

A Desolvation Model for Trifluoroethanol-Induced Aggregation of Enhanced Green Fluorescent Protein

Valerie L. Anderson and Watt W. Webb*

School of Applied and Engineering Physics, Cornell University, Ithaca, New York

ABSTRACT Studies of amyloid disease-associated proteins in aqueous solutions containing 2,2,2-trifluoroethanol (TFE) have shown that the formation of structural intermediates is often correlated with enhanced protein aggregation. Here, enhanced green fluorescent protein (EGFP) is used as a model protein system to investigate the causal relationship between TFE-induced structural transitions and aggregation. Using circular dichroism spectroscopy, light scattering measurements, and transmission electron microscopy imaging, we demonstrate that population of a partially α -helical, monomeric intermediate is roughly correlated with the growth of β -sheet-rich, flexible fibrils for acid-denatured EGFP. By fitting our circular dichroism data to a model in which TFE-water mixtures are assumed to be ideal solutions, we show that increasing entropic costs of protein solvation in TFE-water mixtures may both cause the population of the intermediate state and increase aggregate production. Tertiary structure and electrostatic repulsion also impede aggregation. We conclude that initiation of EGFP aggregation in TFE likely involves overcoming of multiple protective factors, rather than stabilization of aggregation-prone structural elements.

INTRODUCTION

Numerous human diseases, including Alzheimer's disease, Parkinson's disease, and type II diabetes, are associated with protein aggregation (1). Biophysical investigations of aggregation processes frequently employ chemical cosolvents to reduce experimental variability, simulate cellular conditions, or induce the formation of atypical aggregate structures. The fluorinated alcohol 2,2,2-trifluoroethanol (TFE) is one of the most commonly used cosolvents in these studies (2).

The effects of TFE on protein structure have been studied for decades, but the physical interactions underlying TFE-induced aggregation enhancement are not well understood (2,3). Disordered proteins and peptides generally undergo a gradual coil-to-helix transition as TFE is added to a solution, reaching their maximally helical state by ~30–40% TFE. Additionally, TFE can denature globular proteins, typically leading to the formation of nonnative α -helical structure. When present at intermediate (~10–40% v/v) concentrations, TFE often induces or increases the aggregation of both globular and disordered proteins. For globular proteins, tertiary structure disruption often, but not always, precedes aggregation.

Many researchers have hypothesized that fluorinated alcohols promote protein aggregation by stabilizing fibrillogenic structural intermediates. Helical intermediates are

frequently associated with aggregation for disordered and denatured proteins (4–9). However, short (5–6-mer) peptides, which should not be capable of forming significant helical structure, also show enhanced fibril formation in ~7–10% TFE (10). Moreover, β -sheet-rich species have been detected before aggregation for some proteins (11–13). Also, multiple conformational states can lead to similar aggregates for some proteins (14), and TFE-induced aggregation can even occur in the absence of significant tertiary structure disruption (15,16). It is possible that various structural intermediates promote protein aggregation via multiple mechanisms. Alternatively, some solvent-dependent effect may cause reduced protein solubility, with structural transitions being coincidental.

The observation that fluoroalcohol molecules form clusters in water has led to the suggestion that aggregation might result from increased effective hydrophobicity of a protein bound to a cluster of relatively hydrophobic alcohol molecules (17,18). However, protein structural transitions and aggregation can occur at fluoroalcohol concentrations lower than those required for cluster formation (19). The decrease in the dielectric constant of alcohol-water mixtures has also been linked to the formation of H-bonded structures, including structural intermediates and β -sheet-rich aggregates, although polarity alone cannot account for the differences in behavior between fluorinated and simple alcohols (20). Therefore, relationships between bulk properties of TFE-water mixtures and protein structural transitions remain uncertain.

In this article, we investigate TFE-induced structural rearrangements and aggregation using enhanced green fluorescent protein (EGFP) as a model system. Near physiological pH, EGFP possesses β -can tertiary structure, whereas low and high pH conditions denature the protein, leading to

Submitted September 20, 2011, and accepted for publication January 20, 2012.

*Correspondence: www2@cornell.edu

This is an Open Access article distributed under the terms of the Creative Commons-Attribution Noncommercial License (<http://creativecommons.org/licenses/by-nc/2.0/>), which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editor: Kathleen Hall.

© 2012 by the Biophysical Society
0006-3495/12/02/0897/10 \$2.00

doi: 10.1016/j.bpj.2012.01.036

readily detectable loss of green fluorescence (21). Thus, EGFP is a nearly ideal system for examining the roles of tertiary versus secondary structure in TFE-induced conformational rearrangements. By combining circular dichroism (CD) spectroscopy with thermodynamic modeling, we show that changes in the chemical potential of solvent molecules may drive the observed structural transitions. We conclude that increased entropic costs of protein solvation may promote EGFP aggregation at moderate [TFE].

MATERIALS AND METHODS

Materials and solutions

Enhanced GFP (GFPmut1), derived from wild-type *Aequorea victoria* GFP with substitutions F64L and S65T (22), was synthesized by Dr. Cynthia Kinsland and the Cornell University Life Sciences Core Laboratories Center Protein Production Facility. Details of the expression can be found in the Supporting Material.

Acros Organics brand 99.8% pure 2,2,2-trifluoroethanol was purchased from Fisher Scientific (Pittsburgh, PA). Our pH measurements refer to aqueous samples; we did not correct our measurements to account for TFE effects on ionization constants of protein, water, or buffer components. These changes are likely to be small at low pH but can be significant at higher pH (23,24). Samples labeled pH 2.4 contained 10 mM phosphoric acid and pH 7.5 samples contained 10 mM sodium phosphate buffer. Additional solution conditions are specified in the figure legends.

Light scattering

Scattering experiments were performed using a QuantaMaster fluorescence spectrofluorometer (Photon Technology International, Birmingham, NJ) using 600-nm light to avoid EGFP absorption. The scattering angle was

90° and the same quartz cuvette was used to obtain all the data. Before measurement, 50 μ M EGFP samples, along with baseline solutions containing identical ingredients excepting the protein, were maintained under quiescent conditions for 24 ± 2 h in a 37°C incubator. The sample signals were normalized to the baseline signals.

CD spectroscopy

Far-UV CD measurements were performed using an Aviv 400 CD spectrometer (Aviv Biomedical, Lakewood, NJ). Data were collected at 1-nm intervals using a 1-nm bandwidth and a scan speed of 1 s per nm. Buffer-only baseline samples were measured and subtracted from the protein spectra and noise was reduced via a smoothing routine in the instrument software. The protein samples were first mixed with acid or buffer salts, and then TFE was added immediately before the measurements. However, for pH 3–5 samples, which were aggregation-prone, the protein was diluted into water, and the buffer salts and TFE were added together before the measurements.

