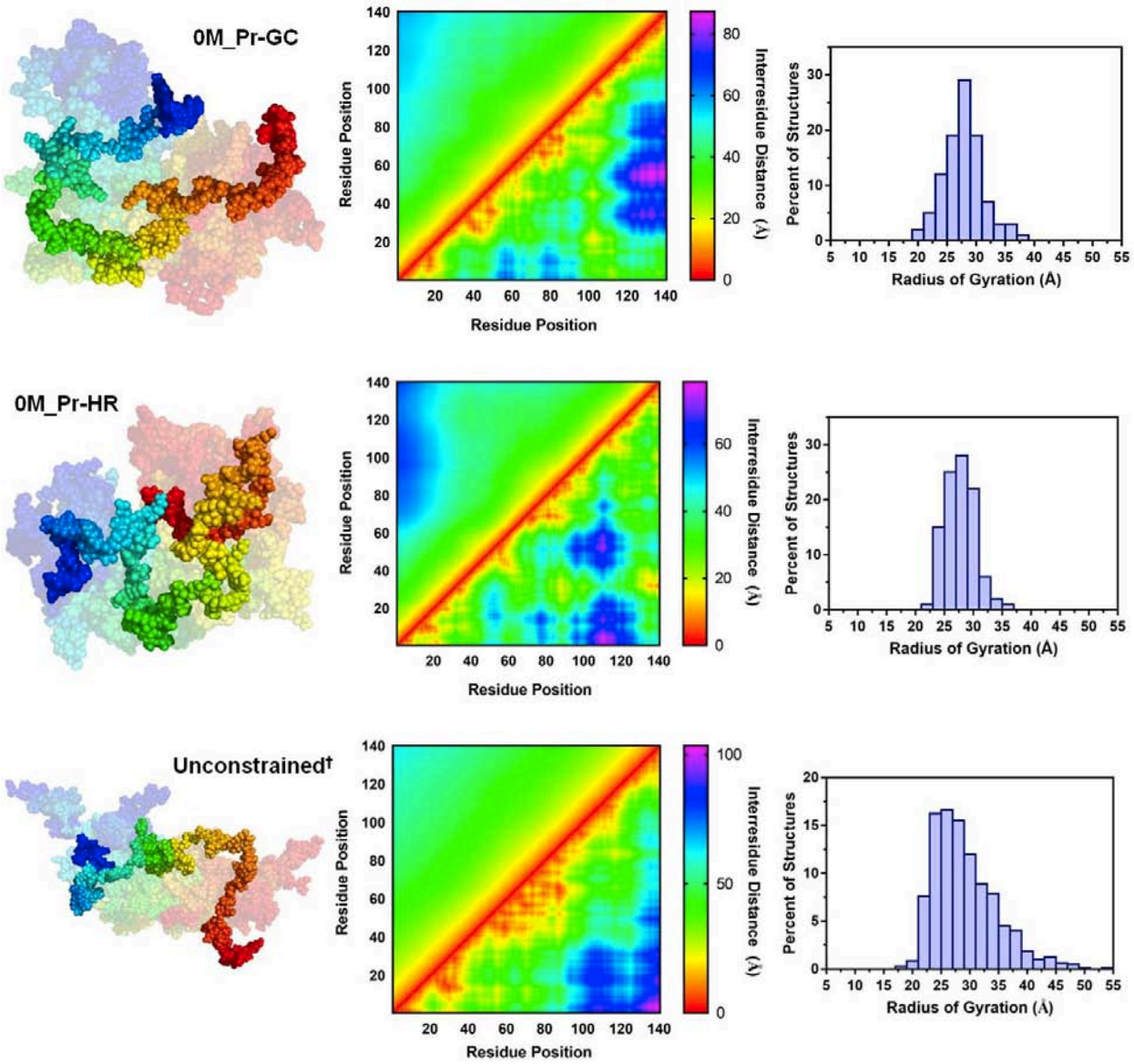

Figure S61: Global structural summary. Plots are derived from 0M_Pr-GC (top set), 0M_Pr-HR (middle set) and unconstrained† (bottom set) simulations. Simulation names represent the concentration of TMAO in which measurements were taken, the probability distribution used in the polymer-scaled Förster equation for obtaining distances (Gaussian-chain = GC or semi-empirical = Pr) and the constraining function employed (Gaussian-chain = GC or harmonic potential = HR), respectively. Structures (left) show the 10 lowest energy structures faded with the single lowest energy structure darkened (blue - red, N-term - C-term). Heat map (middle) shows the average inter-residue distances from all output structures (above diagonal) and the inter-residue distances from the single lowest energy structure (below diagonal) as a function of residue pair. Histogram (right) of the percent of output structures for a given radius of gyration with bin widths of 2 Å. Simulations marked with † indicate that the weight of the solvation term (fa_sol) was set to 0.

