Supporting Information

Amorphous aggregation of cytochrome *c* with inherently low amyloidogenicity is characterized by the metastability of supersaturation and the phase diagram

Yuxi Lin[1], József Kardos[2], Mizue Imai[3], Tatsuya Ikenoue[1], Misaki Kinoshita[1], Toshihiko Sugiki[1], Koichiro Ishimori[3,4,5], Yuji Goto[1], Young-Ho Lee[1,*]

[1]Institute for Protein Research, Osaka University, Yamadaoka 3-2, Suita, Osaka 565-0871, Japan

[2]MTA-ELTE NAP B Neuroimmunology Research Group, Department of Biochemistry, Eötvös Loránd University, Budapest H-1117, Hungary

[3]Graduate School of Chemical Sciences and Engineering, Hokkaido University, Kita 13, Nishi 8, Kita-ku, Sapporo 060-8628, Japan

[4]Division of Chemistry, Graduate School of Science, Hokkaido University, Kita 10, Nishi 8, Kita-ku, Sapporo 060-0810, Japan

[5]Department of Chemistry, Faculty of Science, Hokkaido University, Kita 10, Nishi 8, Kita-ku, Sapporo 060-0810, Japan

Corresponding Author

[*]mr0505@protein.osaka-u.ac.jp

Table of contents

Table S1. Secondary structure analysis of various forms of Cyt*c* by infrared spectroscopy

HoloCyt*c*														
No alcohol			20% TFE (10 min)			20% TFE (15 h)			10% HFIP (10 min)			10% HFIP (15 h)		
cm^{-1}	frac.[a]	str.[b]	cm^{-1}	frac.	str.	cm^{-1}	frac.	str.	cm^{-1}	frac.	str.	cm^{-1}	frac.	str.
1630	0.24	B	1628	0.13	B	1628	0.11	B	1630	0.07	B	1630	0.03	B
1642	0.18	R	1642	0.11	R	1641	0.11	R	1638	0.12	R/B	1637	0.28	R/B
1655	0.44	H	1651	0.64	H/R[c]	1650	0.66	H/R	1650	0.73	H/R	1652	0.61	H/R
1674	0.13	T	1675	0.12	T	1675	0.13	T	1677	0.08	T	1677	0.07	T
ApoCyt*c*														
No alcohol			30% TFE (10 min)			30% TFE (15 h)			10% HFIP (10 min)			10% HFIP (15 h)		
1635	0.09	B	1624	0.05	B	1617	0.11	B	1631	0.04	B	1632	0.10	B
1646	0.53	R	1632	0.12	B	1638	0.35	R/B	1640	0.07	R/B	1638	0.03	R/B
1652	0.19	H	1649	0.76	R/H	1647	0.14	R	1648	0.78	R/H	1645	0.26	R
1672	0.19	T	1675	0.12	T	1654	0.21	H	1676	0.11	T	1651	0.48	H/R
						1673	0.19	T				1674	0.13	T
						1694	0.01	AP						
Ag-apoCyt*c*														
No alcohol			10% TFE (10 min)			10% TFE (15 h)			4% HFIP (10 min)			4% HFIP (15 h)		
1638	0.52	R/B	1617	0.13	B	1617	0.19	B	1617	0.14	B	1617	0.24	B
1646	0.05	R	1638	0.39	R/B	1638	0.40	R/B	1638	0.33	R/B	1638	0.26	R/B
1652	0.15	H	1650	0.24	R/H	1650	0.21	R/H	1649	0.32	R/H	1648	0.27	R/H
1669	0.28	T	1669	0.23	T	1671	0.19	T	1670	0.21	T	1654	0.01	H
1684	0.01	AP	1684	0.01	AP	1694	0.01	AP	1685	0.01	AP	1672	0.22	T
												1686	0.01	AP

Experiments were performed in D_2O containing 25 mM sodium acetate (pD 4.8) in the absence or presence of the given percentage of TFE or HFIP. Baseline- and vapor-corrected spectra were fit by Gaussians positioned at wavenumbers determined from the second derivatives of the spectra. [a]The fraction of the component and [b]assigned secondary structure are presented. The following abbreviations were used: B, β-sheet; H, α-helix; R, random coil; T, turn; AP, high frequency component suggesting the presence of an antiparallel β-sheet. [c]In several cases, it was difficult to distinguish a β-sheet from a random coil or a helix from a random coil when they were largely overlapping. In such cases, both structures were indicated.

Table S2. Summary of secondary structure contents obtained by CD spectroscopy

			H (%)	AP (%)	P (%)	Others (%)
HoloCyt*c*	No alcohol		25	27	0	48
	TFE	10%	27	25	0	48
		20%	40	9	2	49
		40%	44	6	0	50
		60%	47	5	0	48
		70%	52	7	0	41
	HFIP	4%	31	22	1	46
		8%	n.d.	n.d.	n.d.	n.d.
		10%	n.d.	n.d.	n.d.	n.d.
		12%	n.d.	n.d.	n.d.	n.d.
		16%	n.d.	n.d.	n.d.	n.d.
		20%	54	6	0	40
		40%	46	2	0	52
		70%	55	6	0	39
ApoCyt*c*	No alcohol		2	18	0	80
	TFE	10%	10	19	0	71
		20%	25	9	6	60
		30%	n.d.	n.d.	n.d.	n.d.
		50%	n.d.	n.d.	n.d.	n.d.
		70%	n.d.	n.d.	n.d.	n.d.
	HFIP	4%	13	14	7	66
		8%	n.d.	n.d.	n.d.	n.d.
		10%	n.d.	n.d.	n.d.	n.d.
		12%	n.d.	n.d.	n.d.	n.d.
		20%	n.d.	n.d.	n.d.	n.d.
		40%	45	6	0	49
		70%	46	7	0	47
Ag-apoCyt*c*	No alcohol		4	20	0	76
	TFE	10%	10	23	10	57
		20%	22	21	6	51
		40%	30	11	0	59
		50%	34	14	0	52
		60%	30	15	0	55
		4%	13	22	9	56

	HFIP	8%	n.d.	n.d.	n.d.	n.d.
		12%	n.d.	n.d.	n.d.	n.d.
		14%	n.d.	n.d.	n.d.	n.d.
		20%	47	4	2	47
		40%	45	2	2	51
		60%	46	2	1	51

The secondary structure contents of the three types of Cyt*c*s after the ultrasonication treatment in 25 mM sodium acetate buffer (pH 4.8) with and without the given percentage of TFE or HFIP were predicted by the BeStSel algorithm.[1] The following abbreviations were used: H, α-helix; P, parallel β-sheet; AP, antiparallel β-sheet; Others include all non-helical and non β-sheet components, mainly disordered.