Transmission electron microscopy imaging

A 37°C incubator was used to maintain solutions in quiescent conditions, whereas agitated samples were incubated in an orbital shaker operating at 200 RPM and 37°C. Sodium azide (0.02% w/v) was added to solutions incubated for longer than 24 h. To prevent grid damage, samples containing >15% TFE or large amounts of aggregates were diluted with water or buffer immediately before analysis. Images were obtained as described in Anderson et al. (4).

RESULTS

Secondary and tertiary structure of EGFP in TFE

Fig. 1, A and B, shows CD spectra of 0.3 μ M EGFP in the presence of 0–60% TFE. We show spectra that are

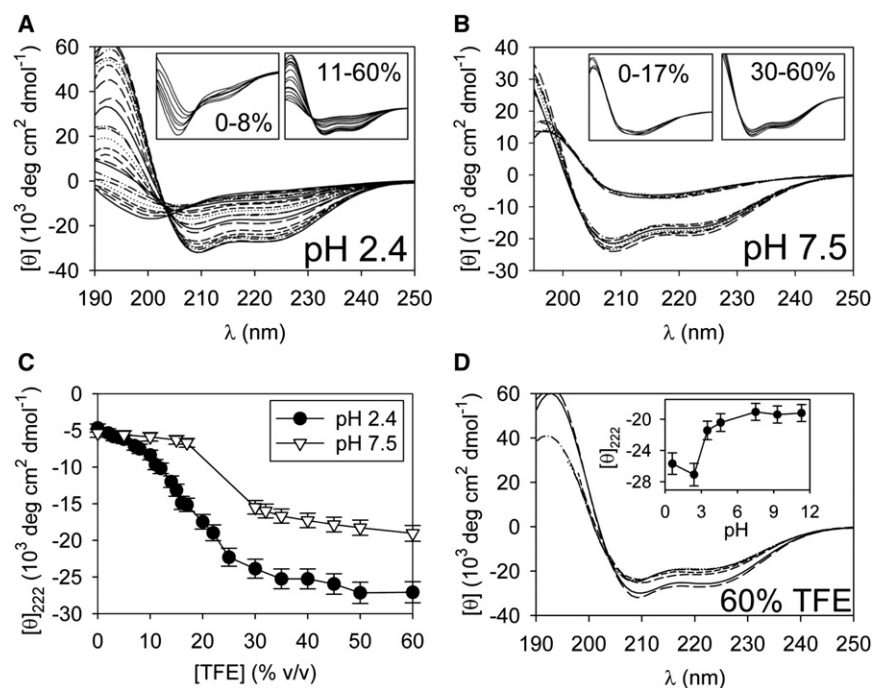


FIGURE 1 TFE-induced secondary structural transitions for EGFP. (A and B) Far-UV CD spectra for 0.3 μ M EGFP in 0–60% TFE at 37°C. (Insets) Selected spectra from the main plots, and the inset x axes' ranges are the same as the main plot ranges. (A) Data for EGFP at pH 2.4. The TFE concentrations for spectra with increasing negative ellipticity at 222 nm are 0, 2, 3, 4, 5, 7, 8, 10, 11, 12, 14, 15, 16, 17, 20, 22, 25, 30, 35, 40, 45, 60, and 50% TFE. (Insets) Spectra that share isodichroics. (Left y -axis range) -18 – $5 \times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$. (Right y -axis range) -37 – $65 \times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$. (B) Data for EGFP at pH 7.5. The TFE concentrations for spectra with increasing negative ellipticity at 222 nm are 0, 5, 10, 15, 17, 30, 32, 35, 40, 45, 50, and 60% TFE. (Left inset) Low-TFE spectra that are nearly invariant. (Right inset) High-TFE spectra that share an isodichroic point. (Left y -axis range) -9 – $18 \times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$. (Right y -axis range) -26 – $26 \times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$. (C) The mean residue ellipticity at 222 nm versus [TFE] for the spectra in A and B. (Error bars) Experimental uncertainties. (D) The CD spectra of EGFP in 60% TFE at various pH. The solution conditions for spectra with increasing negative ellipticity at 222 nm are 10 mM sodium phosphate (pH 7.5), 2 mM NaOH (pH 11.3), 10 mM borax (pH 9.3), 10 mM citrate-phosphate buffer (pH 4.6), 10 mM citrate-phosphate buffer (pH 3.5), 0.25 N sulfuric acid (pH 0.6), and 10 mM phosphoric acid (pH 2.4). (Inset) Ellipticity at 222 nm (in the same units as the main plot) as a function of pH for these samples. (Error bars) Experimental uncertainties.

consistent with monomeric or mostly monomeric protein, based on comparisons of 0.3 μM and 3 μM samples and the observation that the curves do not change significantly during the ~ 20 min (per sample) experimental duration (see Fig. S1 in the Supporting Material).

The pH 2.4, 0% TFE spectrum (Fig. 1 A) features the negative peak near 200 nm that is characteristic of polyproline type II (PPII) or statistical coil structure (25). As [TFE] increases, the ellipticity near 222 nm becomes larger-negative and the 200-nm peak becomes less prominent. At high TFE, the curves show the double minima at 208 and 222 that are expected for α -helical secondary structure; these TFE-induced helical states likely contain little to no tertiary structure (3).

Isodichroic points are apparent for two subsets of the spectra (Fig. 1 A, insets). The wavelength positions of these points are ~ 209 nm for the 0–8% TFE samples and ~ 203 nm for the 11–60% TFE samples. We also verify that EGFP is nonfluorescent at low pH in the presence of 0–60% TFE (see Fig. S2 A), as expected for denatured protein.

For EGFP in pH 7.5 solutions, the low (0–17%) TFE curves (Fig. 1 B, left inset) are consistent with the expected signal from β -can structure (26). Above 30% TFE, the spectra appear α -helical and an isodichroic point occurs near 203 nm (Fig. 1 B, right inset). We verify that EGFP remains fluorescent in pH 7.5 solutions immediately after heating to 37°C for [TFE] $< \approx 20\%$ (see Fig. S2 A). In contrast, rapid loss of fluorescence is observed for $\geq 30\%$ TFE, pH 7.5 samples. Note that we omit the ~ 20 –28% TFE, pH 7.5 spectra from Fig. 1 B because EGFP unfolds and oligomerizes during the experimental time frame in these conditions (see Fig. S1 and Fig. S3).

