Supporting Information for

Protein Vesicles Self-Assembled from Functional Globular Proteins with Different Charge and Size

Dylan R. Dautel and Julie A. Champion

School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, 950 Atlantic Drive NW, Atlanta, Georgia 30332, United States

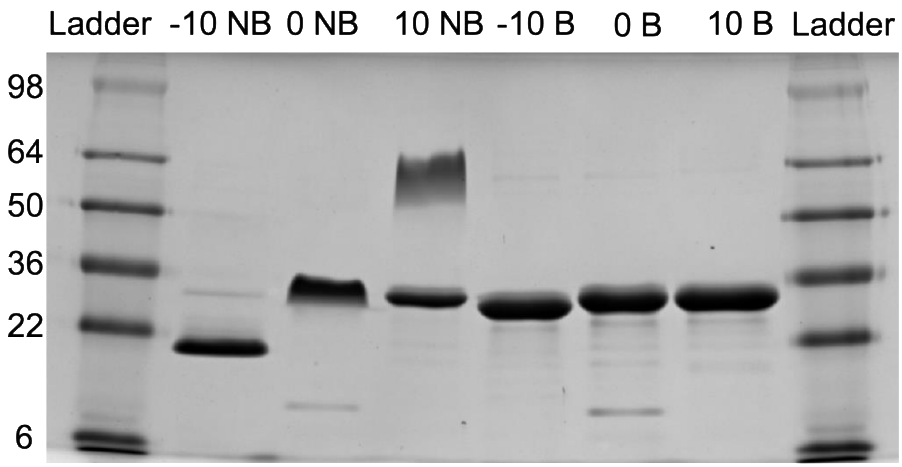

Figure S1. SDS-PAGE gel of sfGFP-Z_E. Each charged variant of sfGFP runs differently when not denatured by boiling. The negative variant runs further because there are more negatively charged amino acids on the surface of the protein. The neutral variant runs nearly the same MW as the denatured proteins and the positive variant has two bands one that arrests at a higher MW and another that gets denatured by SDS and runs at the expected MW. The molecular weights of each band on the ladder is shown in kDa. NB- not boiled, B- boiled.

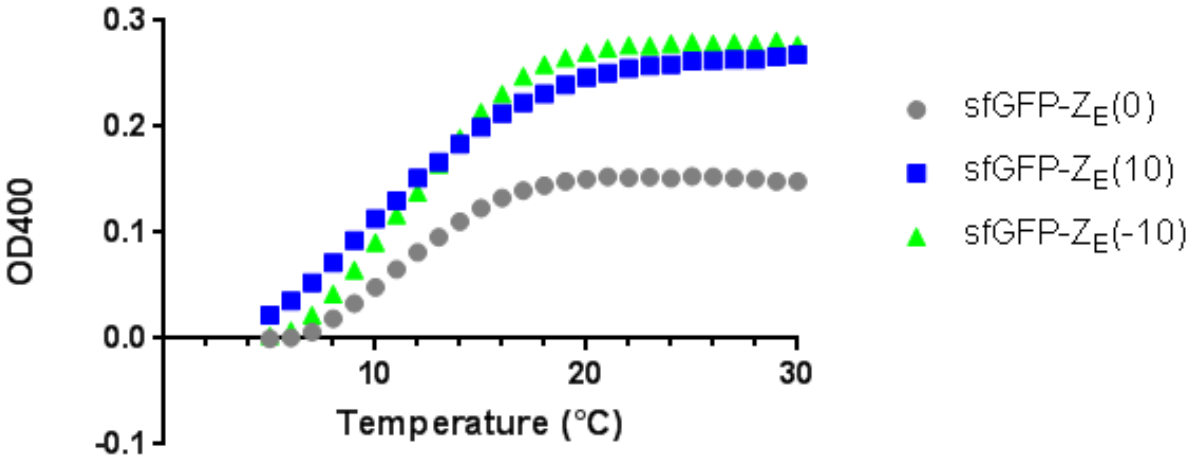

Figure S2. Transition temperature of sfGFP-Z_E charge variants. Solutions contain 30 µM Z_R-ELP, 1.5 µM sfGFP-Z_E, and 1.0 M PBS. The T_t (point of largest slope) of each variant is 12 °C.

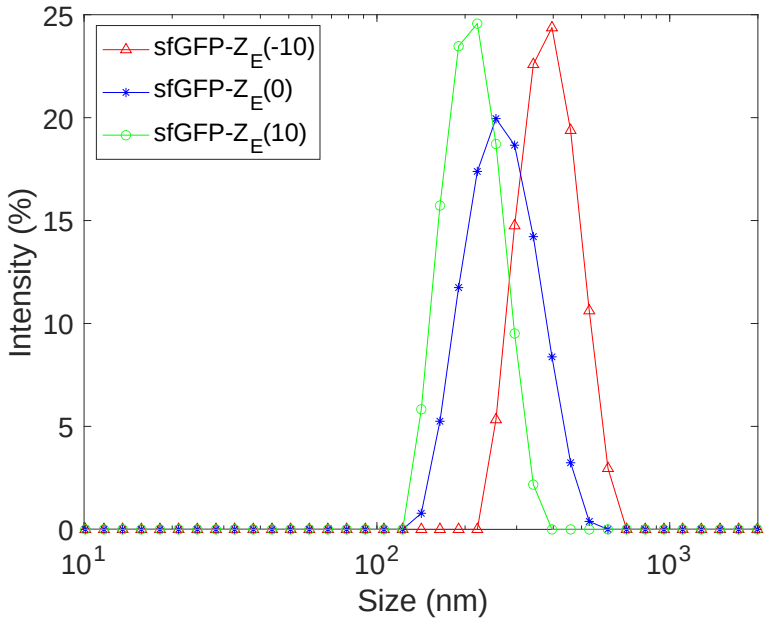

Figure S3. Size intensity plot of sfGFP-Z_E vesicles assembled 1.5 M PBS, 30 µM Z_R-ELP, and Z_E/Z_R of 0.10. Increasing the salt concentration decreases the size of vesicles from greater than 1 µm to ~400 nm.