Table S3. Summary of the degree of supersaturation of protein solutions

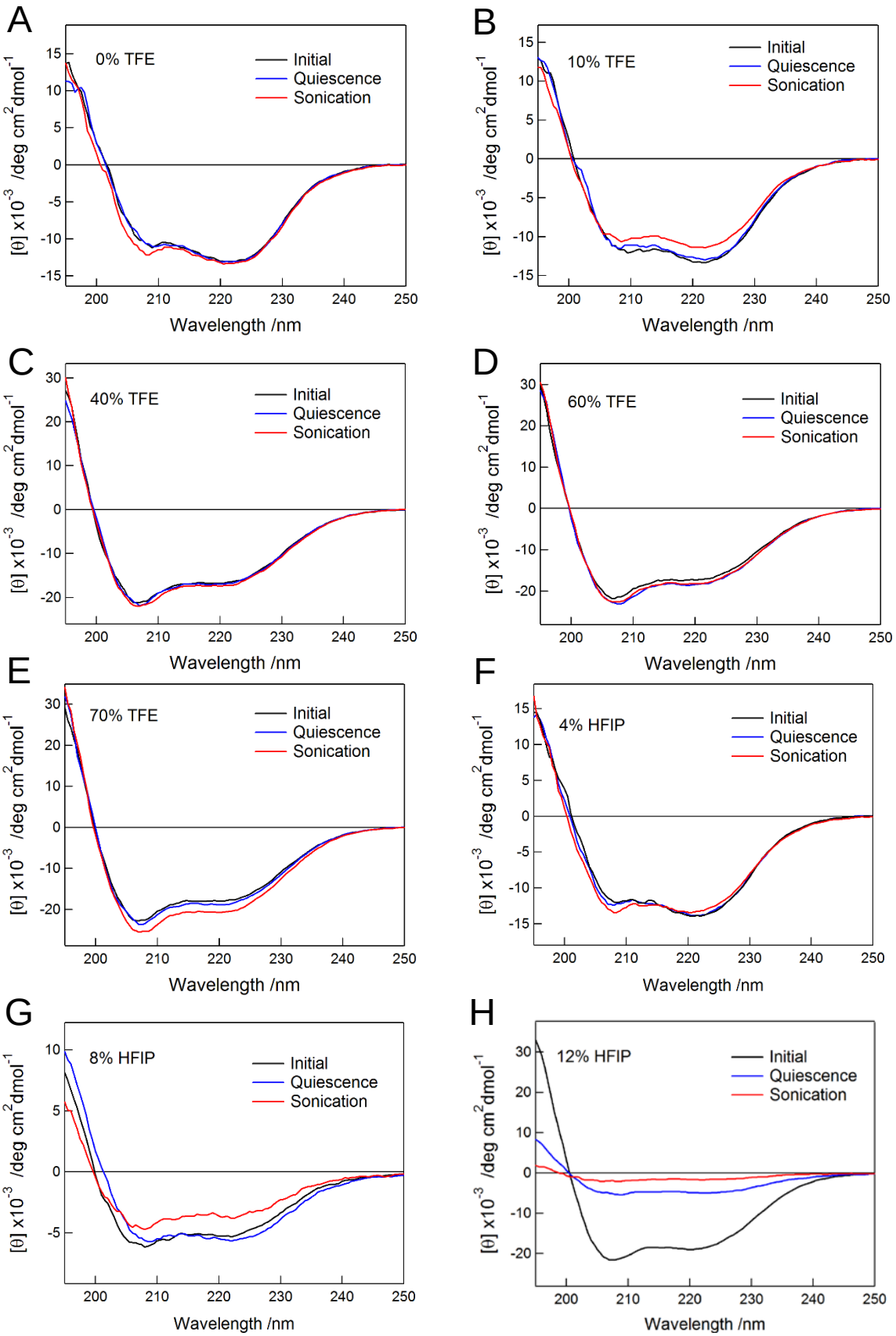
			Degree of supersaturation (σ)
HoloCyt*c*	No alcohol		n.d.
	TFE	10%	n.d.
		20%	n.d.
		40%	n.d.
		60%	n.d.
		70%	n.d.
	HFIP	4%	n.d.
		8%	6.9
		10%	8.7
		12%	15.7
		16%	1.7
		20%	n.d.
		40%	n.d.
		70%	n.d.
ApoCyt*c*	No alcohol		n.d.
	TFE	10%	n.d.
		20%	n.d.
		30%	0.2
		50%	0.3
		70%	0.5
	HFIP	4%	n.d.
		8%	4.4
		10%	7.3
		12%	9.0
		20%	0.4
		40%	n.d.
		70%	n.d.
Ag-apoCyt*c*	No alcohol		n.d.
	TFE	10%	0.1
		20%	0.3
		40%	n.d.
		50%	n.d.
		60%	n.d.
		4%	0.2

Lysozyme	HFIP	8%	14.0
		12%	10.8
		14%	7.3
		20%	n.d.
		40%	n.d.
		60%	n.d.
	No alcohol		n.d.
	TFE	10%	n.d.
		20%	n.d.
		30%	n.d.
		40%	36.5
		50%	74.0
		60%	74.0
		70%	79.0
		80%	n.d.
		90%	1.6
Insulin	No alcohol		23.8
	TFE	10%	66.3
		20%	47.8
		30%	156.1
		40%	1413.2
		50%	87.4
		60%	n.d.
		70%	51.4

The degree of supersaturation (σ) of the three types of Cyt*c*s, lysozyme, and insulin in 25 mM sodium acetate buffer (pH 4.8) with and without the given percentage of TFE or HFIP was predicted by the following equation:

$$\sigma = (C - C^*) / C^*$$

where C and C^* are the protein concentration used and the protein solubility, respectively. C^* corresponds to the concentration of residual protein monomers at the end of aggregation reaction. “n.d.” is shown in cases, in which the protein solutions are unsaturated or the concentration of residual monomers cannot be determined. Our previous results were used for the calculation of lysozyme[2] and insulin.[3]


A
0% TFE
Initial
Quiescence
Sonication
B
10% TFE
C
40% TFE
D
60% TFE
E
70% TFE
F
4% HFIP
G
8% HFIP
H
12% HFIP
[θ] x10^-3 /deg cm^2dmol^-1
Wavelength /nm

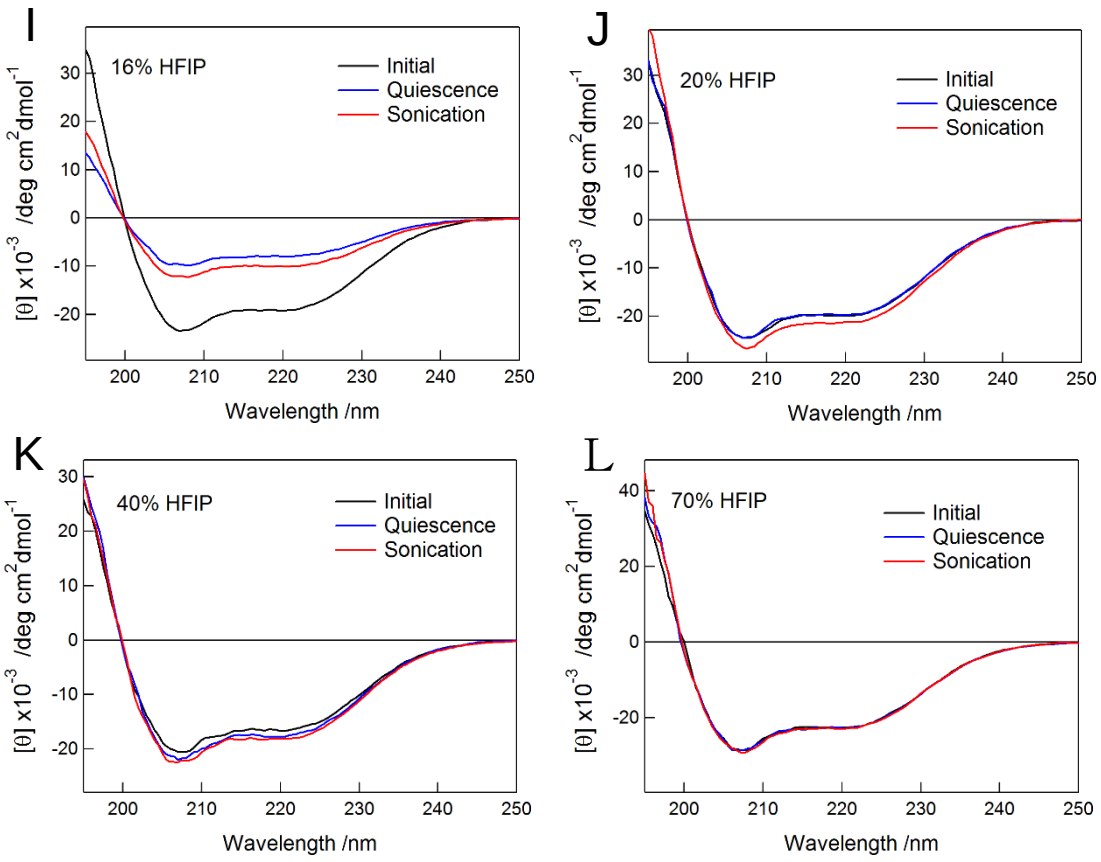


Figure S1. Alcohol-dependent aggregation of holoCyt*c* monitored by far-UV CD. Far-UV CD spectra of holoCyt*c* at ~5 mins (black lines) and 15 h when incubated without (blue lines) and with ultrasonication (red lines) after sample preparation were determined in the presence of TFE concentrations of 0% (A), 10% (B), 40% (C), 60% (D), 70% (E); or HFIP concentrations of 4% (F), 8% (G), 12% (H), 16% (I), 20% (J), 40% (K), and 70% (L).