Plots of the ellipticity at 222 nm versus [TFE] (Fig. 1 C) reveal that the low-pH, high-TFE signal is relatively strong; the estimated number of helical residues (27) is $\sim 50\%$ higher in the acidic, 60% TFE solution compared to pH 7.5 conditions (see Table S1 in the Supporting Material). Examination of additional solution conditions reveals that the ellipticity shift occurs at pH ~ 3.5 (Fig. 1 D), which coincides roughly with the pK_a of acidic residues, although TFE may modify buffer and protein ionization constants (23,24). It is unlikely that an α -helix can propagate through the EGFP chromophore, which involves a covalent bond between Thr-65 and Gly-67 backbone groups. At pH 2.4, ~ 160 residues are estimated to be helical (see Table S1). Therefore, structure formation probably does not involve only the portion of the protein N-terminal to the chromophore. The C-terminal portion may form a continuous helix, or multiple short segments could form in various parts of the protein. Thus, the pH-dependent changes might involve either a change in helix lengths or in the number of helical segments. Electrostatic attraction between oppositely charged residues may tend to favor compact structures, rather than extended helices, near neutral pH

(28,29). However, higher-resolution experiments are necessary to determine the origin of the pH dependence of the helicity.

Reconstruction of the I-state spectrum and estimation of state populations

A transition-diagram (30) plot of the pH 2.4 CD data shows two linear segments (Fig. 2). The existence of two isodichroics in the CD spectra (Fig. 1 A) and two linear segments in the transition diagram indicate that EGFP is probably sampling at least three secondary structural conformations, which appear to include a low-TFE, PPII-like state (U), a high-TFE, highly helical state (HH), and an intermediate conformation (I).

Points derived from 0–17% TFE, pH 7.5 curves (Fig. 1 B) cluster in a region of the transition diagram that corresponds to the native β -can fold (*N* in Fig. 2). EGFP remains fluorescent in these solution conditions (see Fig. S2 A); therefore the slight spectral differences for these samples likely involve changes in loop regions, rather than the core β -can. In contrast, the higher-TFE, pH 7.5 points lie along the I \leftrightarrow HH transition line and reflect loss of tertiary structure in favor of nonnative helical conformations.

Following the procedure from Anderson et al. (4), we use principal component analysis (PCA) to obtain information about the conformations contributing to our CD spectra (see the Supporting Material). We find that acid-denatured EGFP samples three structural states within the resolution of our measurements (see Fig. S4 A), as might be expected given that the spectra contain two isodichroics (Fig. 1 A) and that the corresponding points lie along two lines in the transition diagram (Fig. 2).

We also use PCA to infer the CD spectrum of the I state (see the Supporting Material). The resulting curve (Fig. 3) features weak double minima, which are suggestive of

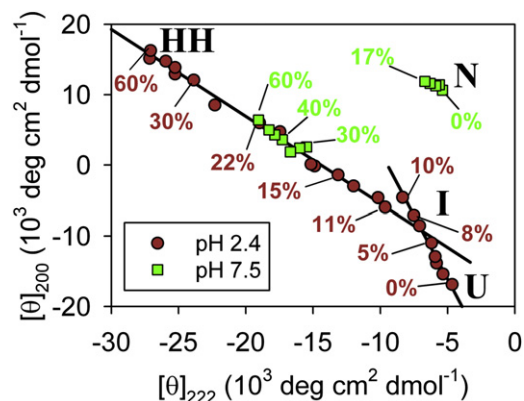


FIGURE 2 Transition diagram constructed from the CD spectra in Fig. 1, A and B. The plot shows the ellipticity at 200 vs. 222 nm. Some points are labeled with their [TFE], and the labels are color-coded to show pH. (Solid lines) Fits to the pH 2.4 points derived from spectra that share isodichroics (Fig. 1 A, insets).

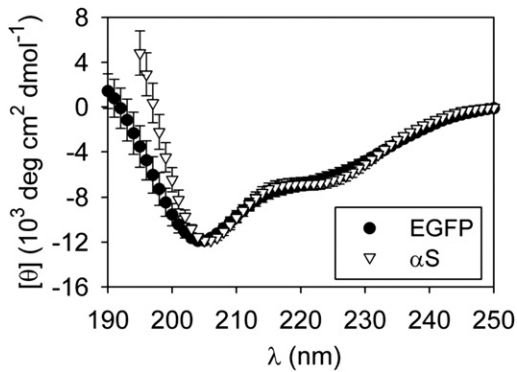


FIGURE 3 Reconstructed I-state spectrum for EGFP. The PCA-inferred EGFP curve is compared to the α S data from Anderson et al. (4). (Error bars) Uncertainties in the PCA estimates.

partial α -helical structure. Deconvolutions via several algorithms also predict an increase in helicity for the I state compared to the 0% TFE conformation (see Table S2). Furthermore, the reconstructed EGFP I state curve is very similar to the TFE-induced intermediate state spectrum for human wild-type α S, which may be partly helical (4).

By fitting our CD data to a linear combination of the inferred I-state spectrum (Fig. 3) and the measured 0% and 60% TFE curves (Fig. 1 A), we can obtain an estimate of the state populations over the full TFE range (Fig. 4 A, symbols; see Fig. S5 for more information about the fits). We find that the EGFP I state is maximally populated near 8% TFE. In contrast, fits to the data in Anderson et al. (4) suggest that peak I-state population for α S occurs at ~17% TFE (Fig. 4 B, symbols). Although the α S structural transitions were measured at 25°C and pH 7.5, whereas our EGFP data reflect 37°C and pH 2.4 conditions, we subsequently found that the TFE concentration at which the α S intermediate is maximally populated is not significantly affected by temperature variations or low pH conditions (V. L. Anderson, W. W. Webb, and D. Eliezer, unpublished data). Therefore, differences in solution conditions do not appear to be primarily responsible for the state population differences between the two proteins.

A desolvation model for TFE-induced secondary structural transitions

We construct a very simple model for the structural changes of denatured and disordered proteins in TFE-water mixtures. We briefly summarize the model here, and a more detailed derivation can be found in the Supporting Material.

Our desolvation model assumes that changes in the water and TFE chemical potentials perturb the grand potentials of the U and HH states. Therefore, this model predicts the relative populations of hydrated, TFE-solvated, and desolvated conformations, but does not account for the structural details of these states.