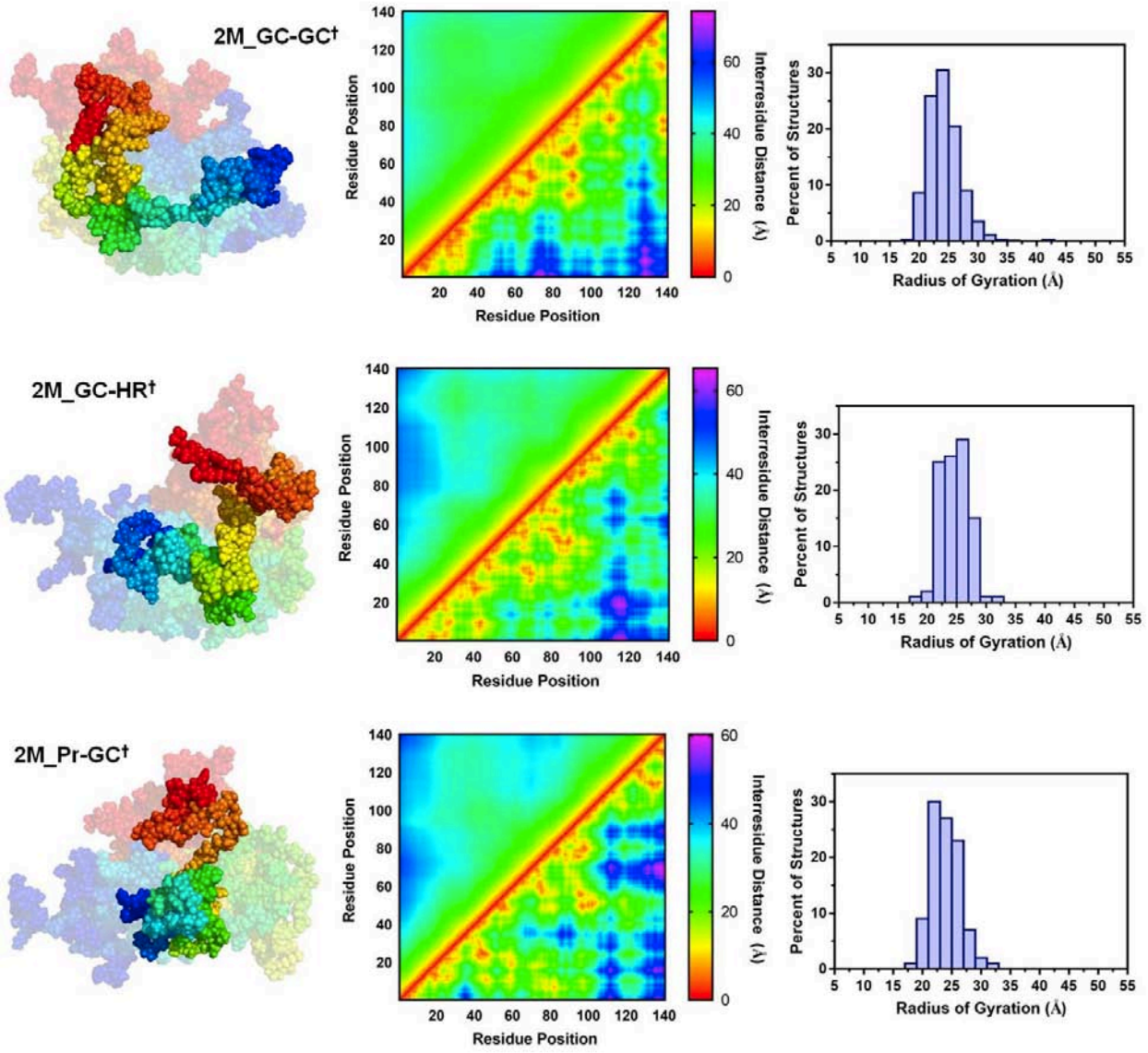

Figure S62: Global structural summary. Plots are derived from 2M_GC-GC† (top set), 2M_GC-HR† (middle set) and 2M_Pr-GC† (bottom set) simulations. Simulation names represent the concentration of TMAO in which measurements were taken, the probability distribution used in the polymer-scaled Förster equation for obtaining distances (Gaussian-chain = GC or semi-empirical = Pr) and the constraining function employed (Gaussian-chain = GC or harmonic potential = HR), respectively. Structures (left) show the 10 lowest energy structures faded with the single lowest energy structure darkened (blue - red, N-term - C-term). Heat map (middle) shows the average inter-residue distances from all output structures (above diagonal) and the inter-residue distances from the single lowest energy structure (below diagonal) as a function of residue pair. Histogram (right) of the percent of output structures for a given radius of gyration with bin widths of 2 Å. Simulations marked with † indicate that the weight of the solvation term (fa_sol) was set to 0.