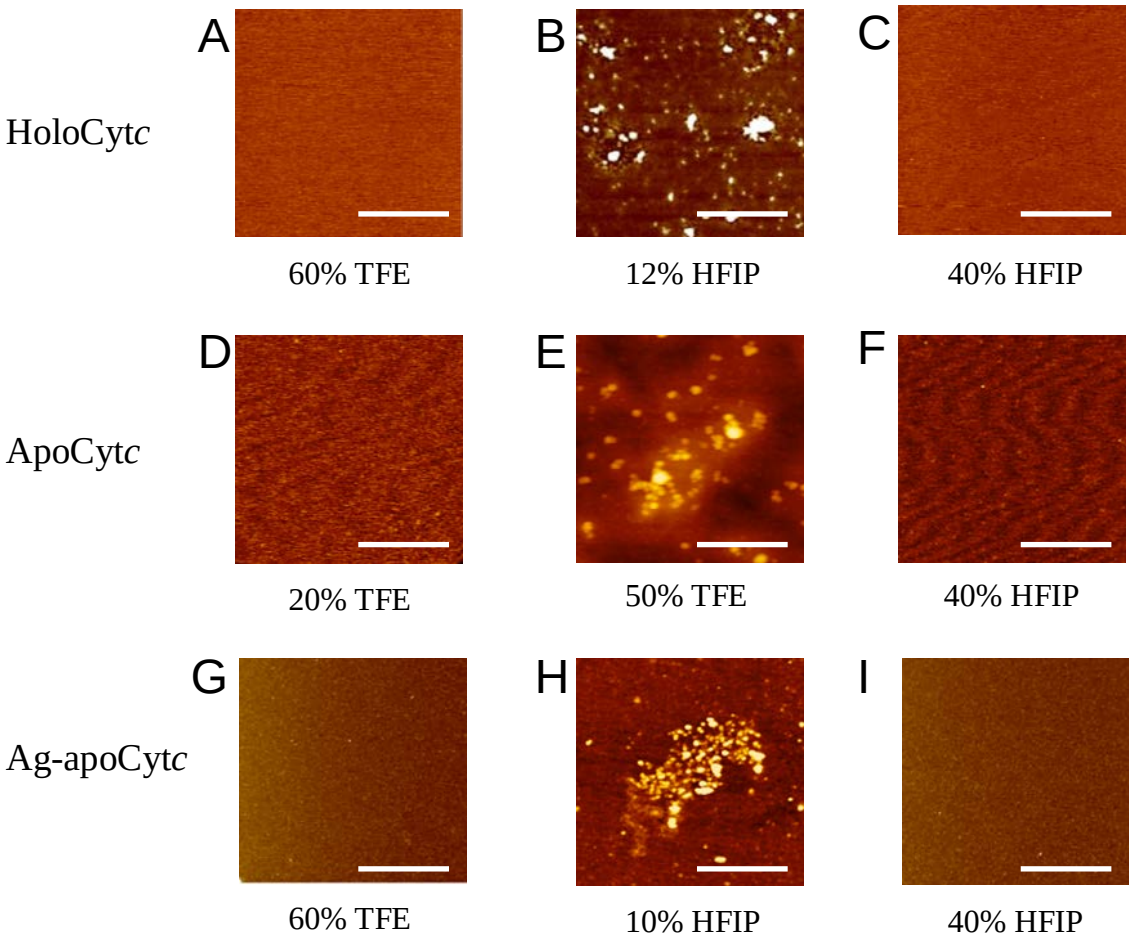


Figure S2. Morphological characterization of the three types of Cyt*c* at water/alcohol mixtures by AFM. (A-I) AFM images of Cyt*c* under distinct alcohol conditions after treatment with ultrasonication are shown for holoCyt*c* (A-C), apoCyt*c* (D-F), and Ag-apoCyt*c* (G-I). The concentrations of TFE and HFIP used are displayed below the AFM images. The white scale bars correspond to 1 μm.

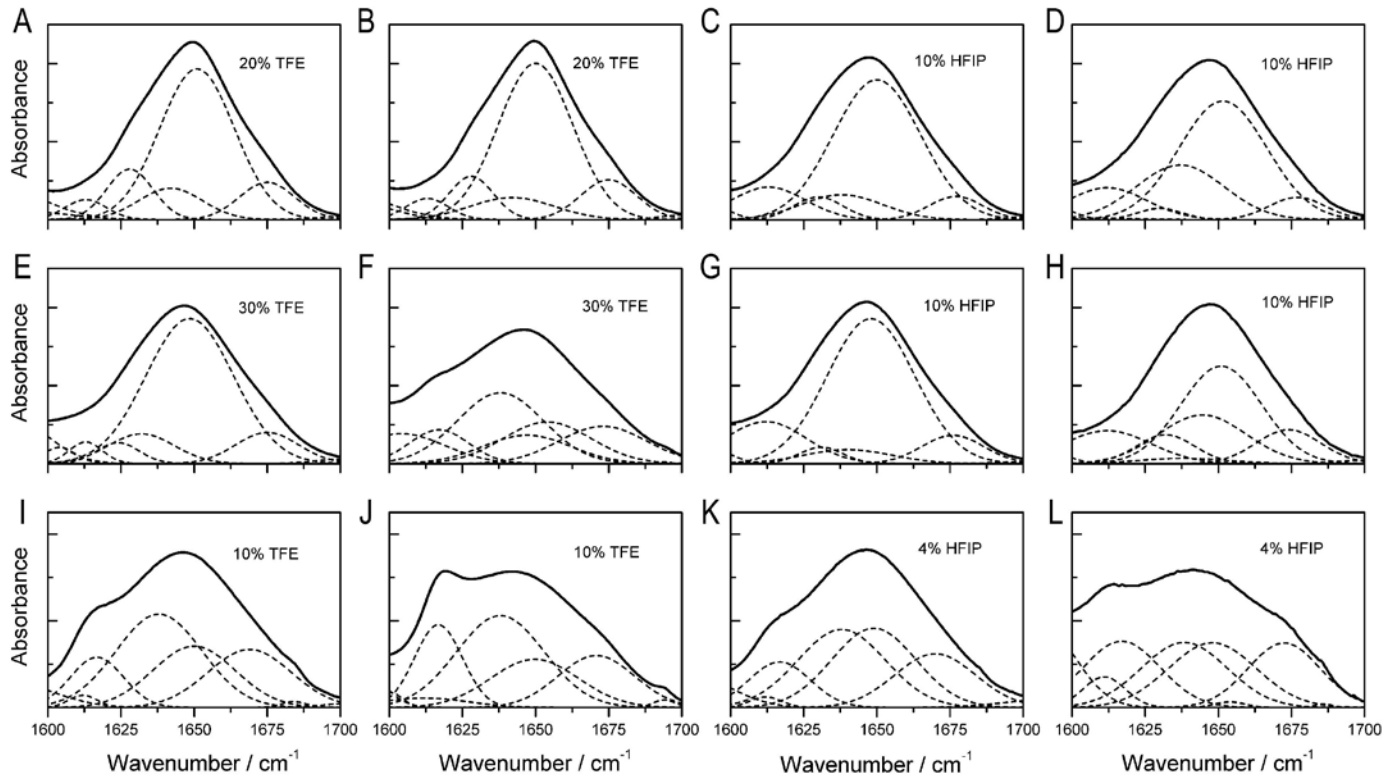
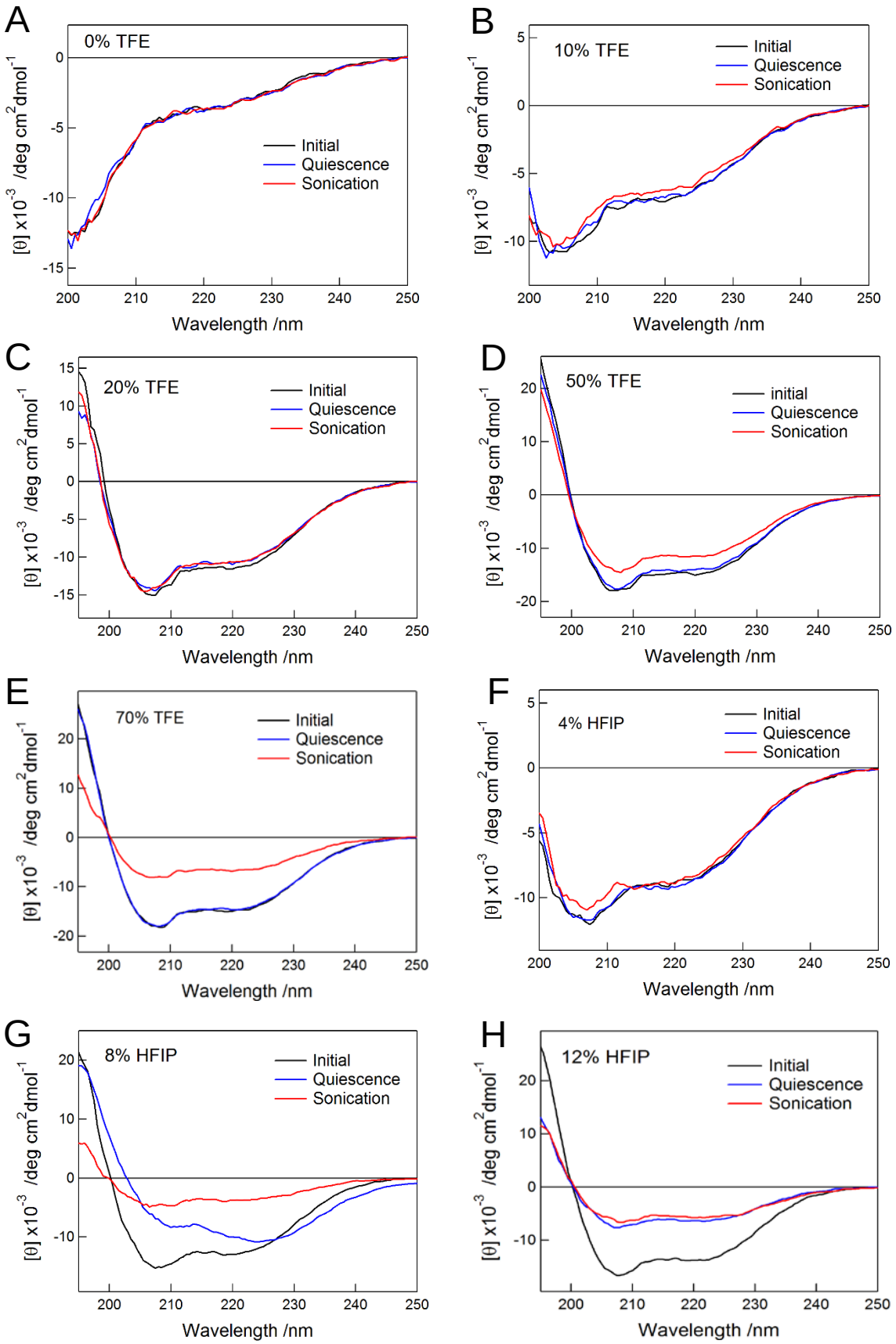