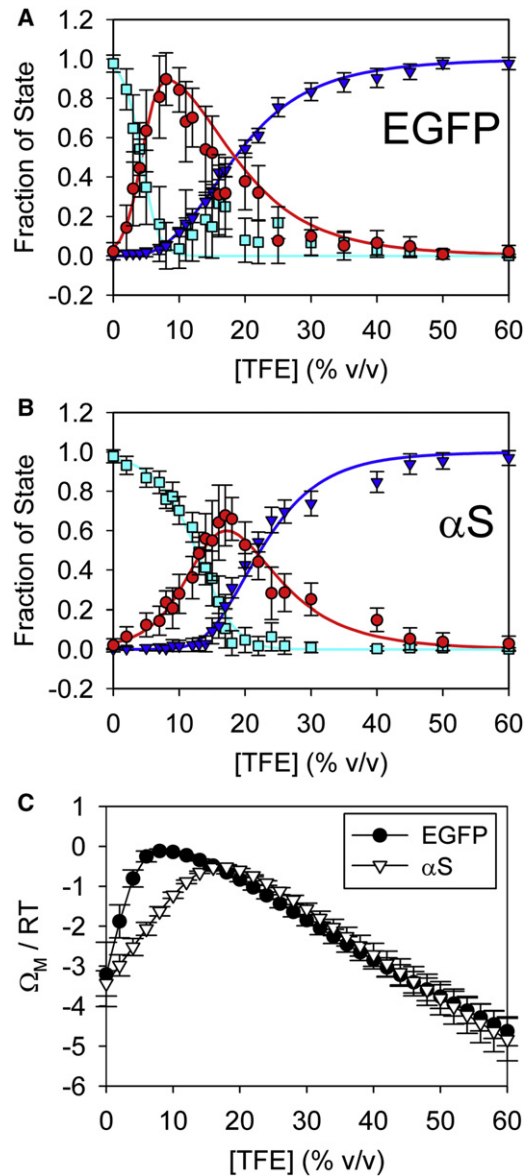


FIGURE 4 State populations and model fits for α S and EGFP. (A and B) State populations (symbols) for (A) acid-denatured EGFP and (B) α S, which were calculated by fitting the experimental spectra to a linear combination of the U (cyan), I (red), and HH (dark blue) state spectra. The error bars reflect estimated uncertainties due to noise in the experimental and I state spectra (see the Supporting Material). The EGFP populations were calculated from the data in Fig. 1 A, and the α S populations were calculated from the data in Anderson et al. (4). The lines show fits of the populations to our desolvation model (Eq. 3). (C) The average monomer grand potentials versus [TFE] for EGFP and α S. The error bars reflect standard errors resulting from the desolvation model fits.

In the ideal solution approximation, the chemical potentials of water and TFE (μ_w and μ_T , respectively) are equal to the pure substance chemical potentials plus terms related to the mixture entropy, so that (31)

$$\begin{aligned}\mu_w &= \mu_0^w + RT \ln(x_w), \\ \mu_T &= \mu_0^T + RT \ln(x_T),\end{aligned}\quad (1)$$

where x_w and x_T are the mole fractions of water and TFE ($x_w + x_T = 1$), T is the temperature in degrees Kelvin, and R is the ideal gas constant. The expressions in Eq. 1 imply that protein-water and protein-TFE interactions are entropically costly in TFE-water mixtures; therefore, the potentials of water and/or TFE solvated states will increase at moderate [TFE].

We consider the dilute-protein limit, so that the properties of the bulk solvent are not affected by protein solvation. We further assume that the 0% TFE, U state corresponds to a hydrated conformation in which some protein region interacts with n_w water molecules, the I state is a desolvated conformation in which the region is buried or solvent-shielded, and the HH state is TFE-solvated so that the region interacts with n_T TFE molecules. Also, we set the arbitrary zero-point potential to correspond to the I-state energy. Then, the grand canonical partition function (Z) for our system is

$$Z = (1 - x_T)^{n_w} \exp(-e_U/RT) + 1 + (x_T)^{n_T} \exp(-e_{HH}/RT), \quad (2)$$

where e_U (e_{HH}) is the grand potential of state U (HH) in pure water (TFE) (see the Supporting Material).

Using Eq. 2, we calculate the fractions of the U, I, and HH states (f_U , f_I , and f_{HH} , respectively):

$$\begin{aligned} f_U &= (1 - x_T)^{n_w} \exp(-e_U/RT)/Z, \\ f_I &= 1/Z, \\ f_{HH} &= (x_T)^{n_T} \exp(-e_{HH}/RT)/Z. \end{aligned} \quad (3)$$

We fit our estimated state populations to Eq. 3, converting between vol % TFE and x_T (Fig. 4, A and B, lines). The expressions in Eq. 3 fit both our EGFP data and the previously obtained α S data (4) fairly well, and we obtain similar reasonable fits for additional α S variants (see Fig. S6 and Table S3). Table 1 shows the parameters resulting from the fits.

In Fig. 4 C, we show the model prediction for the average monomer protein grand potential (Ω_M ; see Eq. S12 in the Supporting Material). Note that $\Omega_M = 0$ corresponds to 100% population of the I state. The range of TFE concentrations over which the monomer protein is destabilized (Ω_M is high) is predicted to be wider for EGFP than for α S.

Aggregate production versus [TFE]

Fig. 5 shows light scattering data, reflecting relative amounts of aggregate production, for 50 μ M EGFP samples incubated for 24 h at 37°C. At pH 2.4 with no added salt, aggregation is minimal over the entire 0–60% TFE range. However, the addition of 75 mM NaCl leads to aggregation enhancement for samples containing ~7.5–30% TFE.

At pH 7.5, aggregation is low for 0–7.5% TFE samples, increases sharply at 15% TFE, and then decreases at higher

TABLE 1 Parameters resulting from fits of the EGFP and α S state populations

	EGFP	α S
n_w	286 \pm 85	92 \pm 10
n_T	2.84 \pm 0.20	3.27 \pm 0.36
e_U/RT	-3.17 \pm 0.84	-3.40 \pm 0.33
e_{HH}/RT	-8.32 \pm 0.62	-9.2 \pm 1.0

State population data (Fig. 4) were fit to Eq. 3. The uncertainties reflect 95% confidence intervals for the fits.

[TFE]. Fluorescence loss occurs above 15% TFE for pH 7.5 samples (see Fig. S2 A); therefore, EGFP aggregation appears to require tertiary structure loss. However, very high TFE conditions stabilize monomeric protein.

Structural features of EGFP aggregates

CD spectra for EGFP samples identical to those examined in Fig. 5 reveal single minima near 216 nm, suggesting the presence of β -sheet structure (Fig. 6). For 50 μ M EGFP in pH 7.5 solutions containing 15–35% TFE (Fig. 6 B), the magnitude of the CD spectra near 216 nm is $\sim 2\times$ larger than the native protein value (Fig. 1 B), suggesting that aggregates are β -structured but not natively folded. The ~ 15 – 25 % TFE samples also appear cloudy-white after 24 h. We verify that removal of large aggregates from 15% TFE samples similar to those examined in Fig. 6 results in CD signal loss, indicating that large aggregates contribute to the observed spectra (see Fig. S7).

Even at low ($\sim \mu$ M) protein concentrations, pH 7.5, 15–25% TFE samples appear to be aggregation-prone, whereas pH 2.4 samples in low salt and/or reduced protein concentrations show CD spectra that are similar to the monomer curves (see Fig. S8).