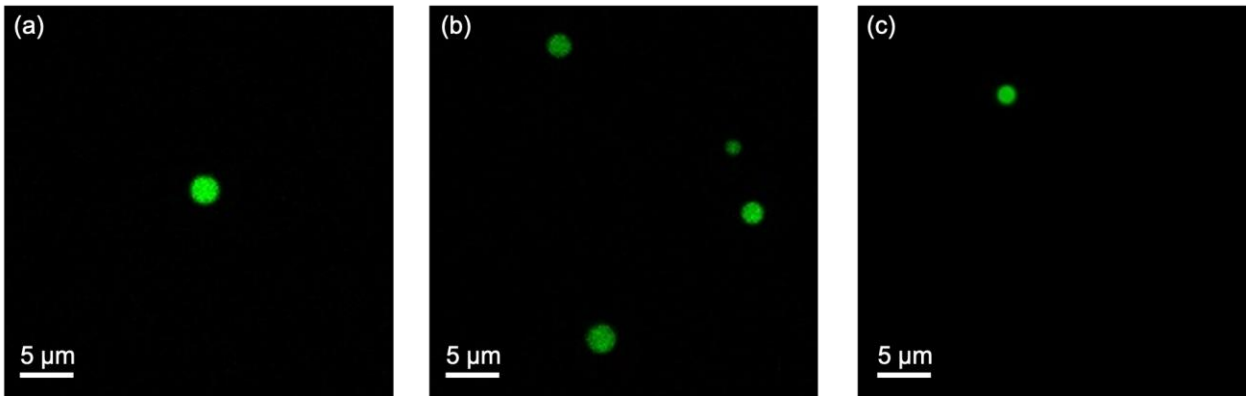

Figure S4. Fluorescent micrographs of HCA-Z_E coacervates. HCA-Z_E coacervates made at 25 °C in solutions containing 0.15 M PBS, 10 µM Z_R-ELP, and (a) 0.01 (b) 0.05 (c) 0.10 Z_E/Z_R molar ratio.