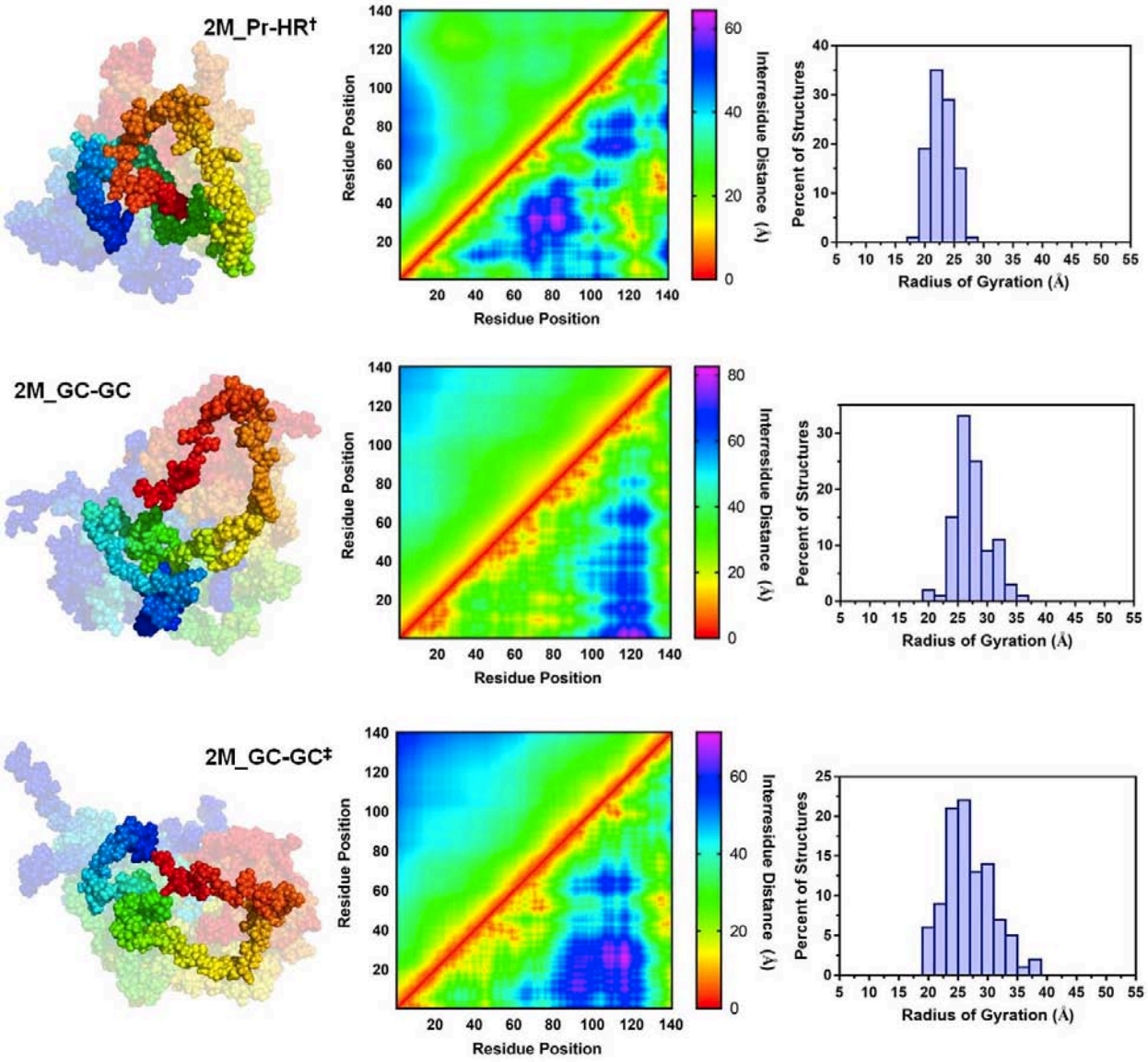

Figure S63: Global structural summary. Plots are derived from 2M_Pr-HR† (top set), 2M_GC-GC (middle set) and 2M_GC-GC‡ (bottom set) simulations. Simulation names represent the concentration of TMAO in which measurements were taken, the probability distribution used in the polymer-scaled Förster equation for obtaining distances (Gaussian-chain = GC or semi-empirical = Pr) and the constraining function employed (Gaussian-chain = GC or harmonic potential = HR), respectively. Structures (left) show the 10 lowest energy structures faded with the single lowest energy structure darkened (blue - red, N-term - C-term). Heat map (middle) shows the average inter-residue distances from all output structures (above diagonal) and the inter-residue distances from the single lowest energy structure (below diagonal) as a function of residue pair. Histogram (right) of the percent of output structures for a given radius of gyration with bin widths of 2 Å. Simulations marked with † indicate that the weight of the solvation term (fa_sol) was set to 0 while simulations marked with ‡ indicated that fragment insertion was incorporated within the search.