Figure S3. Investigation of the secondary structures of different forms of Cyt*c* by infrared spectroscopy. (A-D) HoloCyt*c*, (E-H) apoCyt*c*, and (I-L) Ag-apoCyt*c* in the presence of TFE and HFIP. (A, C, E, G, I, and J) present infrared spectra measured promptly after sample preparation (solid lines) and their deconvolution to Gaussian components (dashed lines). (B, D, F, H, J, and L) shows the spectra of samples after a 15-h incubation with ultrasonication and their deconvolution. TFE or HFIP concentrations are indicated. Fractions of the components and their secondary structure assignments are summarized in Table S1.


A
0% TFE
Initial
Quiescence
Sonication
B
10% TFE
Initial
Quiescence
Sonication
C
20% TFE
Initial
Quiescence
Sonication
D
50% TFE
initial
Quiescence
Sonication
E
70% TFE
Initial
Quiescence
Sonication
F
4% HFIP
Initial
Quiescence
Sonication
G
8% HFIP
Initial
Quiescence
Sonication
H
12% HFIP
Initial
Quiescence
Sonication
[θ] x10^-3 /deg cm^2dmol^-1
Wavelength /nm

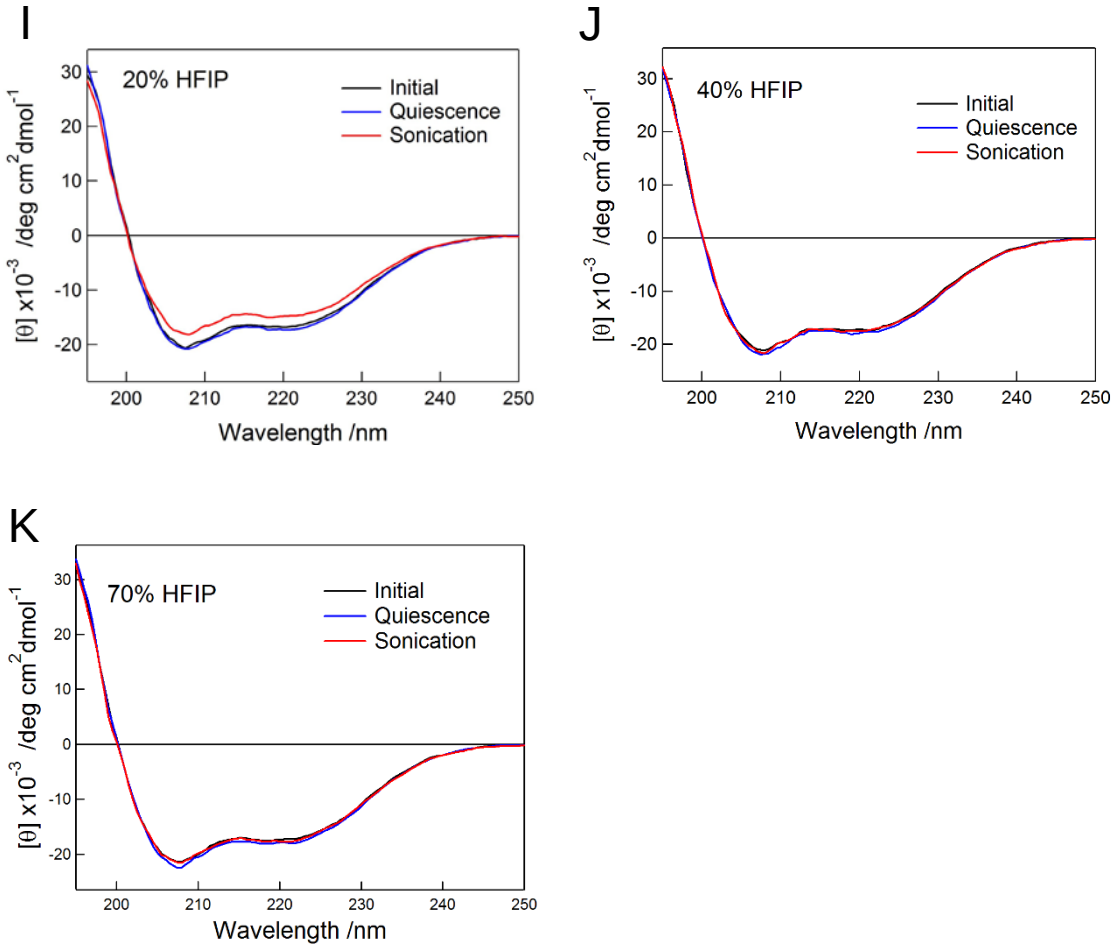
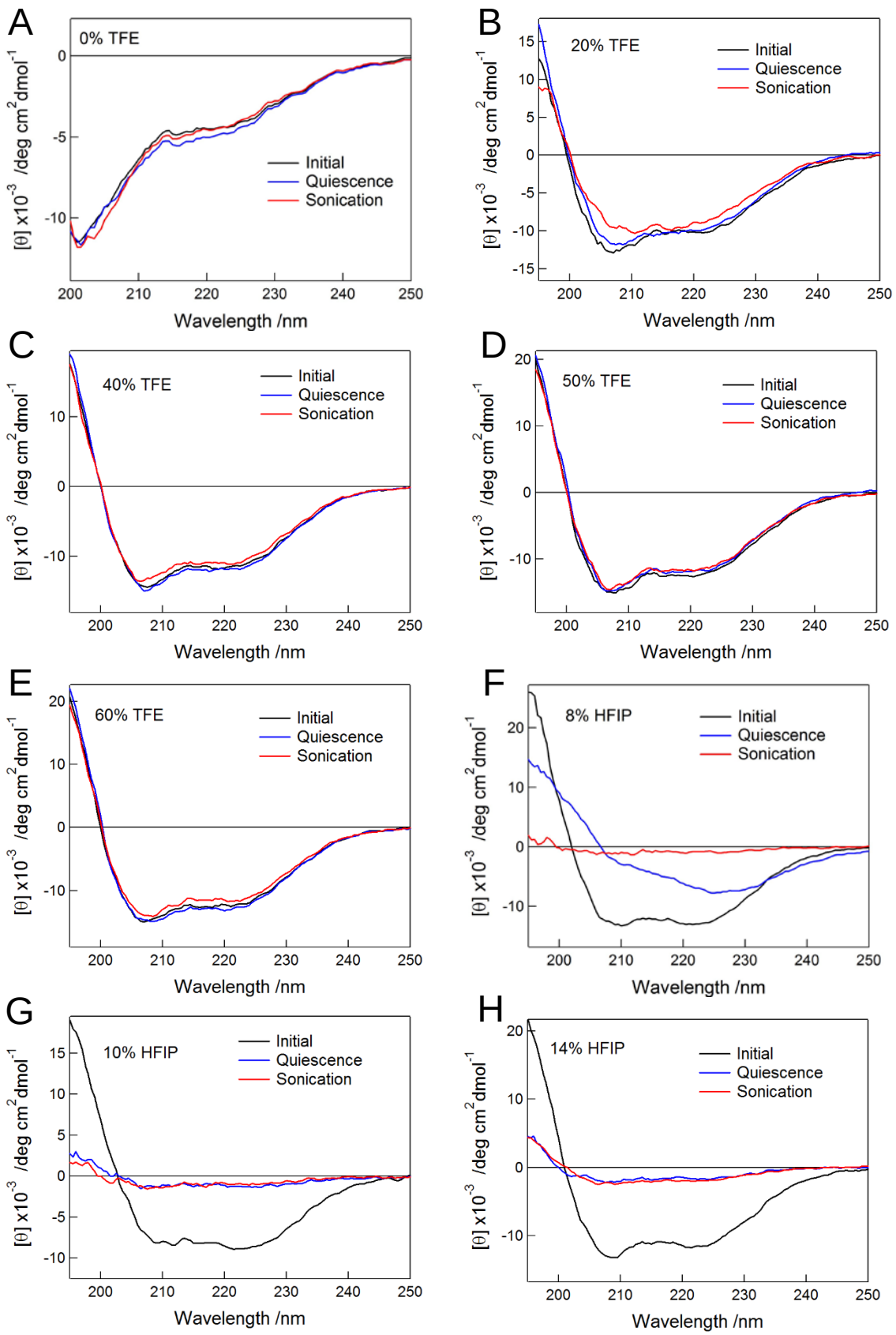