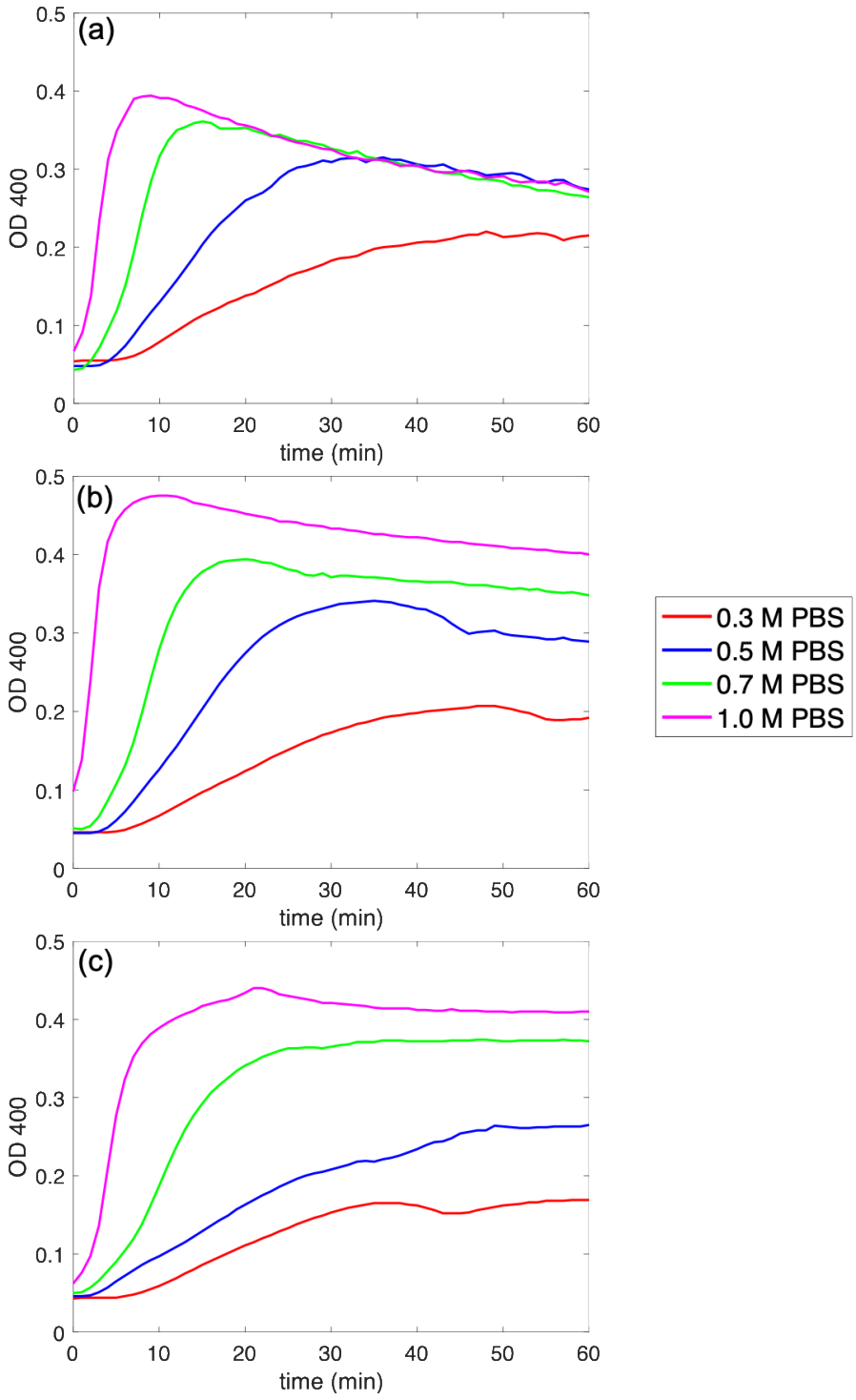

Figure S5. HCA-Z_E turbidity profiles of solutions containing 10 μM Z_R-ELP with varying Z_E/Z_R. (a) The turbidity decreases over time for solutions containing 0.01 Z_E/Z_R, signaling the formation of unstable structures. (b) As salt concentration increases, the structures with 0.05 Z_E/Z_R become more stable. (c) Stable HCA-Z_E vesicles with 0.1 Z_E/Z_R form from solutions containing salt concentrations above 0.3 M PBS.

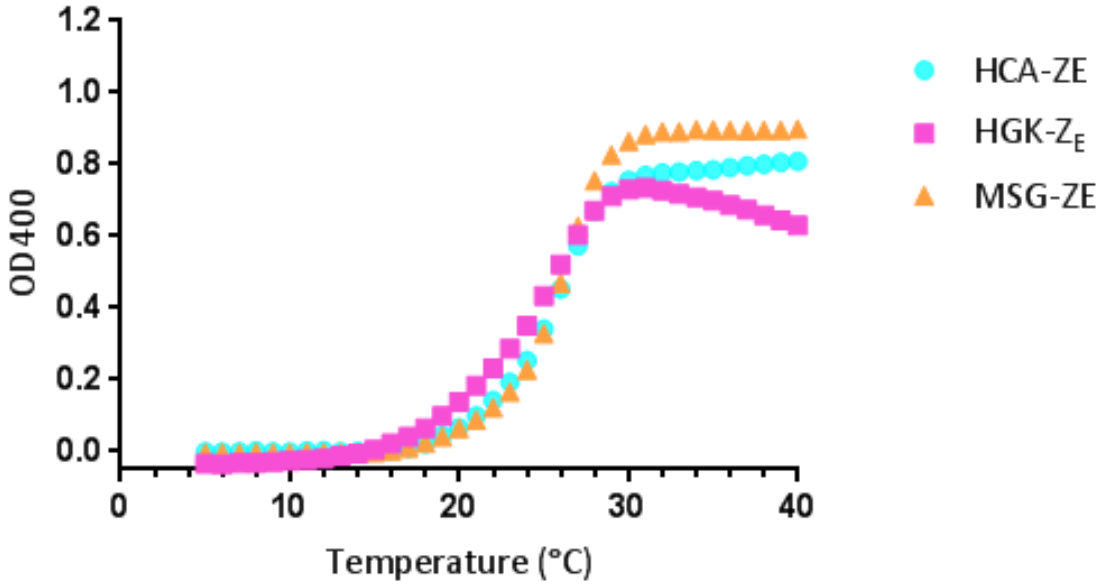

Figure S6. The transition temperature of solutions containing 30 µM Z_R-ELP, 1.5 µM Globule-Z_E, and 0.3 M PBS. The T_t of HCA-Z_E, HGK-Z_E, and MSG-Z_E is 27 °C, 25 °C and 27 °C, respectively.