Transmission electron microscopy (TEM) imaging reveals that the TFE-induced aggregates are thin, fibrillar

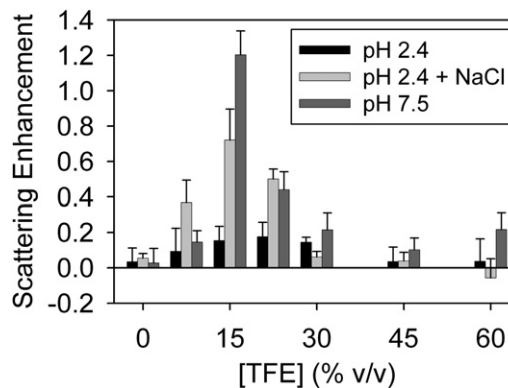


FIGURE 5 EGFP aggregation as a function of [TFE]. The plot shows the scattering signal from 50 μ M EGFP at pH 2.4 (solid bars), pH 7.5 (dark-shaded bars), and pH 2.4 with 75 mM NaCl (light-shaded bars). The samples were incubated for 24 \pm 2 h at 37°C in the presence of 0, 7.5, 15, 22.5, 30, 45, and 60% TFE. (Error bars) Standard deviations of measurements of three identical samples and baseline uncertainties.

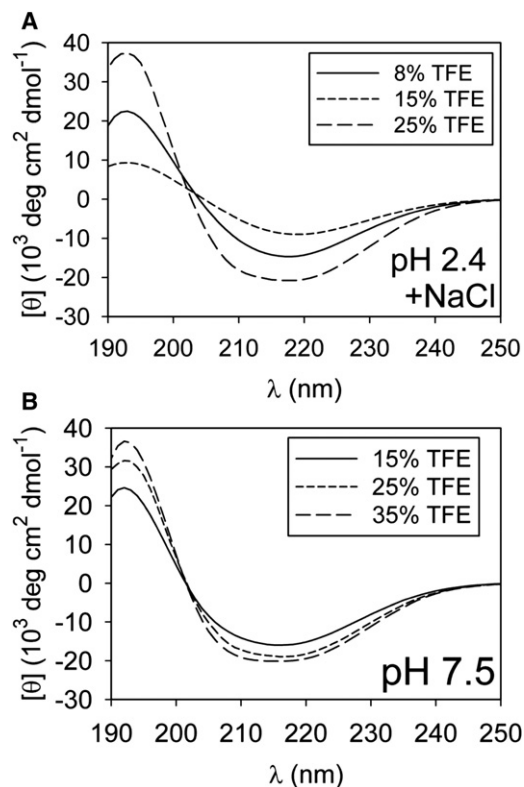


FIGURE 6 CD spectra of EGFP in aggregation-promoting conditions. 50 μ M EGFP samples were incubated for 24 ± 2 h at 37°C in various [TFE]. (A) Spectra for pH 2.4 solutions containing 75 mM NaCl. (B) Spectra for pH 7.5 samples.

structures, which often clump together (Fig. 7, A–C, and see Fig. S9). Rigid fibrils were sometimes observed after extended incubations (Fig. 7 D, and see Fig. S9).

DISCUSSION

We have demonstrated that TFE-induced EGFP aggregation requires tertiary structure disruption and is associated with reduced PPII-like and/or α -helical secondary structure. Using a simplified model, we showed that changes in solvent component chemical potentials may drive protein structural transitions. The free energy of the TFE-water mixture decreases at intermediate [TFE]. Consequently, it becomes costly to move solvent molecules from the bulk solvent into the protein solvation layer. The elevated entropic costs of protein solvation at moderate [TFE] may result in reduced population of solvated states in favor of solvent-shielded monomeric or oligomeric structures.

Entropic costs of protein solvation may drive secondary structural transitions in TFE-water mixtures

For both acid-denatured EGFP (Fig. 1 A) and disordered α S (4), the presence of small amounts ($< \approx 10$ –15%) of TFE

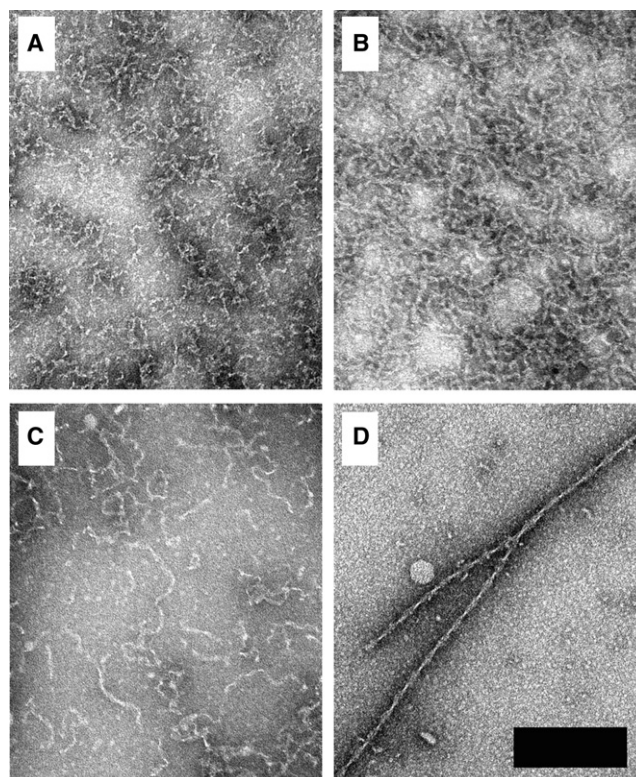


FIGURE 7 TEM images of EGFP aggregates. The scale bar is 200-nm wide and all images are shown at the same magnification. All solutions were incubated at 37°C . Unless otherwise specified, the samples were incubated for 24 ± 2 h. (A) Flexible fibrils grown from 50 μ M EGFP at pH 2.4 with 15% TFE and 75 mM NaCl. (B) Flexible fibrils grown from 50 μ M EGFP at pH 7.5 with 15% TFE. (C) Similar to panel B, except the solution contained 45% TFE. (D) Rigid fibrils grown from 50 μ M EGFP after seven weeks with shaking at pH 2.4 in the absence of TFE or NaCl.

causes a decrease in the magnitude of the negative peak near 200 nm in the CD spectra. This peak is thought to reflect a conformational bias toward PPII structure in the ensemble of disordered conformations (25). Water-protein H-bonding may enable the formation of sterically favorable PPII structure (32); therefore, the loss of the 200-nm CD peak is consistent with the reduction of protein hydration that is predicted by our model (Eq. 3).

The inferred CD spectrum of the EGFP I state appears partially α -helical (Fig. 3, and see Table S2). Partial helical structure has been observed before TFE-induced aggregation for many additional disordered or denatured proteins and peptides (4,5,7,9). If our desolvation model is correct, the intermediate structure may be a relatively compact, solvent-avoiding conformation containing partial secondary structure. However, additional experiments are necessary to test this hypothesis.