Figure S4. Alcohol-dependent aggregation of apoCyt*c* monitored by far-UV CD. Far-UV CD spectra of apoCyt*c* at ~5 mins (black lines) and 15 h when incubated without (blue lines) and with ultrasonication (red lines) after sample preparation were determined in the presence of TFE concentrations of 0% (A), 10% (B), 20% (C), 50% (D), 70% (E); or HFIP concentrations of 4% (F), 8% (G), 12% (H), 20% (I), 40% (J), and 70% (K).


A
0% TFE
Initial
Quiescence
Sonication
B
20% TFE
C
40% TFE
D
50% TFE
E
60% TFE
F
8% HFIP
G
10% HFIP
H
14% HFIP
[θ] x10^-3 /deg cm^2dmol^-1
Wavelength /nm

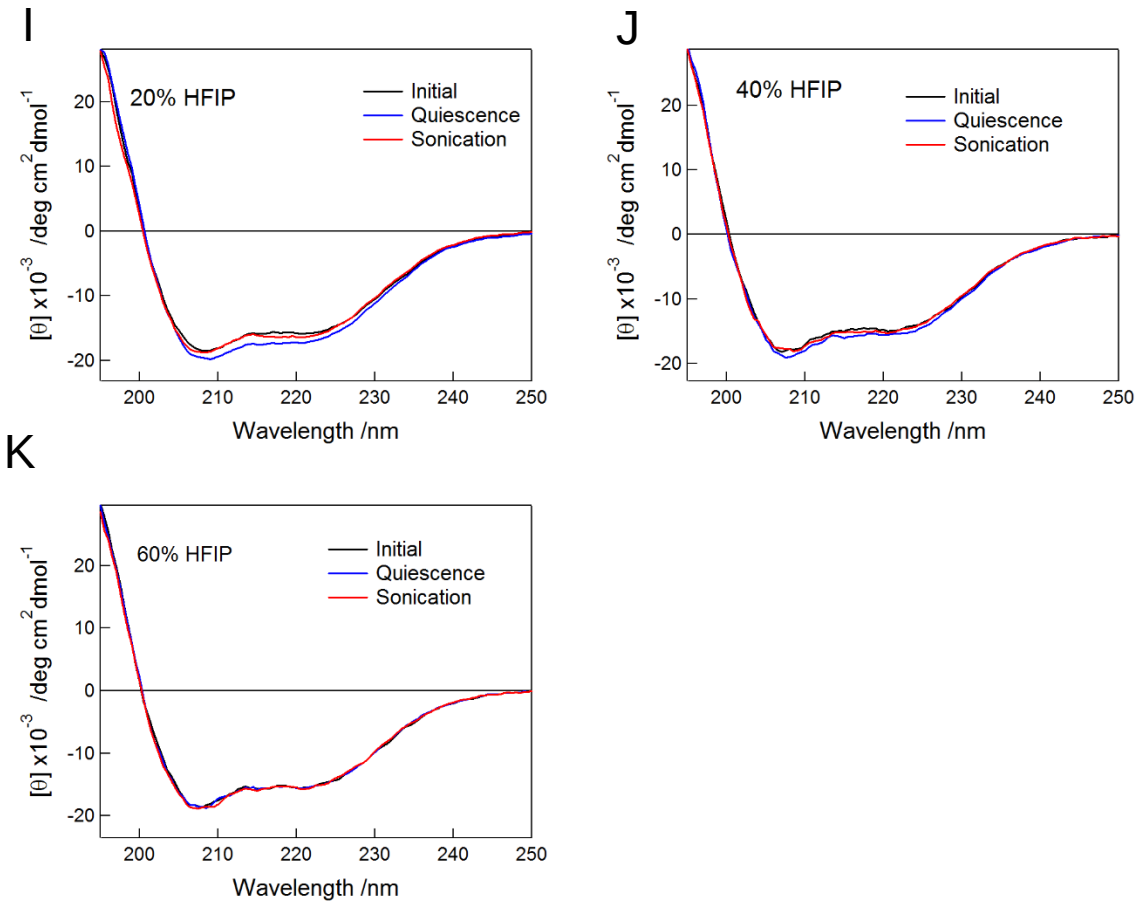


Figure S5. Alcohol-dependent aggregation of Ag-apoCyt*c* monitored by far-UV CD. Far-UV CD spectra of Ag-apoCyt*c* at ~5 mins (black lines) and 15 h when incubated without (blue lines) and with ultrasonication (red lines) after sample preparation were determined in the presence of TFE concentrations of 0% (A), 20% (B), 40% (C), 50% (D), 60% (E); or HFIP concentrations of 8% (F), 10% (G), 14% (H), 40% (I), 20% (J), and 70% (K).