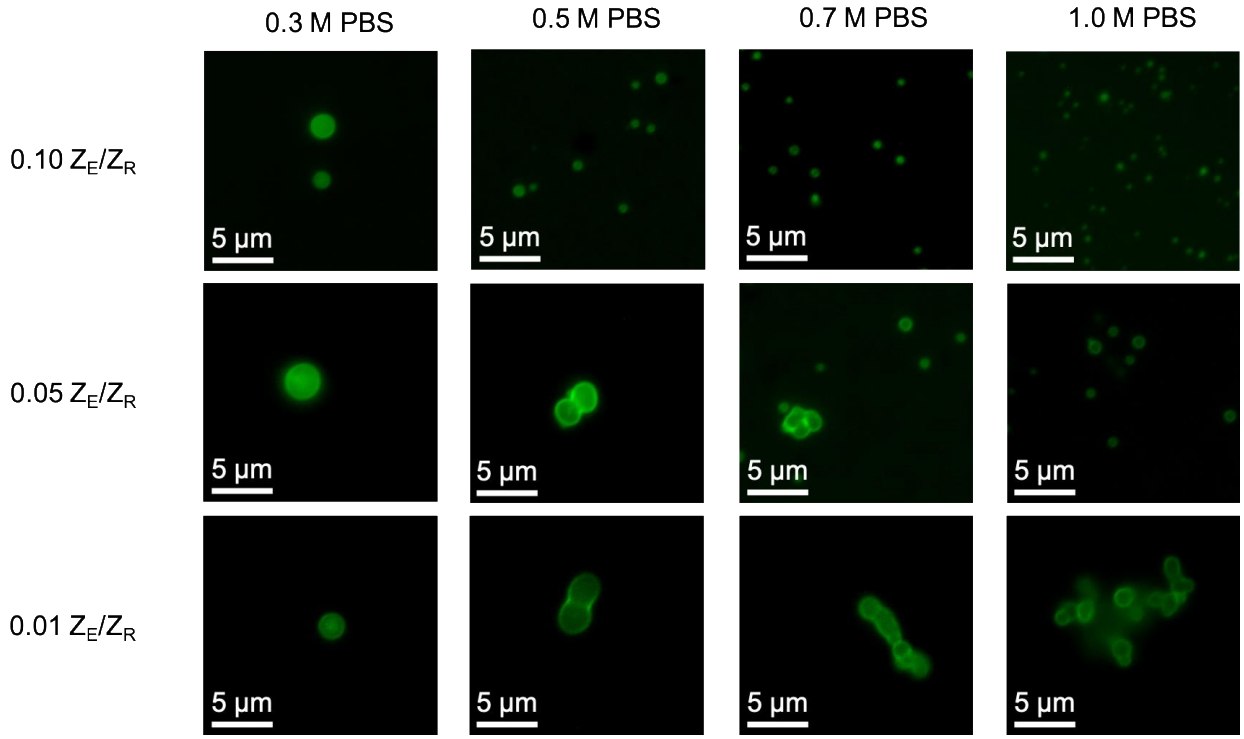

Figure S7. Epifluorescent micrographs of structures in the HCA-Z_E phase diagram containing 10 µM Z_R-ELP.

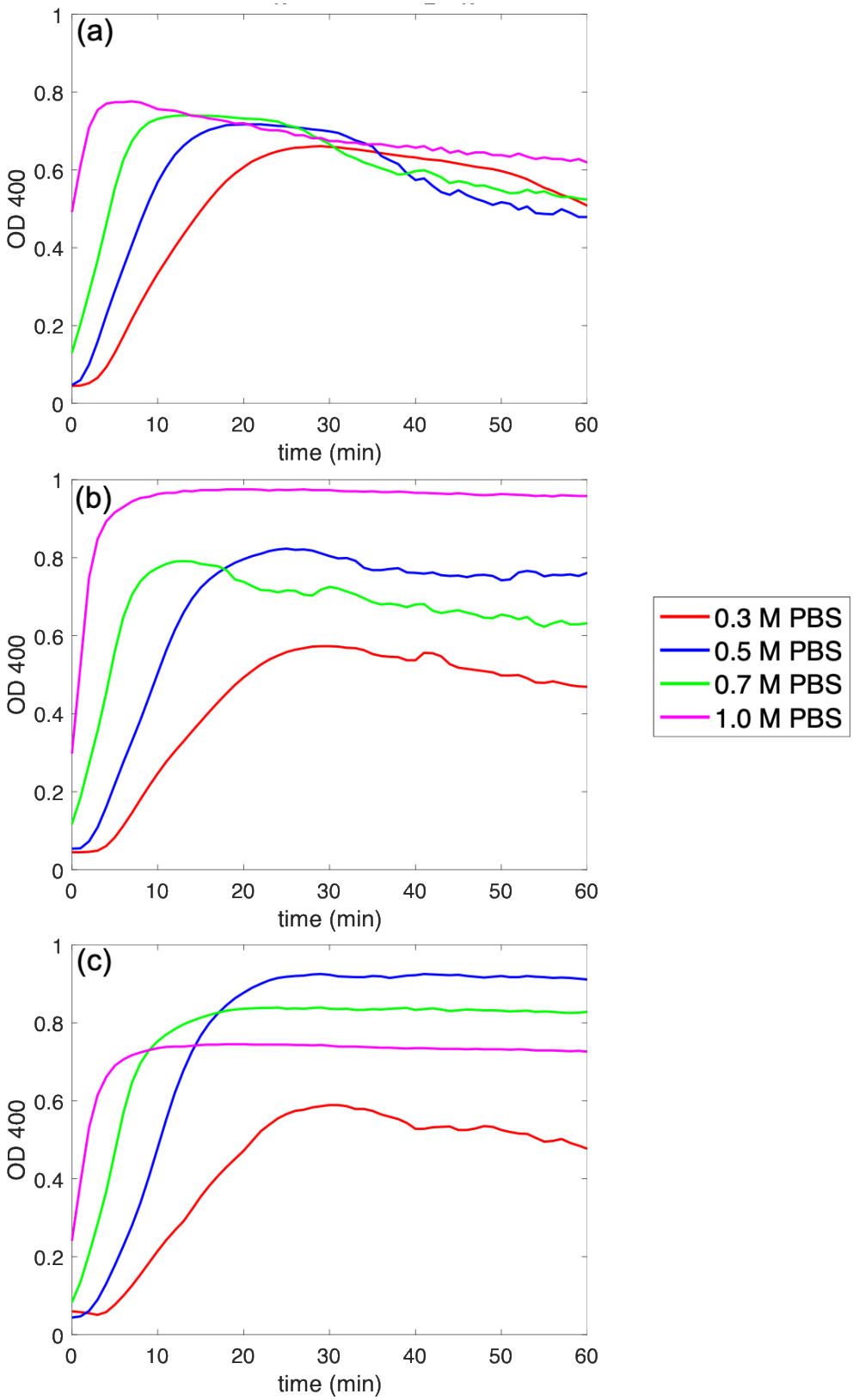

Figure S8. Turbidity profiles of structures assembled at 25 °C from HCA-Z_E with 30 μM Z_R-ELP at (a) 0.01, (b), 0.05, and (c) 0.10 Z_E/Z_R ratio.