Our $>30\%$ TFE CD spectra (Fig. 1, A and B), which appear highly α -helical, are suggestive of preferential solvation of EGFP by TFE. Both experimental (33,34) and theoretical (35) studies have found that TFE binds to or coats proteins when present at medium-to-high ($> \approx 30\%$ v/v)

concentrations. Preferential solvation of proteins by TFE stabilizes α -helical structure, possibly as a result of chaotropic interactions or the low dielectric constant of TFE compared to water (36,37).

In the desolvation model, preferential TFE solvation is partly a consequence of favorable protein-TFE interactions (i.e., negative e_{HH} values in Table 1). Interestingly, the interactions that determine e_{HH} in our model are expected to be the same interactions that are responsible for protein solubilization and structure in pure TFE, and the fact that e_{HH} is larger-negative than e_U may reflect the high solubility of many proteins in neat TFE compared to water (38,39). Moreover, it is less entropically unfavorable to move a smaller number of molecules from the bulk solvent into the protein solvation shell; therefore the relatively small n_T values compared to n_W values (Table 1) tend to favor TFE, rather than water, solvation. Notably, for both EGFP and α S, the n_T/n_W ratios (Table 1) are larger than might be expected given the relative molecular sizes. However, changes in protein solvation are associated with changes in protein conformation; therefore, solvent-accessible surface areas may be different for the U and HH states.

Interestingly, n_T , e_U , and e_{HH} are similar for EGFP and α S. The major difference between the fit results for the two proteins is the number of water molecules predicted to be bound in the U state. The value n_W is approximately three times larger for EGFP than for α S. Both proteins contain a comparable fraction of nonpolar amino acids (38% for α S and 37% for EGFP when the numbers of Ala, Ile, Leu, Met, Phe, Tyr, and Trp residues are counted). However, α S contains a large proportion of Ala residues, whereas EGFP has more bulky nonpolar residues. Consequently, more water molecules should be required to form a solvation shell around denatured EGFP compared to α S. In addition, the sequence context of nonpolar residues in EGFP may favor water exclusion, given that EGFP adopts a globular fold near neutral pH.

The relative helix induction capabilities of some alcohols, hexafluoro-2-propanol > TFE > isopropanol > ethanol > methanol (18), appears to follow the order of the sizes of the molecules. Our model predicts that smaller n_T values will reduce the entropic costs of alcohol solvation, leading to HH-state population at relatively low alcohol concentrations. Thus, if fewer larger molecules are needed to solvate a protein, it may be expected that larger alcohols may be better helix inducers. However, the strengths of alcohol-water interactions (e_{HH}) should also influence helix formation, and the structural details of the high-alcohol states will be affected by the specific properties of the alcohols.

Our desolvation model assumes that mixing entropy alters the chemical potentials of TFE and water, leading to structural transitions. We do not take into account changes in the dielectric constant or viscosity of the bulk solvent, which may also affect the energies of the various states. Despite its simplicity, our model can reproduce many key

features of our data. However, it is possible that corrections to account for solvent properties will be needed to refine the model. Measurements of protein structures and solubility in various alcohols could help to assess the validity of the ideal solution approximation and determine the primary factors responsible for the high helix induction capabilities of fluoroalcohols.

Desolvation may initiate protein aggregation at moderate [TFE]

Enhanced aggregation of acid-denatured EGFP in 75 mM NaCl (Fig. 5) occurs for TFE ranges that are roughly similar to those in which the I-state population is high in 0 mM NaCl solutions (Fig. 4 A). Note that we compare EGFP monomer structures at 0 mM NaCl with aggregate production at 75 mM NaCl, although the protein conformations may not be identical in the different solution conditions. We are unable to measure monomeric EGFP spectra for ~0–20% TFE, 75 mM NaCl solutions because of aggregation (see Fig. S10). However, the high-TFE spectra are qualitatively similar in the presence and absence of NaCl, and EGFP may sample a similar partially α -helical intermediate state in both conditions. Partially helical states appear to be more aggregation-prone than highly helical states at for acid-denatured EGFP at 75 mM NaCl, as well as for high TFE, pH 7.5 samples (Fig. 5). A similar correlation between the formation of partial α -helical intermediates and aggregation has been observed for many additional disordered and denatured proteins in fluoroalcohols (4–9).

Our model (Eq. 3) predicts that I-state population and aggregation enhancement both may result from a solvent entropy-driven increase in the grand potentials of solvated conformations. At moderate [TFE], the average potential of the monomer protein increases (Fig. 4 C), favoring aggregated conformations. However, the amount of aggregate produced will also depend on energies of oligomeric and fibrillar species, which may vary with [TFE]. Therefore, peak aggregation may not exactly coincide with maximal destabilization of monomeric protein.

For acid-denatured EGFP in 75 mM NaCl, we find that the peak scattering signal may be slightly shifted rightward to ~15% TFE, compared to the ~8% TFE I-state peak (Fig. 4 B), although the differences are within the measurement uncertainties. The presence of NaCl may affect the state populations, or the relatively low dielectric constant of TFE compared to water may shift aggregation peaks rightward as a result of favorable H-bond formation (20).

Decreases in the free energy of the bulk solvent are expected to favor the formation of solvent-shielded and oligomeric structures (40). In addition, desolvation is consistent with previous reports that hydration impedes amyloid aggregation (41–44). Therefore, we believe that the simplicity of our model, combined with the fact that it offers a straightforward explanation for both the structural changes and the

aggregation enhancement observed for proteins in TFE, make it a reasonable initial analysis. However, we do not directly measure protein-solvent interactions, and the fact that the desolvation model fits our data does not prove that it is correct. Moreover, because changes in solvent properties and protein conformation are likely to occur in parallel, it may be difficult to definitively separate correlation from causation in aggregation reactions. Experiments and simulations that directly address the hydration status of proteins might help verify the primary factors driving aggregation in the presence of TFE (41,43).

Multiple protective factors impede EGFP aggregation

Previous studies have identified electrostatic repulsion and globular structures as natural protective mechanisms, which likely evolved to prevent abnormal protein aggregation (45). Hydration is also an important stabilizing interaction for disordered proteins (46). Our data suggest that several types of protective interactions impede EGFP aggregation.

We find that aggregation of acid-denatured EGFP is minimal in low ionic strength solutions (Fig. 5). The net charge per EGFP molecule is predicted to be $+36e$ pH 2.4, compared to $-7.7e$ at pH 7.5 (47); therefore electrostatic repulsion likely opposes aggregation of the acid-denatured protein. The addition of 75 mM NaCl enables significant aggregation in solutions containing ~7.5–30% TFE, indicating that electrostatic screening can overcome these protective repulsive forces.

For EGFP in pH 7.5 solutions, aggregation jumps sharply at 15% TFE (Fig. 5), where tertiary structure disruption occurs (see Fig. S2 A). However, increased α -helical structure at higher TFE (Fig. 1) is associated with reduced fibrilization. Therefore both intact tertiary structure and TFE-protein interactions may impede EGFP fibrilization at neutral pH.