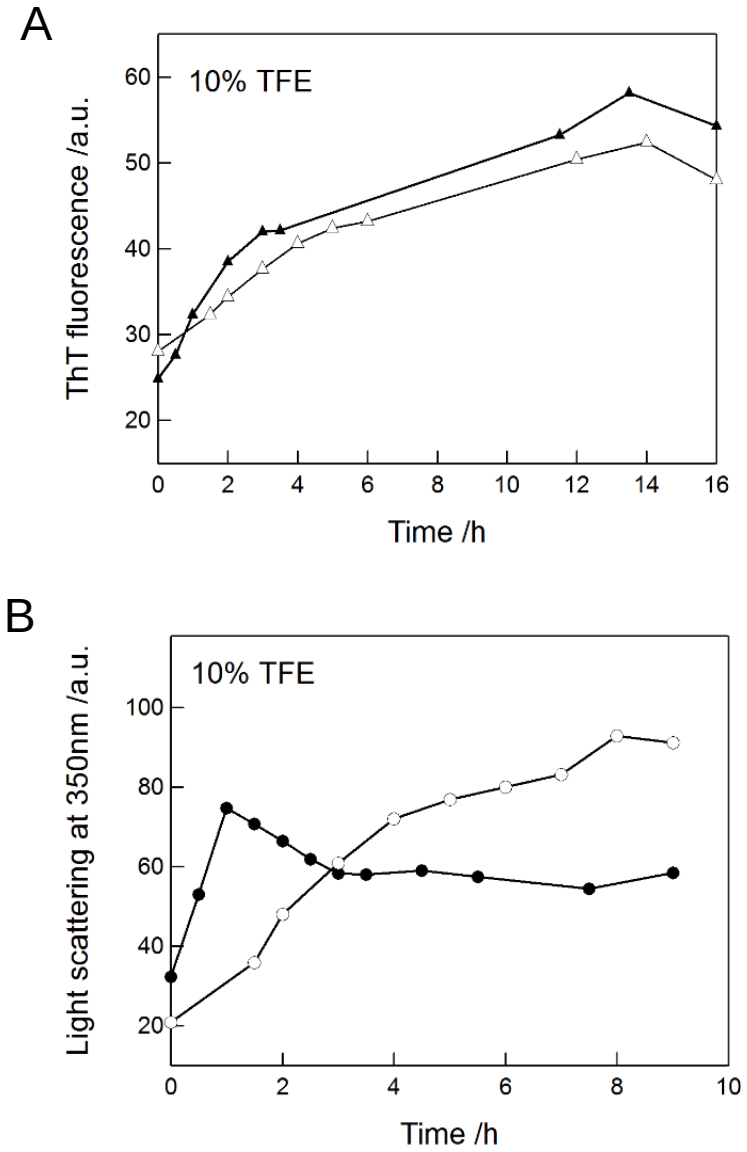


Figure S6. Aggregation kinetics of Ag-apoCyt*c* at a TFE concentration of 10%. The process of Ag-apoCyt*c* aggregating into protofibrils was monitored by ThT fluorescence (A) and light scattering (B) with (▲, ●) and without sonication (Δ, ○).

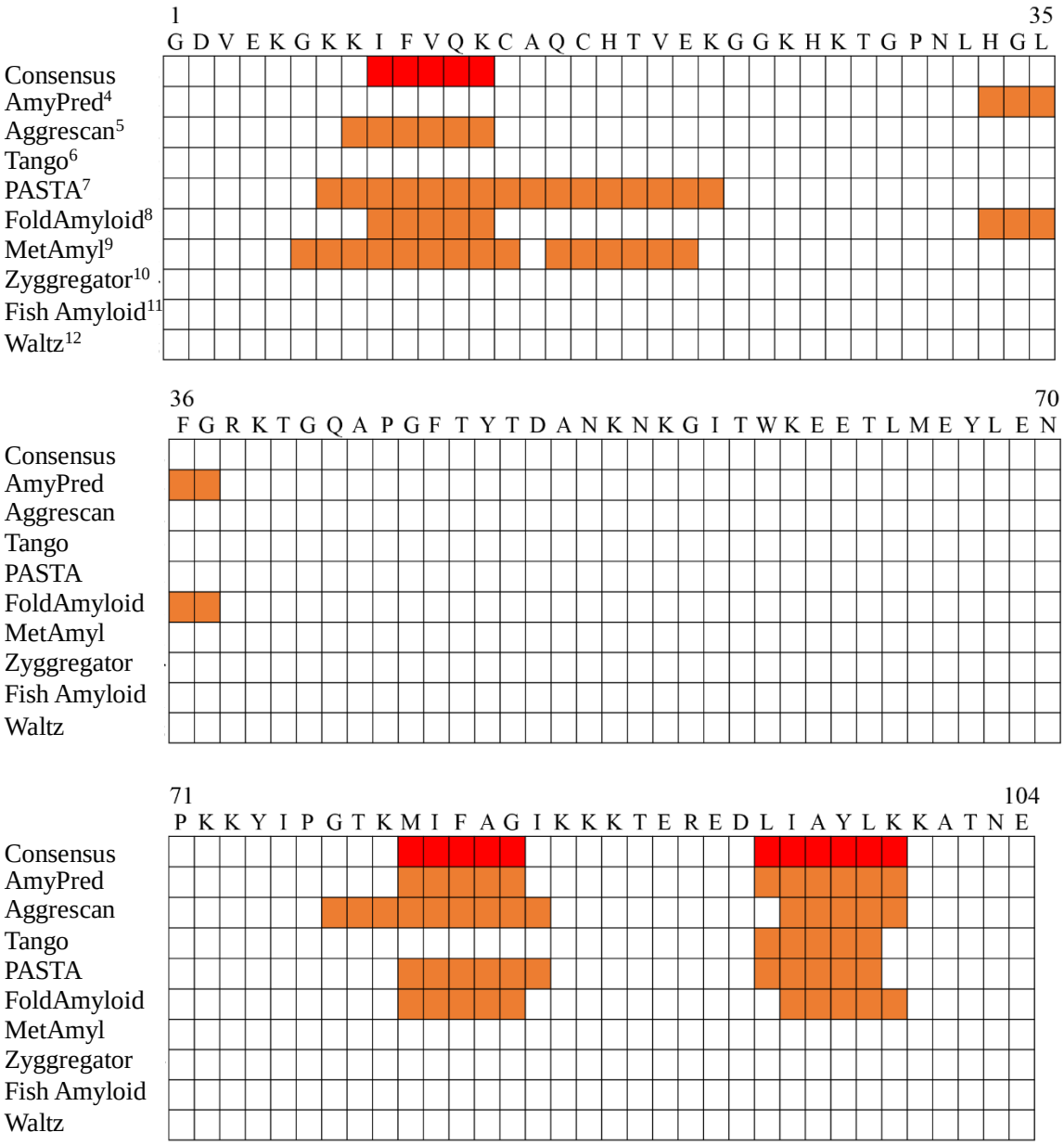


Figure S7. Predictions of aggregation-prone regions in horse heart Cyt*c* using various computational algorithms. An orange color indicates the predicted regions prone to form amyloid fibrils. Default values and thresholds were used for predictions. A red color indicates the consensus region with amyloidgenicity between different algorithms.