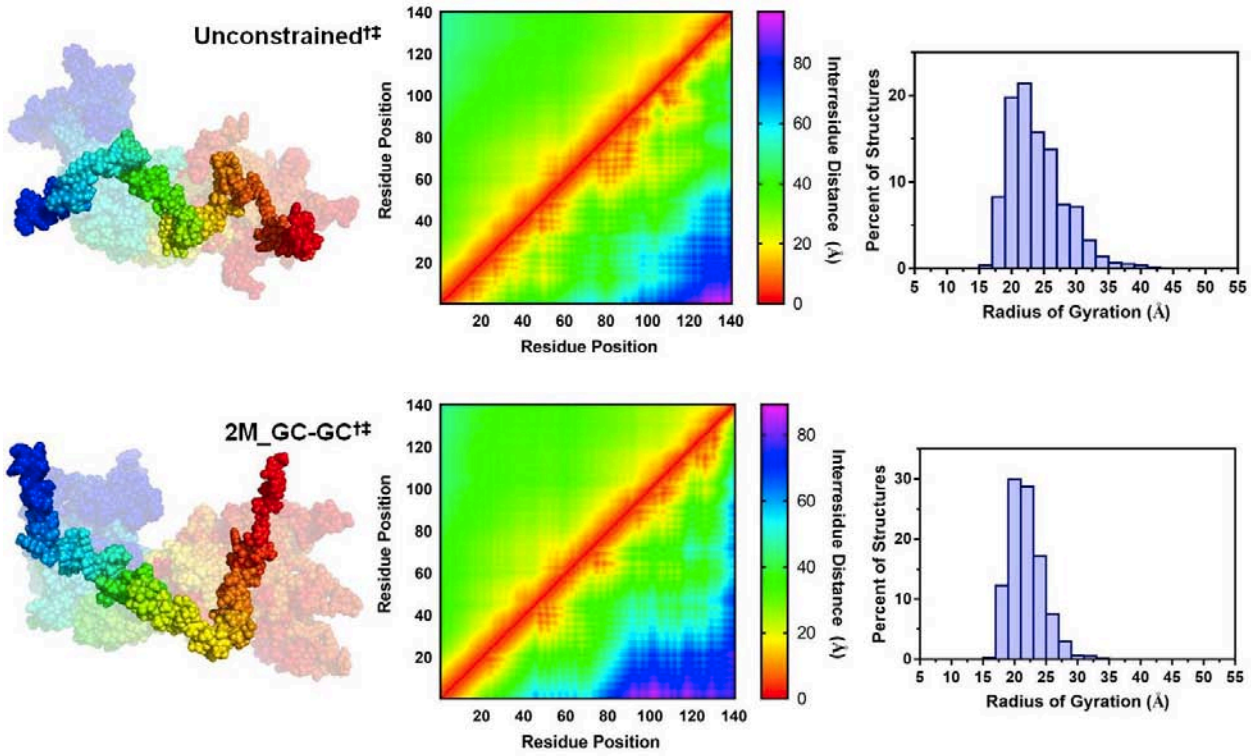

Figure S64: Global structural summary. Plots are derived from unconstrained†‡ (top set) and 2M_GC-GC†‡ (bottom set) simulations. Simulation names represent the concentration of TMAO in which measurements were taken, the probability distribution used in the polymer-scaled Förster equation for obtaining distances (Gaussian-chain = GC or semi-empirical = Pr) and the constraining function employed (Gaussian-chain = GC or harmonic potential = HR), respectively. Structures (left) show the 10 lowest energy structures faded with the single lowest energy structure darkened (blue - red, N-term - C-term). Heat map (middle) shows the average inter-residue distances from all output structures (above diagonal) and the inter-residue distances from the single lowest energy structure (below diagonal) as a function of residue pair. Histogram (right) of the percent of output structures for a given radius of gyration with bin widths of 2 Å. Simulations marked with † indicate that the weight of the solvation term (fa_sol) was set to 0 while simulations marked with ‡ indicated that fragment insertion was incorporated within the search.

Comparison of 0 M GC-GC Model to Literature Experiments and Simulations:

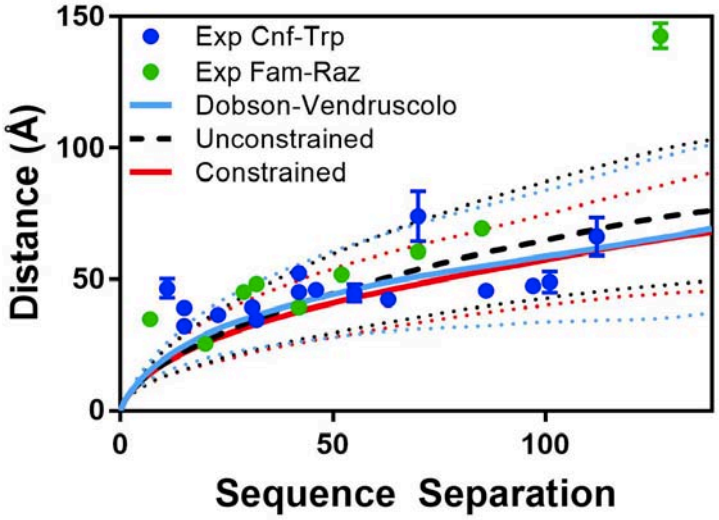

Figure S65: Flory Scaling Plot. Plot shows the average distance as a function of sequence separation for simulated ensembles (PED9AAC) and unconstrained and constrained simulations reported herein (12, 13). Distance constraints were derived from data acquired in 0 M TMAO and interpreted using the Gaussian-chain polymer scaled Förster equation. Constraints were implemented using the Gaussian-chain constraint function.