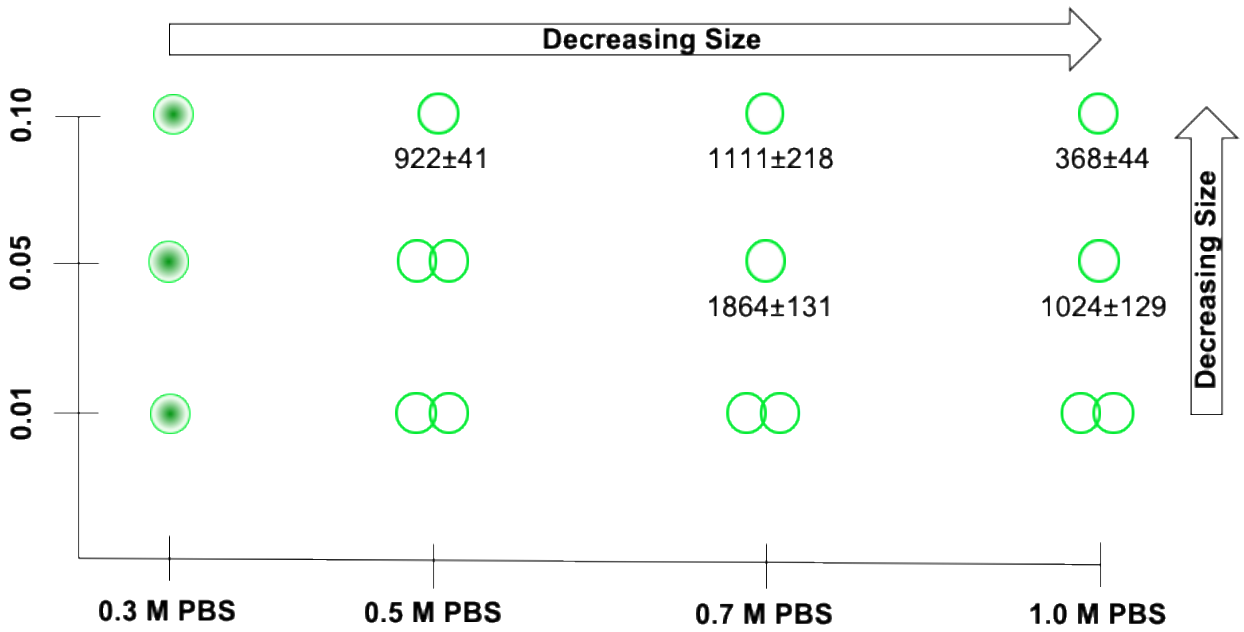

Figure S9. Phase diagram of structures assembled from HCA-Z_E at 25 °C with 30 µM Z_R-ELP. Partially filled circles represent the formation of coacervate-vesicle hybrid structures. Fused rings represent the formation of fused vesicles. Hollow circles represent the formation of vesicles. The average hydrodynamic diameters and standard deviation of stable structures are shown in nm.

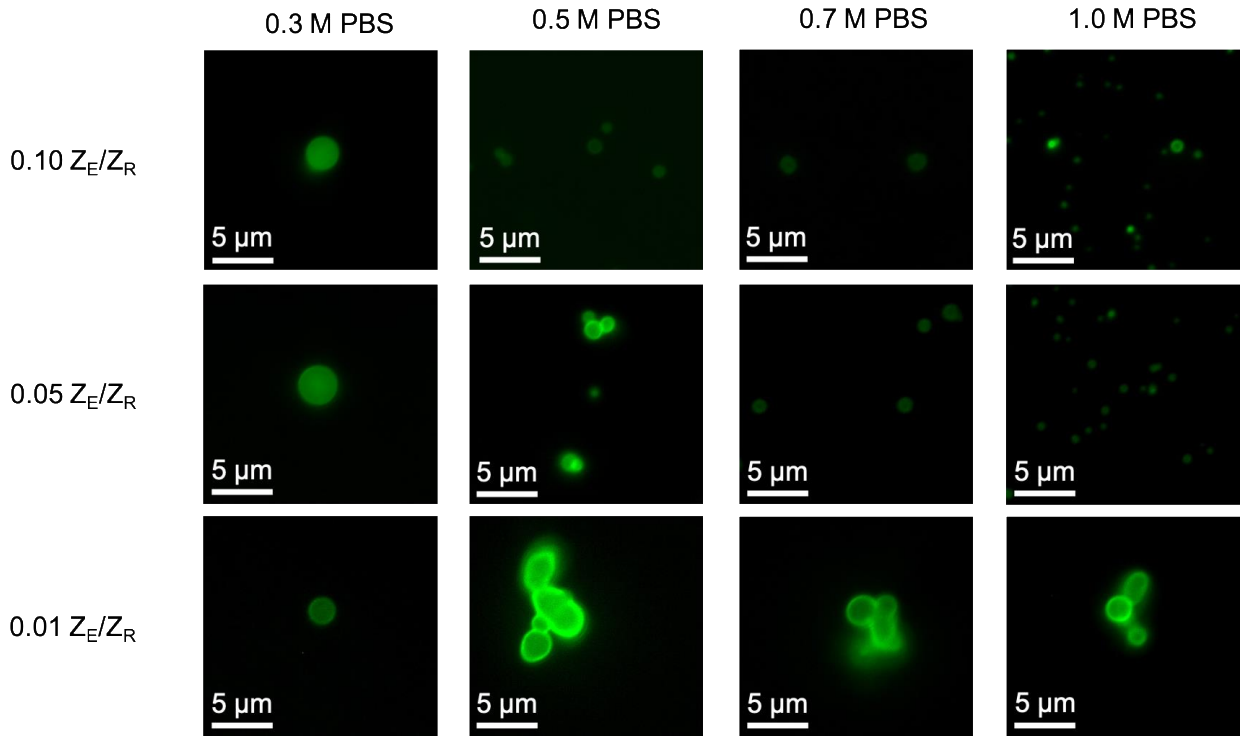

Figure S10. Epifluorescent micrographs of structures in phase diagram assembled from HCA-Z_E at 25 °C with 30 µM Z_R-ELP.