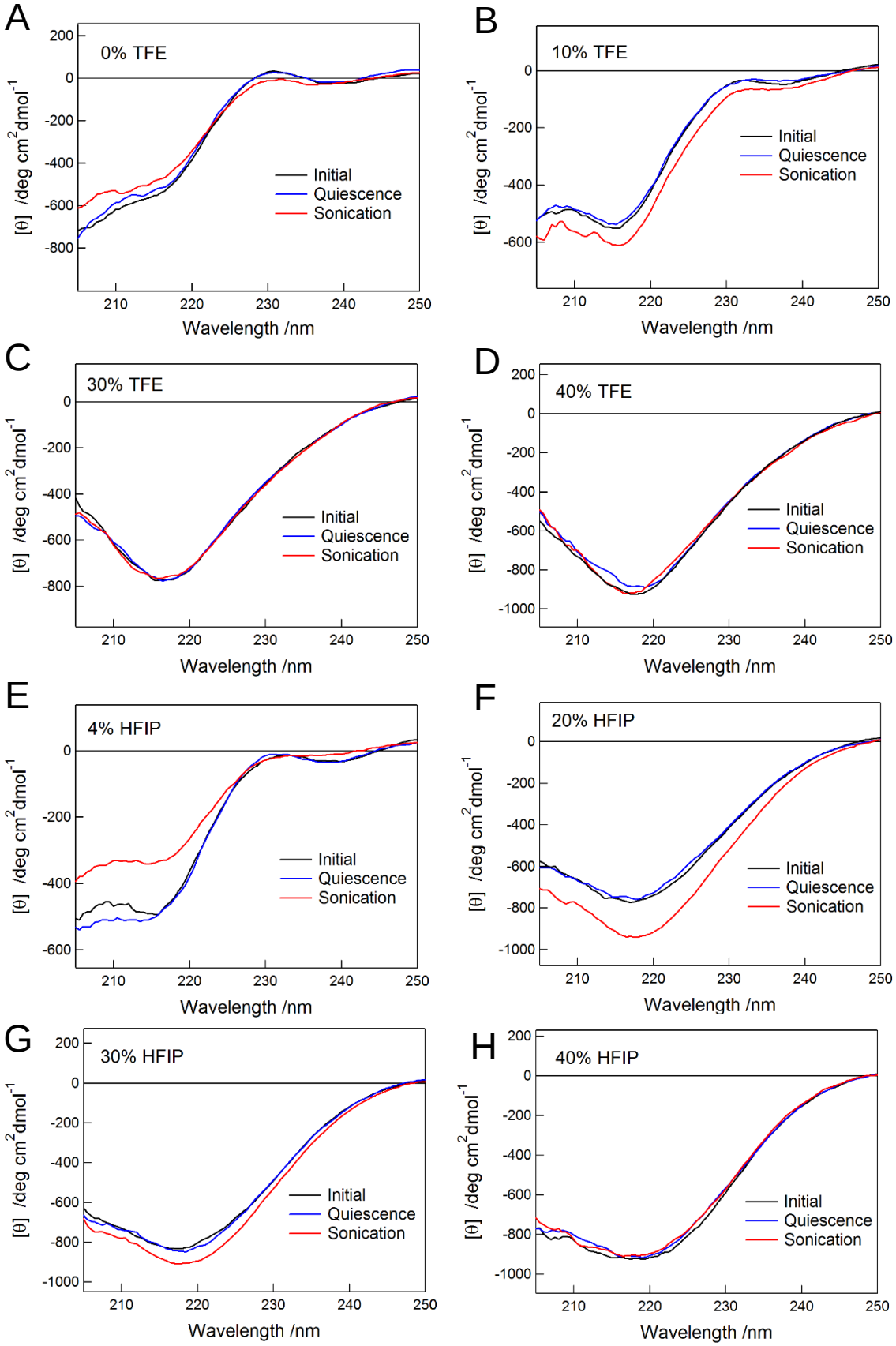

A
0% TFE
B
10% TFE
C
30% TFE
D
40% TFE
E
4% HFIP
F
20% HFIP
G
30% HFIP
H
40% HFIP
Initial
Quiescence
Sonication
[θ] /deg cm2dmol-1
Wavelength /nm

Figure S8. Alcohol-dependent aggregation of LIAYLK monitored by far-UV CD. Far-UV CD spectra of LIAYLK at ~5 mins (black lines) and 15 h when incubated without (blue lines) and with ultrasonication (red lines) after samples preparation were determined in the presence of TFE concentrations of 0% (A), 10% (B), 30% (C), 40% (D); or HFIP concentrations of 4% (E), 20% (F), 30% (G), and 40% (H).

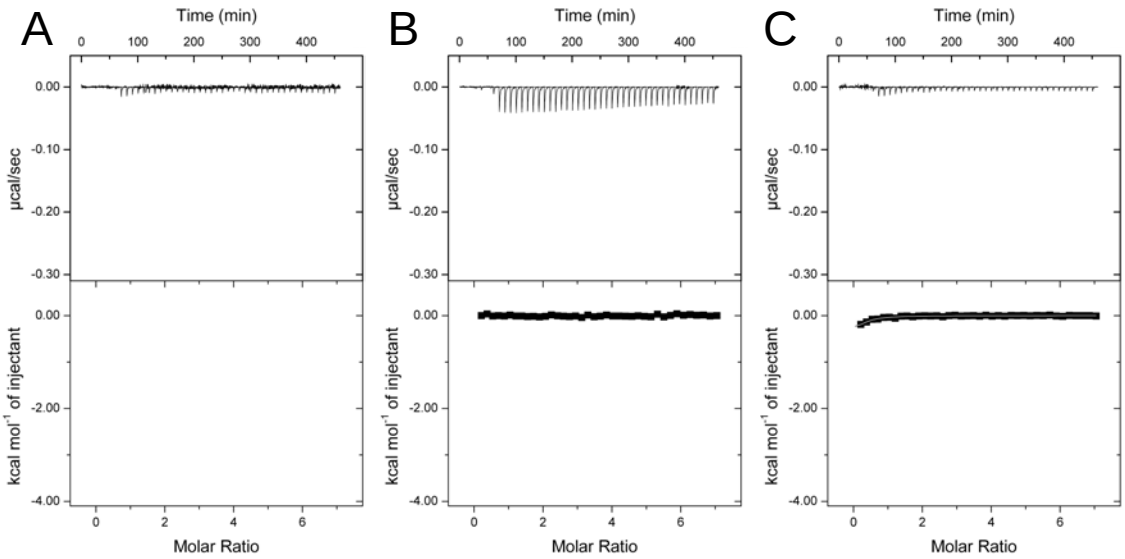


Figure S9. Intermolecular interactions between the two types of Cyt*c* and iron examined by isothermal titration calorimetry. (A-C) ITC thermograms of the titration of Fe^{3+} to 25 mM sodium acetate buffer (pH 4.8) (A), holoCyt*c* (B), and apoCyt*c* (C) are shown in the upper panel. Normalized heat values were plotted against the molar ratio ([Fe^{3+}]/[Cyt*c*]) in the lower panel. Thermodynamic parameters of iron binding to Cyt*c* could not be determined due to too weak intermolecular interaction.

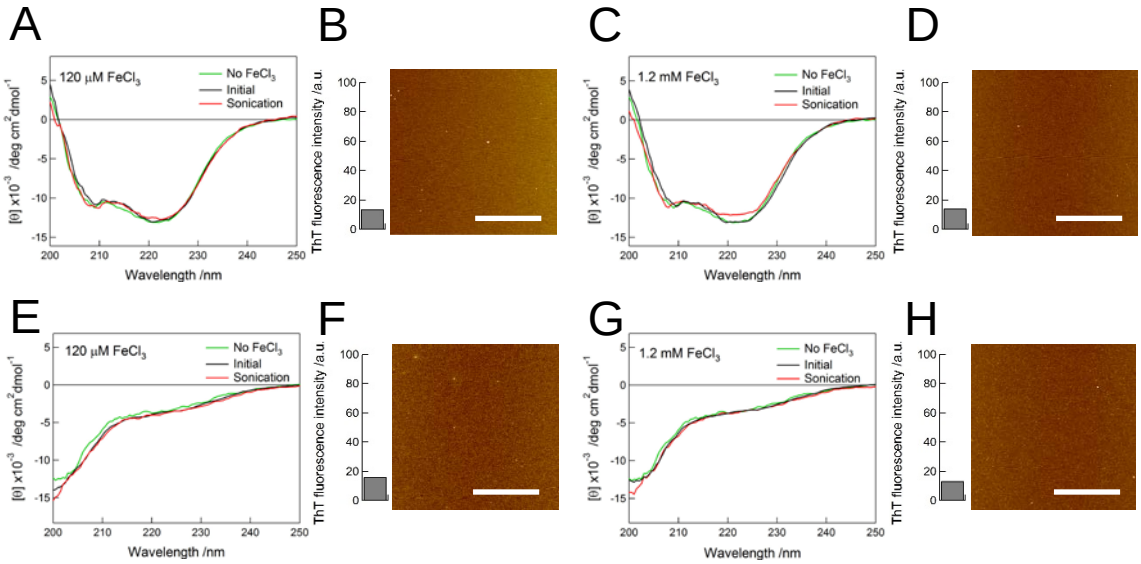


Figure S10. Effects of iron on the aggregation of the physiologically relevant two types of Cyt*c* monitored by far-UV CD, ThT fluorescence, and AFM. (A, C, E, and G) Far-UV CD spectra of holoCyt*c* (A, C) and apoCyt*c* (E, G) in the presence of 120 μM (A and E) or 1.2 mM (C and G) iron ions are shown. The CD spectra were recorded at ~5 mins (black lines) and 15 h when incubated with ultrasonication (red lines) after sample preparation. For comparison, the spectra of each Cyt*c* in the absence of iron ions are also presented (green line). The concentration of Cyt*c* for incubation was 120 μM. (B, D, F, and H) The ThT fluorescence (left) and AFM images and of Cyt*c* (right) in the presence of 120 μM (B and F) or 1.2 mM (D and H) irons after 15 h incubation with ultrasonication. The white scale bars indicate 1 μm.