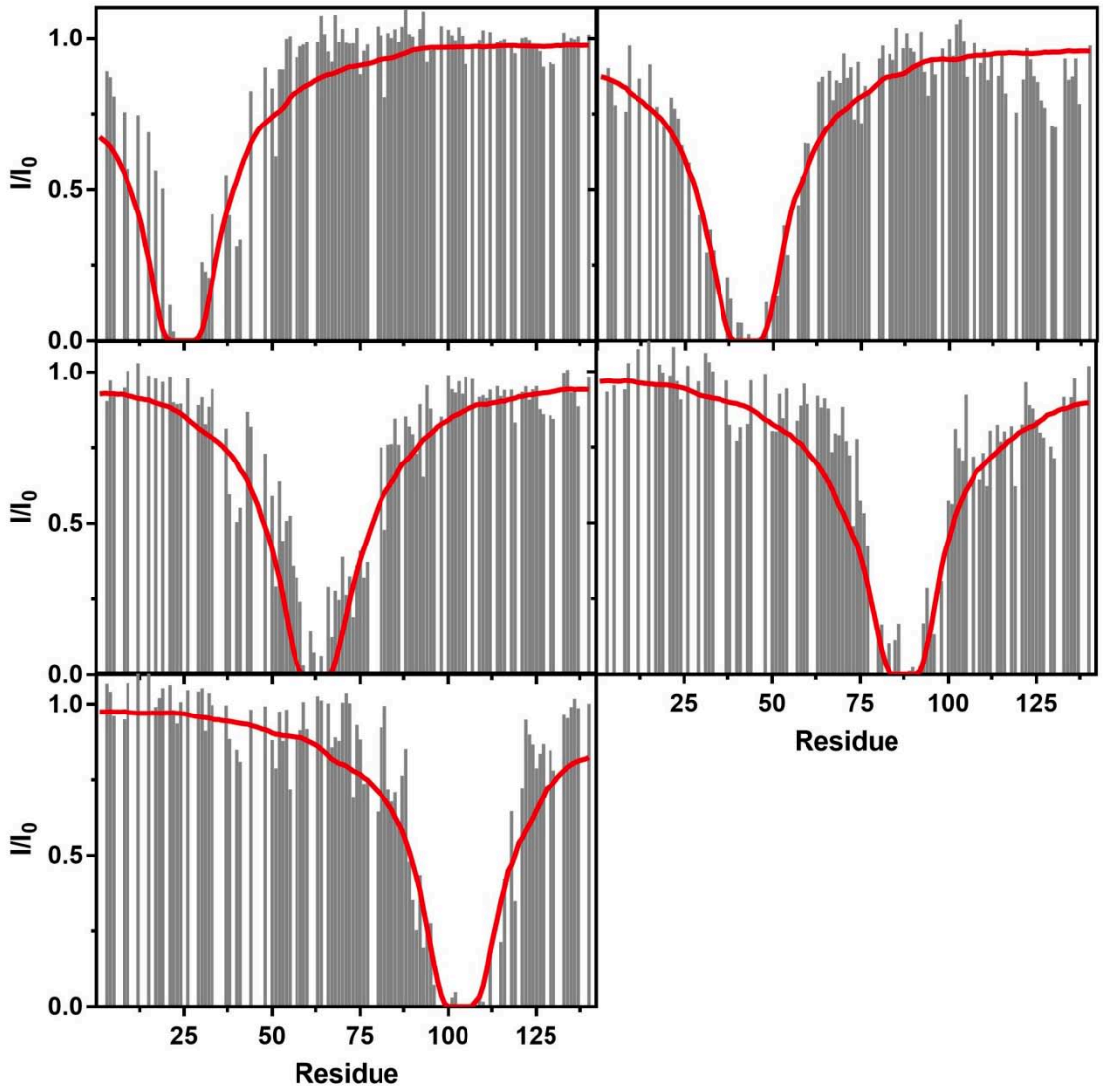

Figure S66: Intensity ratios from PRE measurements. Each plot shows the measured intensity ratio from PRE data obtained by Allison et. al. (grey bars) and the calculated average intensity ratio from the constrained simulation (red lines) were derived from data acquired in 0 M TMAO and interpreted using the Gaussian-chain polymer scaled Förster equation (12). Plots show data for spin-label placement at residues 24 (top left), 42 (top right), 62 (middle left), 87 (middle right) and 103 (bottom right).

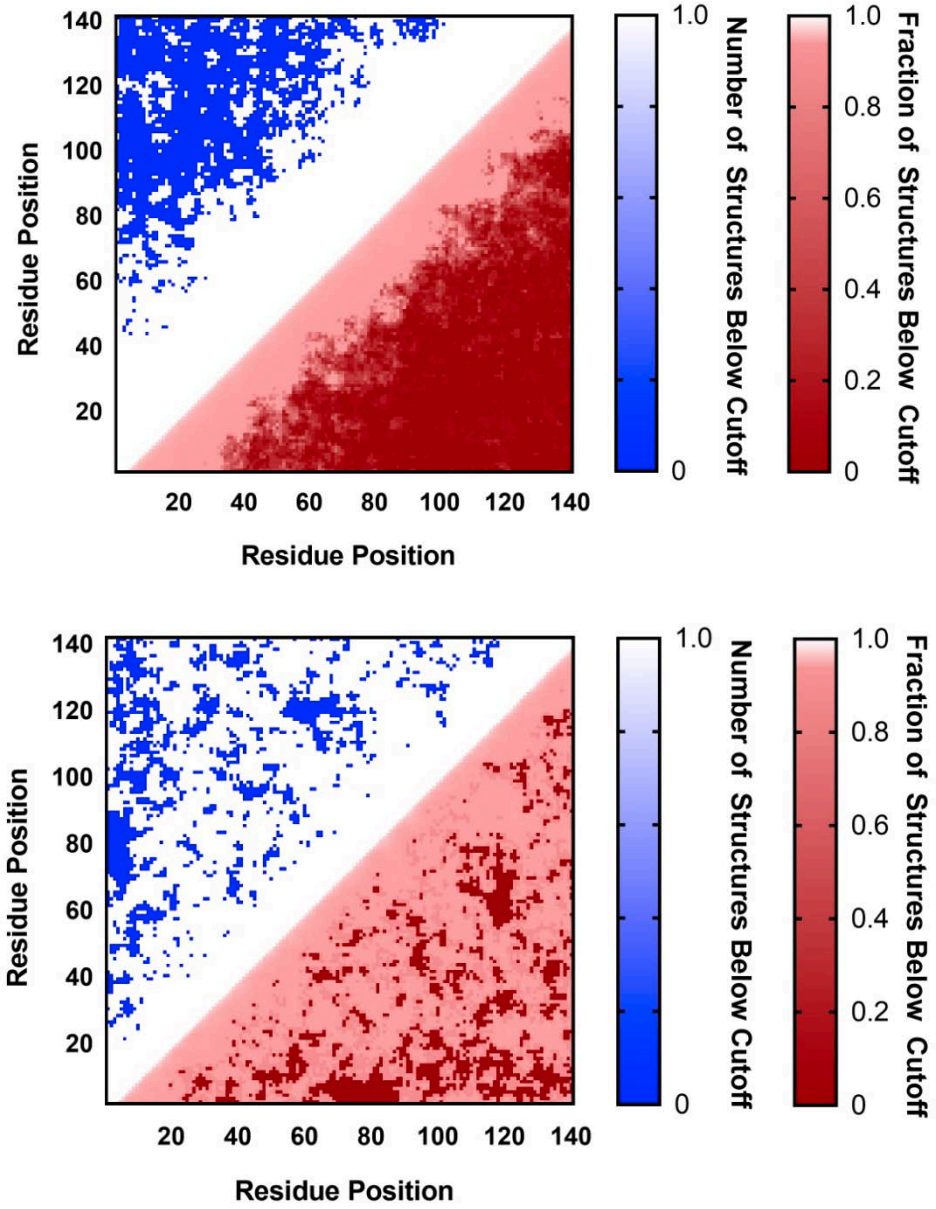

Figure S67: Contact maps from the 0M_GC-GC (top) and 2M_GC-GC† (bottom) ensembles. A cutoff distance of 10 Å was set as a contact. Each half of each heat map are analyses of the same ensemble and show all residue pairs which make contact in one or more structures from the ensemble as white while residue pairs which do not make contact are shown in blue (top left). The fraction of structures of the ensemble which make contact are shown over the full fractional scale with a red/white gradient (bottom right).