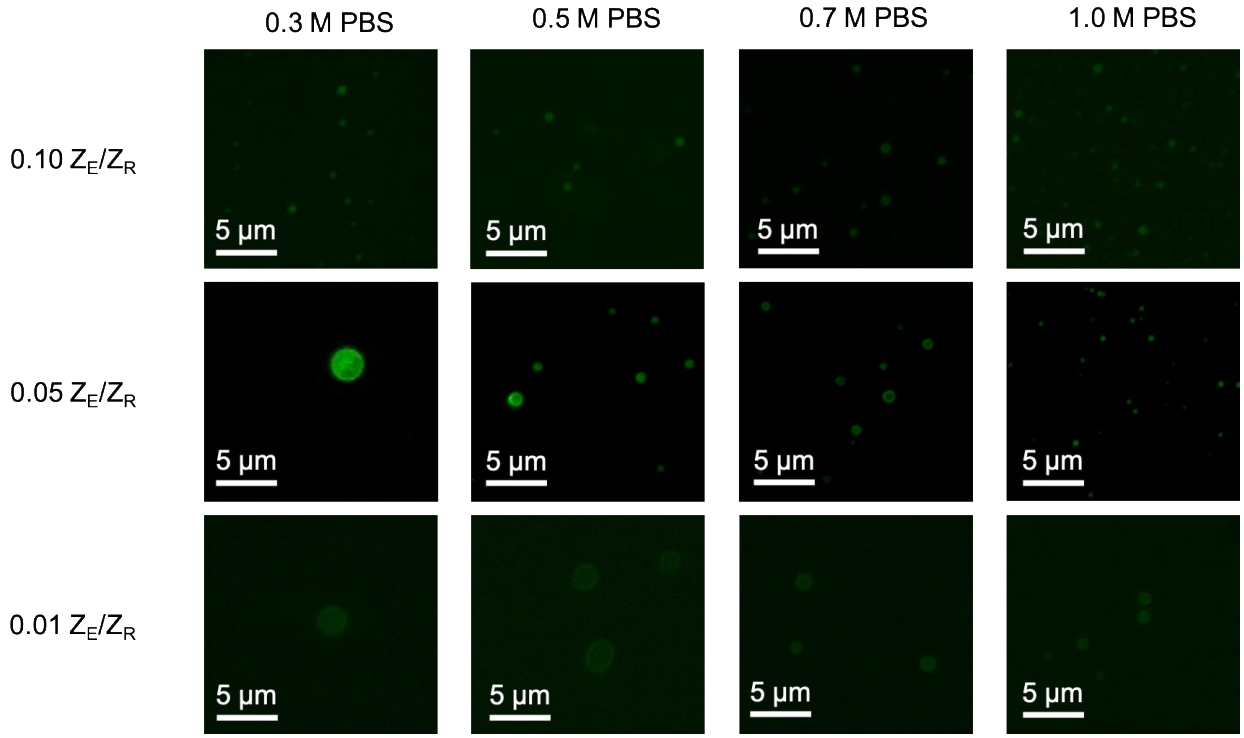

Figure S11. Epifluorescent micrographs of structures in the MSG-Z_E phase diagram at 25 °C with 10 µM Z_R-ELP.