Our model predicts that increased entropic costs of protein solvation overcome protective protein-solvent interactions, leading to monomer destabilization at intermediate [TFE] (Fig. 4 C). Moreover, if a modification could be made to a protein that would prevent the formation of the I state while leaving the potentials of other states unchanged, the ensemble-average monomer potential would increase in the TFE-water mixture (see Fig. S11). Therefore, the I state may be somewhat protective, in the sense that adoption of solvent-shielded intermediate conformations could reduce aggregate production.

CONCLUSION

The secondary structural transitions for acid-denatured EGFP in TFE are qualitatively similar to those previously observed for α S (4) and are consistent with dehydration leading to loss of PPII structure at low-to-moderate TFE.

At high TFE, favorable TFE-protein interactions induce the formation of α -helical structure and inhibit EGFP aggregation. We propose that monomer destabilization initiates fibrilization at intermediate [TFE]. However, intact tertiary structure and electrostatic repulsion impede aggregation. Hence, TFE-induced aggregation likely involves overcoming multiple protective factors, rather than the stabilization of aggregation-prone conformations.

SUPPORTING MATERIAL

Supporting text with 11 figures, four tables, and references (48–61) is available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(12\)00150-6](http://www.biophysj.org/biophysj/supplemental/S0006-3495(12)00150-6).

The authors thank Prof. David Piston and Dr. Gert-Jan Kremers at Vanderbilt University for their generous gift of GFP variant vectors. We also are grateful to Mark Williams for help with manuscript preparation, and to Prof. D. Eliezer, Prof. J. Sethna, Prof. L. Nicholson, J. Grazul, and Y. Zhang for helpful discussions.

This research made use of the Hudson Mesoscale Processing facility of the Cornell Center for Materials Research with support from the National Science Foundation Materials Research Science and Engineering Centers program (DMR 1120296 and DMR 0520404). Funding was provided by grants from the National Institutes of Health (5 R21 AG026650) and the National Science Foundation (Science and Technology Centers program under agreement No. ECS-9876771).

REFERENCES

1. Chiti, F., and C. M. Dobson. 2006. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75:333–366.
2. Otzen, D. E. 2010. Amyloid formation in surfactants and alcohols: membrane mimetics or structural switchers? *Curr. Protein Pept. Sci.* 11:355–371.
3. Buck, M. 1998. Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* 31:297–355.
4. Anderson, V. L., T. F. Ramlall, ..., D. Eliezer. 2010. Identification of a helical intermediate in trifluoroethanol-induced α -synuclein aggregation. *Proc. Natl. Acad. Sci. USA.* 107:18850–18855.
5. Fezoui, Y., and D. B. Teplow. 2002. Kinetic studies of amyloid β -protein fibril assembly. Differential effects of α -helix stabilization. *J. Biol. Chem.* 277:36948–36954.
6. Zerovnik, E., M. Skarabot, ..., R. A. Staniforth. 2007. Amyloid fibril formation by human stefin B: influence of pH and TFE on fibril growth and morphology. *Amyloid.* 14:237–247.
7. Sen, P., B. Ahmad, ..., R. H. Khan. 2010. 2,2,2-Trifluoroethanol induces simultaneous increase in α -helicity and aggregation in alkaline unfolded state of bovine serum albumin. *Int. J. Biol. Macromol.* 46:250–254.
8. Williamson, J. A., J. P. Loria, and A. D. Miranker. 2009. Helix stabilization precedes aqueous and bilayer-catalyzed fiber formation in islet amyloid polypeptide. *J. Mol. Biol.* 393:383–396.
9. Liu, W., J. M. Prausnitz, and H. W. Blanch. 2004. Amyloid fibril formation by peptide LYS (11–36) in aqueous trifluoroethanol. *Biomacromolecules.* 5:1818–1823.
10. Chaudhary, N., S. Singh, and R. Nagaraj. 2009. Morphology of self-assembled structures formed by short peptides from the amyloidogenic protein τ depends on the solvent in which the peptides are dissolved. *J. Pept. Sci.* 15:675–684.