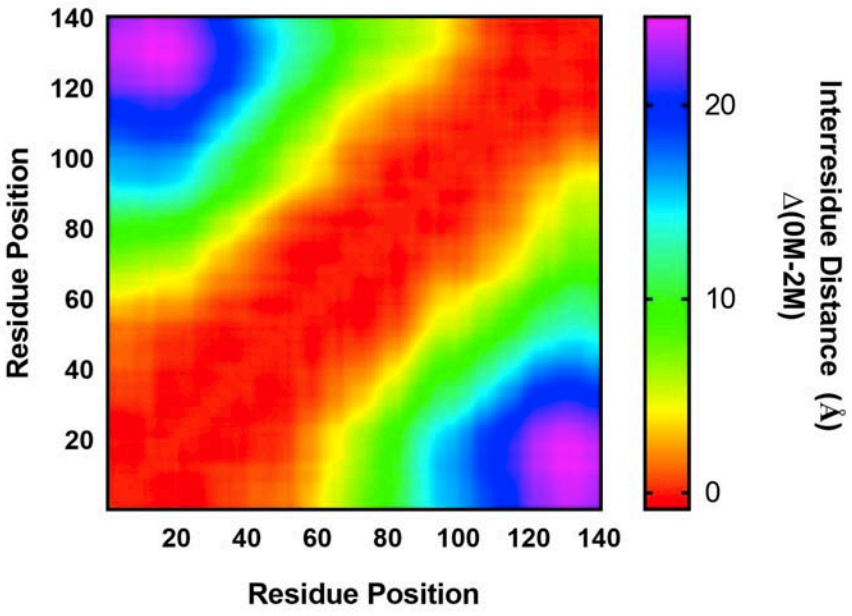

Figure S68: Heat map showing the average difference in interresidue distance between the 0M_GC-GC and 2M_GC-GC† ensembles.

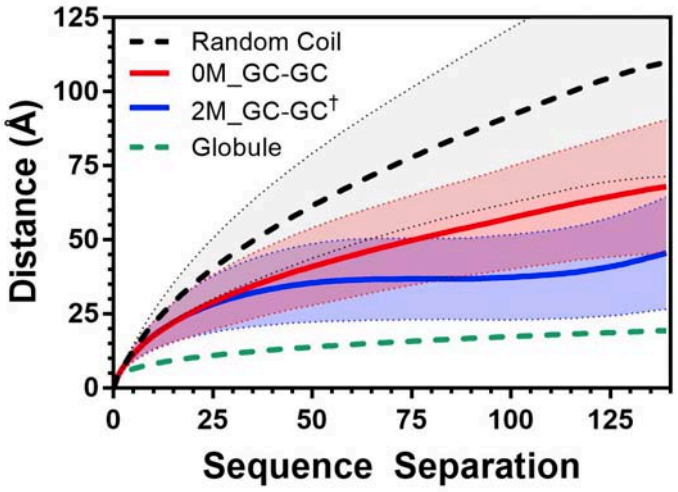

Figure S69: Flory scaling plots of 0M_GC-GC (red) and 2M_GC-GC† (blue) simulated ensembles along with the random coil simulation (black) used for the construction of $P_2(r,x)$ and a curve representing the scaling of a globule protein(green) (14).