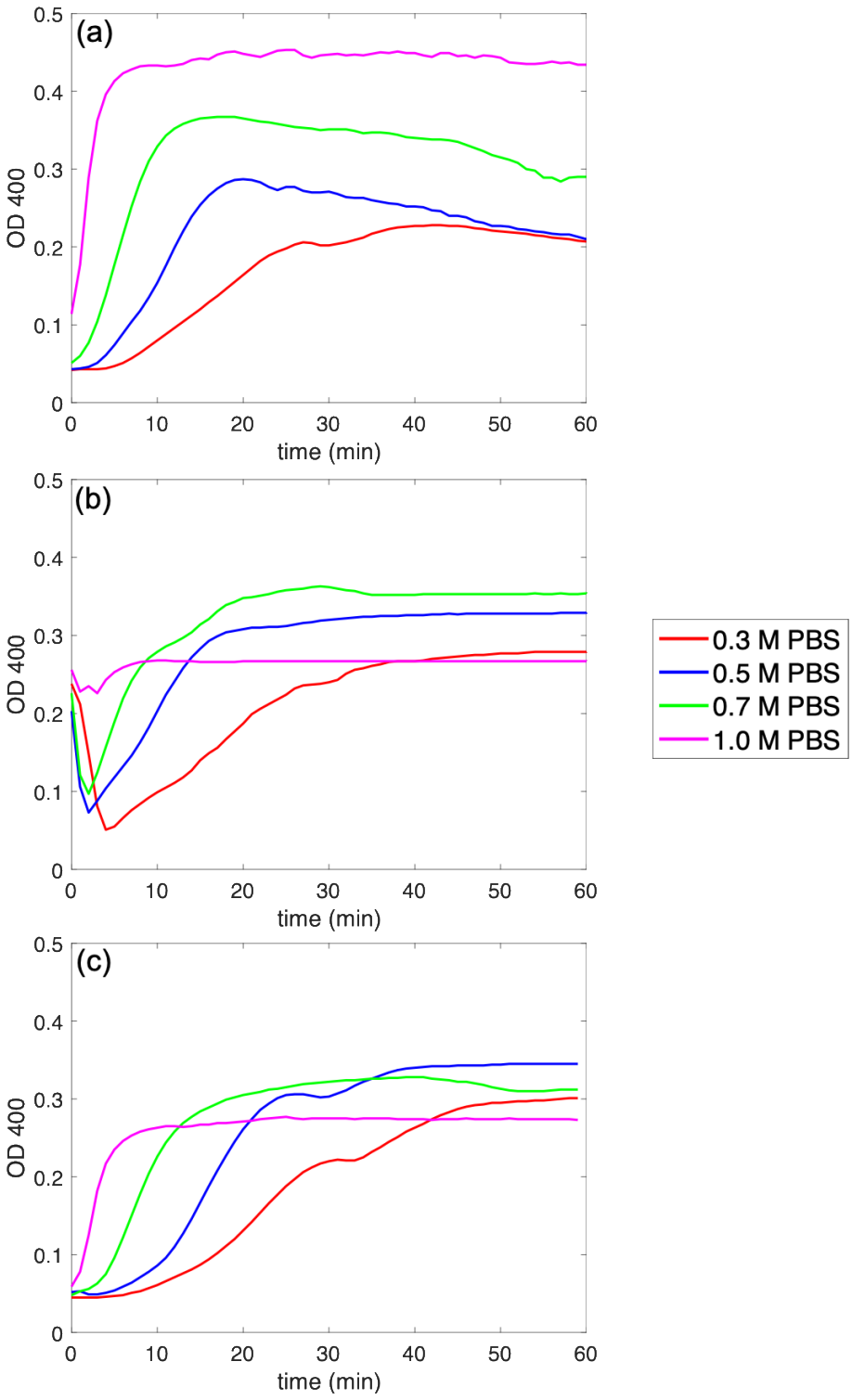

Figure S12. MSG-Z_E turbidity profiles of solutions containing 10 μM Z_R-ELP. (a) The turbidity decreases over time for solutions containing 0.01 Z_E/Z_R, signaling the formation of unstable structures except at 1.0 M PBS. (b) As Z_E/Z_R increases to 0.05 and salt concentration increases, the structures become more stable. (c) Stable HCA-Z_E vesicles form from solutions of all salt concentrations above 0.10 Z_E/Z_R.

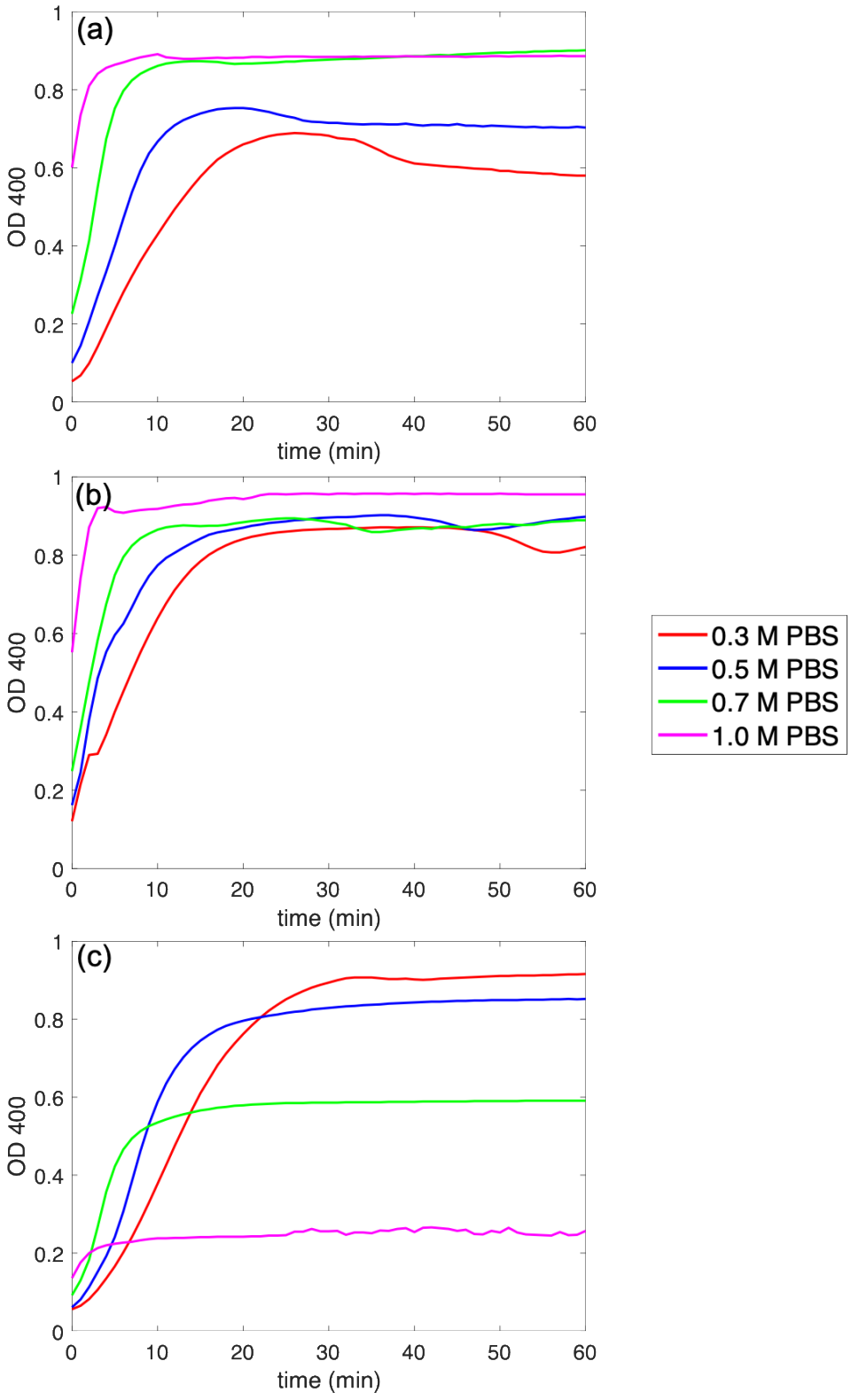

Figure S13. Turbidity profiles of structures assembled from MSG-Z_E at 25 °C with 30 μM Z_R-ELP at (a) 0.01, (b), 0.05, and (c) 0.10 Z_E/Z_R ratio.