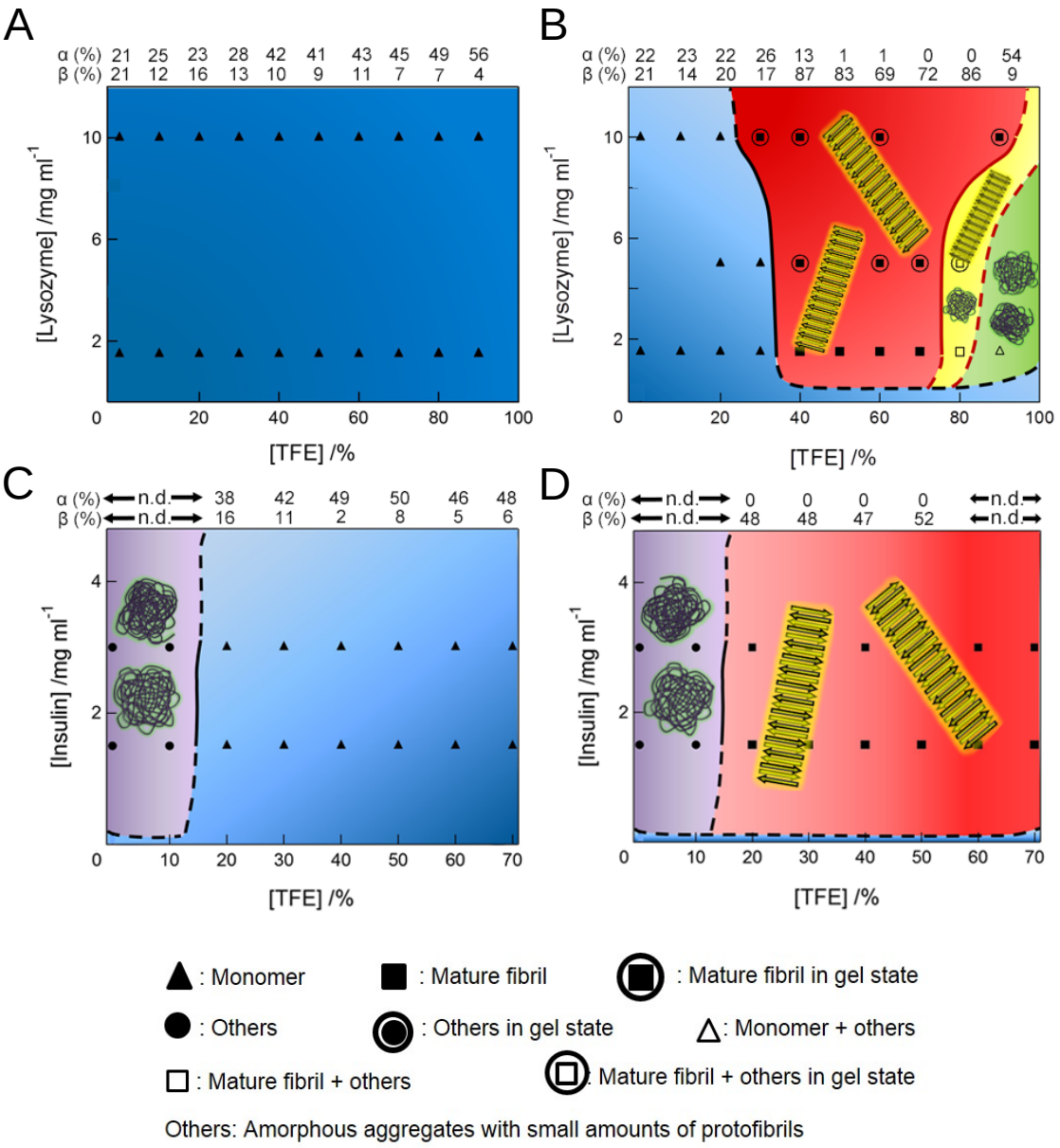


Figure S11. Phase diagrams of the aggregation of lysozyme and insulin in TFE/water mixtures. Phase diagrams of lysozyme (A and B) and insulin (C and D) at pH 4.8 before (A and C) and after ultrasonication (B and D) in TFE/water mixtures. Blue, monomers; red, amyloid fibrils; yellow, dominantly mature fibrils with a small quantity of amorphous

aggregates and protofibrils; Magenta, amorphous aggregates; and green, monomers with a small quantity of protofibrils. Dominant species detected at various TFE concentrations are indicated by symbols. The secondary structure contents of lysozyme and insulin, predicted by the BeStSel algorithm[1] from far-UV CD spectra, at 1.5 mg ml^{-1} are noted at the top of the phase diagrams. α and β indicate the α-helical and β-sheet contents, respectively. Ambiguous results due to aggregation and precipitation are marked with "n.d." The phase diagram of lysozyme after ultrasonication (B) was reproduced with slight modifications[2] and the figures (A, C, and D) were newly constructed based on our previous findings.[2,3]

Supplementary Reference

(1) Micsonai, A.; Wien, F.; Kernya, L.; Lee, Y. H.; Goto, Y.; Refregiers, M.; Kardos, J. Accurate Secondary Structure Prediction and Fold Recognition for Circular Dichroism Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, E3095-3103.

(2) Lin, Y.; Lee, Y. H.; Yoshimura, Y.; Yagi, H.; Goto, Y. Solubility and Supersaturation-dependent Protein Misfolding Revealed by Ultrasonication. *Langmuir* **2014**, *30*, 1845-1854.

(3) Muta, H.; Lee, Y. H.; Kardos, J.; Lin, Y.; Yagi, H.; Goto, Y. Supersaturation-limited Amyloid Fibrillation of Insulin Revealed by Ultrasonication. *J. Biol. Chem.* **2014**, *289*, 18228-18238.

(4) Tsolis, A. C.; Papandreou, N. C.; Iconomidou, V. A.; Hamodrakas, S. J. A Consensus Method for the Prediction of 'Aggregation-prone' Peptides in Globular Proteins. *PLoS One* **2013**, *8*, e54175.

(5) Conchillo-Sole, O.; de Groot, N. S.; Aviles, F. X.; Vendrell, J.; Daura, X.; Ventura, S. AGGRESCAN: a Server for the Prediction and Evaluation of "Hot Spots" of Aggregation in Polypeptides. *BMC bioinformatics* **2007**, *8*, 65.

(6) Fernandez-Escamilla, A. M.; Rousseau, F.; Schymkowitz, J.; Serrano, L. Prediction of Sequence-dependent and Mutational Effects on the Aggregation of Peptides and Proteins. *Nat. Biotechnol.* **2004**, *22*, 1302-1306.

(7) Trovato, A.; Seno, F.; Tosatto, S. C. The PASTA Server for Protein Aggregation Prediction. *Protein Eng. Des. Sel.* **2007**, *20*, 521-523.

(8) Garbuzynskiy, S. O.; Lobanov, M. Y.; Galzitskaya, O. V. FoldAmyloid: a Method of Prediction of Amyloidogenic Regions from Protein Sequence. *Bioinformatics* **2010**, *26*, 326-332.

(9) Emily, M.; Talvas, A.; Delamarche, C. MetAmyl: a METa-predictor for AMYLoid

Proteins. *PLoS One* **2013**, *8*, e79722.

(10) Tartaglia, G. G.; Pawar, A. P.; Campioni, S.; Dobson, C. M.; Chiti, F.; Vendruscolo, M. Prediction of Aggregation-prone Regions in Structured Proteins. *J. Mol. Biol.* **2008**, *380*, 425-436.

(11) Gasior, P.; Kotulska, M. FISH Amyloid - a New Method for Finding Amyloidogenic Segments in Proteins Based on Site Specific Co-occurrence of Aminoacids. *BMC bioinformatics* **2014**, *15*, 54.

(12) Maurer-Stroh, S.; Debulpaep, M.; Kuemmerer, N.; Lopez de la Paz, M.; Martins, I. C.; Reumers, J.; Morris, K. L.; Copland, A.; Serpell, L.; Serrano, L.; Schymkowitz, J. W.; Rousseau, F. Exploring the Sequence Determinants of Amyloid Structure Using Position-specific Scoring Matrices. *Nat. Methods* **2010**, *7*, 237-242.