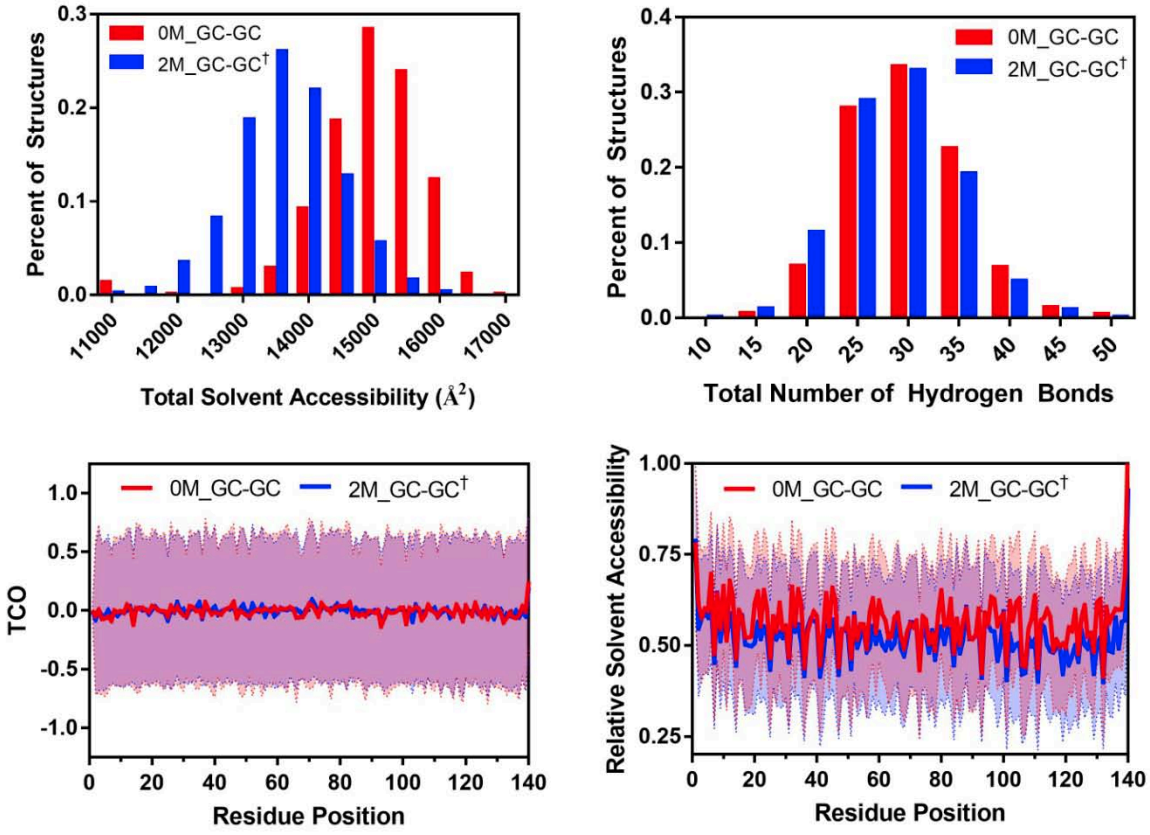

Figure S70: DSSP analysis of 0M_GC-GC (red) and 2M_GC-GC† (blue) ensembles for comparison. (Top Left) Histogram showing the percent of structures binned based on the total solvent accessible surface area of a given structure (Å^2). (Top Right) Histogram showing the percent of structures binned based on the total number of intramolecular hydrogen bonds present in a given structure. (Bottom Left) Plot of the average value of the cosine of the angle between the backbone carbonyl of the current residue and the carbonyl of the previous residue (TCO) for each residue for each structure within a given ensemble. (Bottom Right) Plot of the average value relative solvent accessibility for each residue for each structure within a given ensemble.

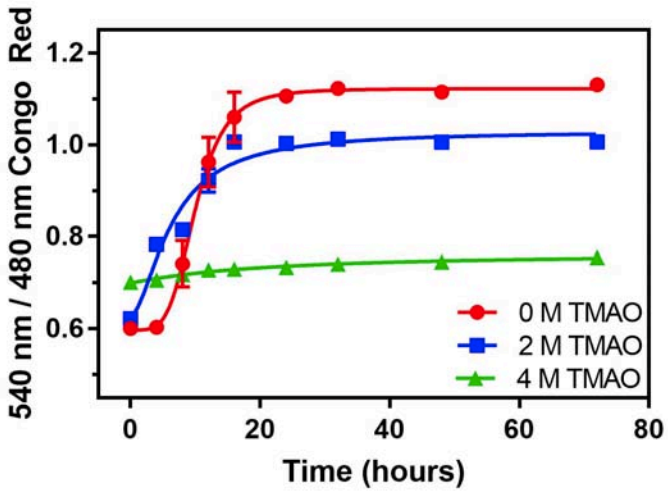

Figure S71: Aggregation kinetics of αS monitored by Congo Red. Aggregation was performed with 100 μM wild-type αS in 20 mM Tris 100 mM NaCl pH 7.5. TMAO was added to each buffer and the pH was readjusted. Samples were agitated in an orbital shaker at 1500 rpm at 37 °C. At each timepoint, 10 μL of sample was removed and added to 140 μL of 20 μM Congo Red in water and incubated for 20 minutes at room temperature prior to measurement. Absorption spectra (230-700 nm) were measured in a 96-well black CoStar clear bottom plate on a Tecan M1000 plate reader. Kinetics were fit to Eq. S22 in Prism with the detailed values in Table S7.

$$Y = A + (B - A)/(1 + (C/x)\text{^}D) \qquad \text{(Eq. S22)}$$

Table S7: Values from fits of Congo Red aggregation kinetics.

	0 M TMAO	2 M TMAO	4 M TMAO
A	0.5977	0.6289	0.7003
B	1.123	1.03	0.7656
C	9.989	6.397	19.72
D	4.357	1.708	1.104
R^2	0.9855	0.9526	0.9062

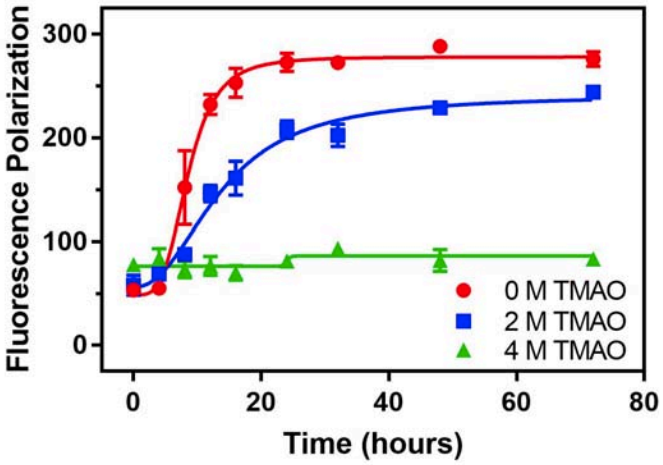

Figure S72: Aggregation kinetics of αS monitored by fluorescence polarization. Aggregation was performed with 100 μM αS-Fam_{136} in 20 mM Tris 100 mM NaCl pH 7.5. TMAO was added to each buffer and the pH was readjusted. Samples were agitated in an orbital shaker at 1500 rpm at 37 °C. At each timepoint, 10 μL of sample was removed and added to 90 μL of 20 mM Tris 100 mM NaCl pH 7.5. Fluorescence polarization measurements were taken in a 96-well black CoStar clear bottom plate on a Tecan F200 plate reader. Kinetics were fit to Eq. S22 in Prism with the detailed values in Table S8.

Table S8: Values from fits of fluorescence polarization aggregation kinetics.