11. Srisailam, S., T. K. S. Kumar, ..., C. Yu. 2003. Amyloid-like fibril formation in an all β -barrel protein. Partially structured intermediate state(s) is a precursor for fibril formation. *J. Biol. Chem.* 278:17701–17709.
12. Pallarès, I., J. Vendrell, ..., S. Ventura. 2004. Amyloid fibril formation by a partially structured intermediate state of α -chymotrypsin. *J. Mol. Biol.* 342:321–331.
13. Lim, K. H., Y. T. Le, ..., J. M. Kenney. 2010. Characterization of amyloidogenic intermediate states through a combined use of CD and NMR spectroscopy. *Biophys. Chem.* 151:155–159.
14. Calamai, M., F. Chiti, and C. M. Dobson. 2005. Amyloid fibril formation can proceed from different conformations of a partially unfolded protein. *Biophys. J.* 89:4201–4210.
15. Plakoutsi, G., N. Taddei, ..., F. Chiti. 2004. Aggregation of the acylphosphatase from *Sulfolobus solfataricus*: the folded and partially unfolded states can both be precursors for amyloid formation. *J. Biol. Chem.* 279:14111–14119.
16. Soldi, G., F. Bemporad, ..., F. Chiti. 2005. Amyloid formation of a protein in the absence of initial unfolding and destabilization of the native state. *Biophys. J.* 89:4234–4244.
17. Yanagi, K., M. Ashizaki, ..., Y. Goto. 2011. Hexafluoroisopropanol induces amyloid fibrils of islet amyloid polypeptide by enhancing both hydrophobic and electrostatic interactions. *J. Biol. Chem.* 286:23959–23966.
18. Hirota, N., K. Mizuno, and Y. Goto. 1997. Cooperative α -helix formation of β -lactoglobulin and melittin induced by hexafluoroisopropanol. *Protein Sci.* 6:416–421.
19. Gast, K., A. Siemer, ..., G. Damaschun. 2001. Fluoroalcohol-induced structural changes of proteins: some aspects of cosolvent-protein interactions. *Eur. Biophys. J.* 30:273–283.
20. Munishkina, L. A., C. Phelan, ..., A. L. Fink. 2003. Conformational behavior and aggregation of α -synuclein in organic solvents: modeling the effects of membranes. *Biochemistry.* 42:2720–2730.
21. Hsu, S. T., G. Blaser, and S. E. Jackson. 2009. The folding, stability and conformational dynamics of β -barrel fluorescent proteins. *Chem. Soc. Rev.* 38:2951–2965.
22. Cormack, B. P., R. H. Valdivia, and S. Falkow. 1996. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene.* 173(1 Spec No):33–38.
23. Espinosa, S., E. Bosch, ..., K. Valkó. 2002. Change of mobile phase pH during gradient reversed-phase chromatography with 2,2,2-trifluoroethanol-water as mobile phase and its effect on the chromatographic hydrophobicity index determination. *J. Chromatogr. A.* 954:77–87.
24. Zagorski, M. G., and C. J. Barrow. 1992. NMR studies of amyloid β -peptides: proton assignments, secondary structure, and mechanism of an α -helix- β -sheet conversion for a homologous, 28-residue, N-terminal fragment. *Biochemistry.* 31:5621–5631.
25. Tiffany, M. L., and S. Krimm. 1969. Circular dichroism of the “random” polypeptide chain. *Biopolymers.* 8:347–359.
26. Visser, N. V., M. A. Hink, ..., A. J. Visser. 2002. Circular dichroism spectroscopy of fluorescent proteins. *FEBS Lett.* 521:31–35.
27. Luo, P., and R. L. Baldwin. 1997. Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry.* 36:8413–8421.
28. Fan, F., and K. H. Mayo. 1995. Effect of pH on the conformation and backbone dynamics of a 27-residue peptide in trifluoroethanol. An NMR and CD study. *J. Biol. Chem.* 270:24693–24701.
29. Valerio, M., F. Porcelli, ..., F. Conti. 2008. pH effects on the conformational preferences of amyloid β -peptide (1–40) in HFIP aqueous solution by NMR spectroscopy. *ChemMedChem.* 3:833–843.
30. Kuznetsova, I. M., K. K. Turoverov, and V. N. Uversky. 2004. Use of the phase diagram method to analyze the protein unfolding-refolding reactions: fishing out the “invisible” intermediates. *J. Proteome Res.* 3:485–494.
31. Klotz, I. M., and R. M. Rosenberg. 1972. *Chemical Thermodynamics: Basic Theory and Methods*, 3rd Ed. W. A. Benjamin, editor. Menlo Park, CA.
32. Drozdov, A. N., A. Grossfield, and R. V. Pappu. 2004. Role of solvent in determining conformational preferences of alanine dipeptide in water. *J. Am. Chem. Soc.* 126:2574–2581.
33. Chatterjee, C., and J. T. Gerig. 2007. Interactions of trifluoroethanol with the Trp-cage peptide. *Biopolymers.* 87:115–123.
34. Díaz, M. D., M. Fioroni, ..., S. Berger. 2002. Evidence of complete hydrophobic coating of bombesin by trifluoroethanol in aqueous solution: an NMR spectroscopic and molecular dynamics study. *Chemistry.* 8:1663–1669.
35. Fioroni, M., M. D. Diaz, ..., S. Berger. 2002. Solvation phenomena of a tetrapeptide in water/trifluoroethanol and water/ethanol mixtures: a diffusion NMR, intermolecular NOE, and molecular dynamics study. *J. Am. Chem. Soc.* 124:7737–7744.
36. Walgers, R., T. C. Lee, and A. Cammers-Goodwin. 1998. An indirect chaotropic mechanism for the stabilization of helix conformation of peptides in aqueous trifluoroethanol and hexafluoro-2-propanol. *J. Am. Chem. Soc.* 120:5073–5079.
37. Uversky, V. N., N. V. Narizhneva, ..., G. Löber. 1997. Conformational transitions provoked by organic solvents in β -lactoglobulin: can a molten globule-like intermediate be induced by the decrease in dielectric constant? *Fold. Des.* 2:163–172.
38. Chin, J. T., S. L. Wheeler, and A. M. Klibanov. 1994. On protein solubility in organic solvent. *Biotechnol. Bioeng.* 44:140–145.
39. Malavolta, L., M. R. Pinto, ..., C. R. Nakaie. 2006. Interpretation of the dissolution of insoluble peptide sequences based on the acid-base properties of the solvent. *Protein Sci.* 15:1476–1488.
40. Eggers, D. K. 2011. A bulk water-dependent desolvation energy model for analyzing the effects of secondary solutes on biological equilibria. *Biochemistry.* 50:2004–2012.
41. Rauscher, S., S. Baud, ..., R. Pomès. 2006. Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. *Structure.* 14:1667–1676.
42. Balbirnie, M., R. Grothe, and D. S. Eisenberg. 2001. An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated β -sheet structure for amyloid. *Proc. Natl. Acad. Sci. USA.* 98:2375–2380.
43. Zhang, J., and Y. B. Yan. 2008. Oligomerization and aggregation of bovine pancreatic ribonuclease A: backbone hydration probed by infrared band-shift. *Protein Pept. Lett.* 15:650–657.
44. Mukherjee, S., P. Chowdhury, and F. Gai. 2009. Effect of dehydration on the aggregation kinetics of two amyloid peptides. *J. Phys. Chem. B.* 113:531–535.
45. Monsellier, E., and F. Chiti. 2007. Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep.* 8:737–742.
46. Uversky, V. N., J. R. Gillespie, and A. L. Fink. 2000. Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins.* 41:415–427.
47. Putnam, C. D. 2006. Protein Calculator V.3.3. Web access 1 June 2011. <<http://www.scripps.edu/~cdputnam/protcalc.html>>.
48. Anderson, V. L., and W. W. Webb. 2011. Transmission electron microscopy characterization of fluorescently labeled amyloid β 1–40 and α -synuclein aggregates. *BMC Biotechnol.* 11:125.
49. Andrade, M. A., P. Chacón, ..., F. Morán. 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng.* 6:383–390.
50. Brahms, S., and J. Brahms. 1980. Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J. Mol. Biol.* 138:149–178.
51. Cattell, R. 1966. The screen test for the number of factors. *Multivariate Behav. Res.* 1:245–276.
52. Compton, L. A., and W. C. Johnson, Jr. 1986. Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem.* 155:155–167.

53. Enoki, S., K. Saeki, ..., K. Kuwajima. 2004. Acid denaturation and refolding of green fluorescent protein. *Biochemistry*. 43:14238–14248.
54. Greenfield, N. J. 1996. Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Anal. Biochem.* 235:1–10.
55. Provencher, S. W., and J. Glöckner. 1981. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry*. 20:33–37.
56. Spiess, A. N., and N. Neumeyer. 2010. An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. *BMC Pharmacol.* 10:6.
57. Sreerama, N., S. Y. Venyaminov, and R. W. Woody. 2000. Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis. *Anal. Biochem.* 287:243–251.
58. Sreerama, N., and R. W. Woody. 2000. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* 287:252–260.
59. van Stokkum, I. H., H. J. Spoelder, ..., F. C. Groen. 1990. Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Anal. Biochem.* 191:110–118.
60. Whitmore, L., and B. A. Wallace. 2004. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* 32(Web Server issue):W668–W673.
61. Whitmore, L., and B. A. Wallace. 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers.* 89:392–400.