	0 M TMAO	2 M TMAO	4 M TMAO
A	48.65	58.28	76.27
B	278.1	240.9	86.22
C	8.495	13.74	23.99
D	3.82	2.248	80.41
R^2	0.9806	0.9734	0.2733

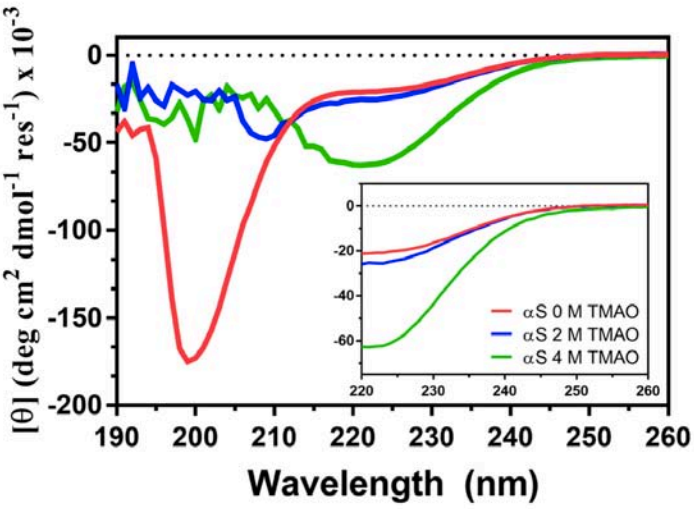

Figure S73: CD spectra of αS in varying concentrations of TMAO. Full plot and insert show the same data set, where the insert is adjusted to show difference in spectra above 220 nm. Measurements were taken on an Aviv model 410 circular dichorism spectrometer in 1 mm path length quartz cuvettes and were performed with 20 μM wild-type αS in 20 mM Tris 100 mM NaCl pH 7.5. TMAO was added to each buffer and the pH was readjusted at 25 °C. Scans from 190-260 nm were performed with a 1 nm bandwidth, 1 nm step size and an averaging time of 10 seconds.

REFERENCES:

1. Deiters, A., T. A. Cropp, M. Mukherji, J. W. Chin, J. C. Anderson, and P. G. Schultz. 2003. Adding Amino Acids with Novel Reactivity to the Genetic Code of Saccharomyces Cerevisiae. J Am Chem Soc 125:11782-11783.
2. Batjargal, S., C. R. Walters, and E. J. Petersson. 2015. Inteins as Traceless Purification Tags for Unnatural Amino Acid Proteins. J Am Chem Soc 137:1734-1737.
3. Haney, C. M., C. L. Cleveland, R. F. Wissner, L. Owei, J. Robustelli, M. J. Daniels, M. Canyurt, P. Rodriguez, H. Ischiropoulos, T. Baumgart, and E. J. Petersson. 2017. Site-Specific Fluorescence Polarization for Studying the Disaggregation of α-Synuclein Fibrils by Small Molecules. Biochemistry 56:683-691.
4. Haney, C. M., R. F. Wissner, J. B. Warner, Y. J. Wang, J. J. Ferrie, D. J. Covell, R. J. Karpowicz, V. M. Y. Lee, and E. James Petersson. 2016. Comparison of strategies for non-perturbing labeling of a-synuclein to study amyloidogenesis. Org Biomol Chem 14:1584-1592.
5. Gill, S. C., and P. H. von Hippel. 1989. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 182:319-326.
6. Middleton, E. R., and E. Rhoades. 2010. Effects of Curvature and Composition on α-Synuclein Binding to Lipid Vesicles. Biophysical Journal 99:2279-2288.
7. Glasscock, J. M., Y. Zhu, P. Chowdhury, J. Tang, and F. Gai. 2008. Using an Amino Acid Fluorescence Resonance Energy Transfer Pair To Probe Protein Unfolding: Application to the Villin Headpiece Subdomain and the LysM Domain. Biochemistry 47:11070-11076.
8. Horne, W. S., L. M. Johnson, T. J. Ketas, P. J. Klasse, M. Lu, J. P. Moore, and S. H. Gellman. 2009. Structural and biological mimicry of protein surface recognition by α/β-peptide foldamers. Proc Natl Acad Sci USA 106:14751-14756.
9. Kim, D. E., D. Chivian, and D. Baker. 2004. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 32:W526-W531.
10. Li, X.-H., J. A. Culver, and E. Rhoades. 2015. Tau Binds to Multiple Tubulin Dimers with Helical Structure. J Am Chem Soc 137:9218-9221.
11. Piana, S., A. G. Donchev, P. Robustelli, and D. E. Shaw. 2015. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. J Phys Chem B 119:5113-5123.
12. Allison, J. R., P. Varnai, C. M. Dobson, and M. Vendruscolo. 2009. Determination of the Free Energy Landscape of α-Synuclein Using Spin Label Nuclear Magnetic Resonance Measurements. J Am Chem Soc 131:18314-18326.
13. Varadi, M., S. Kosol, P. Lebrun, E. Valentini, M. Blackledge, A. K. Dunker, I. C. Felli, J. D. Forman-Kay, R. W. Kriwacki, R. Pierattelli, J. Sussman, D. I. Svergun, V. N. Uversky, M. Vendruscolo, D. Wishart, P. E. Wright, and P. Tompa. 2014. pE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins. Nucleic Acids Res 42:D326-D335.
14. Nath, A., M. Sammalkorpi, D. C. DeWitt, A. J. Trexler, S. Elbaum-Garfinkle, C. S. O'Hern, and E. Rhoades. 2012. The conformational ensembles of alpha-synuclein and tau: combining single-molecule FRET and simulations. Biophys J 103:1940-1949.