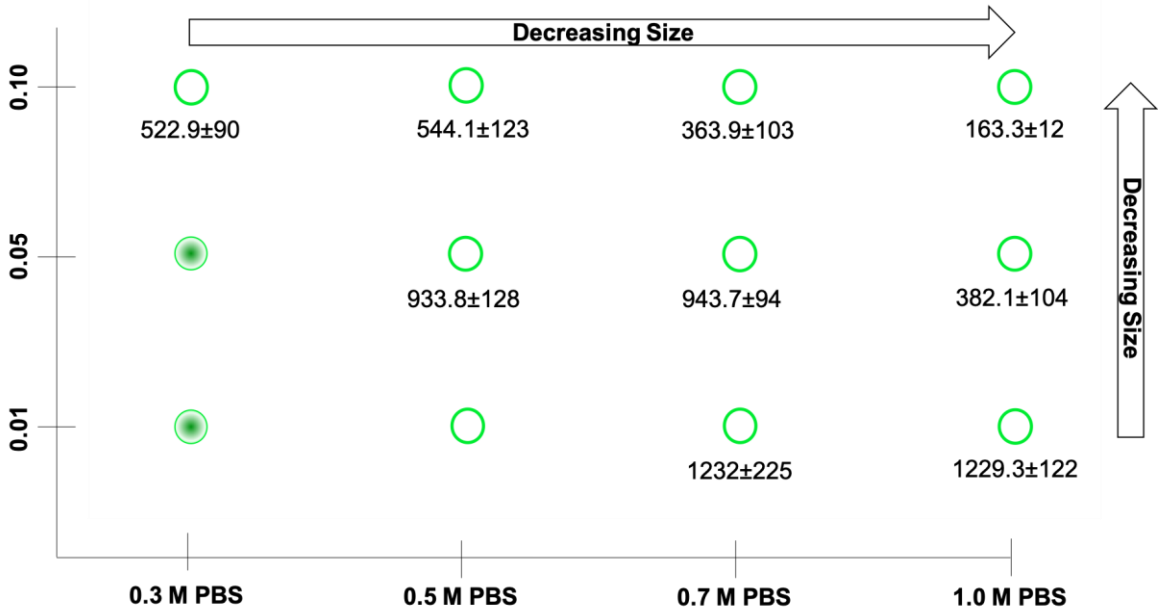

Figure S14. Phase diagram of structures assembled from MSG-Z_E at 25 °C with 30 µM Z_R-ELP. Partially filled circles represent coacervate-vesicle hybrids. Hollow circles represent the formation of vesicles. The average hydrodynamic diameters and standard deviation of stable structures are shown in nm.

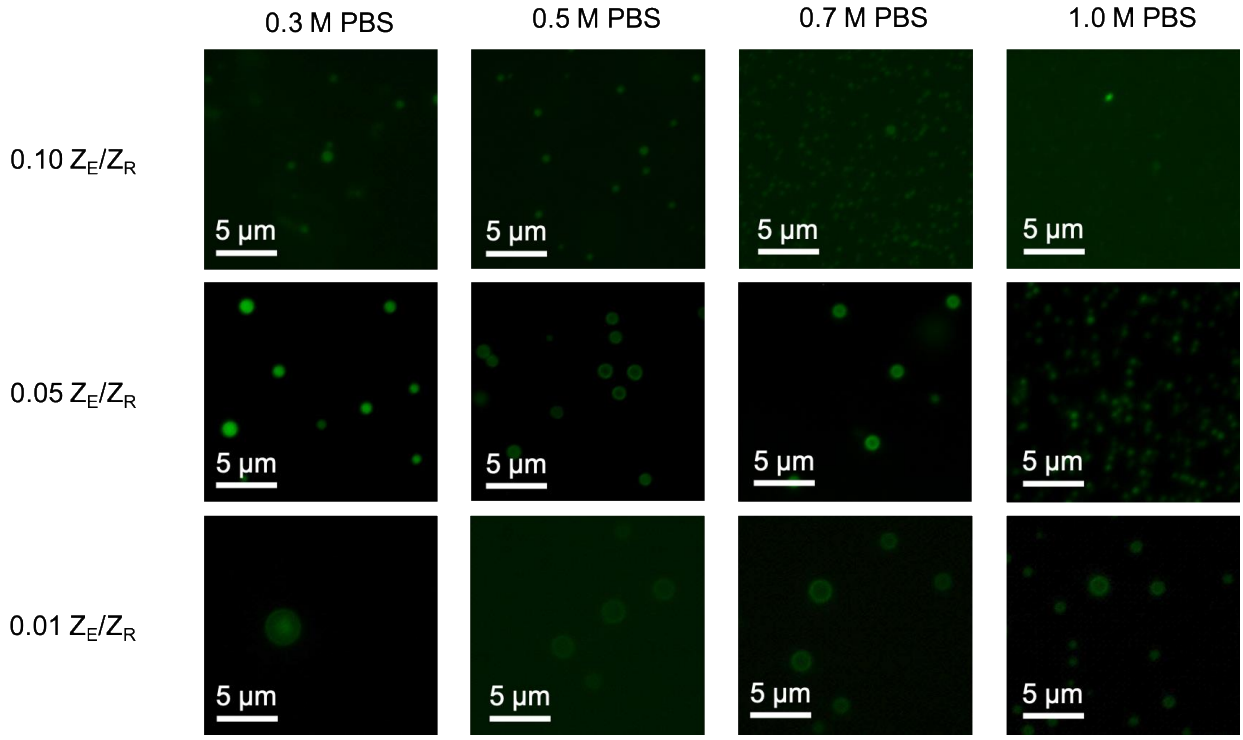

Figure S15. Epifluorescent micrographs of structures in phase diagram assembled from MSG-Z_E at 25 °C with 30 µM Z_R-ELP.

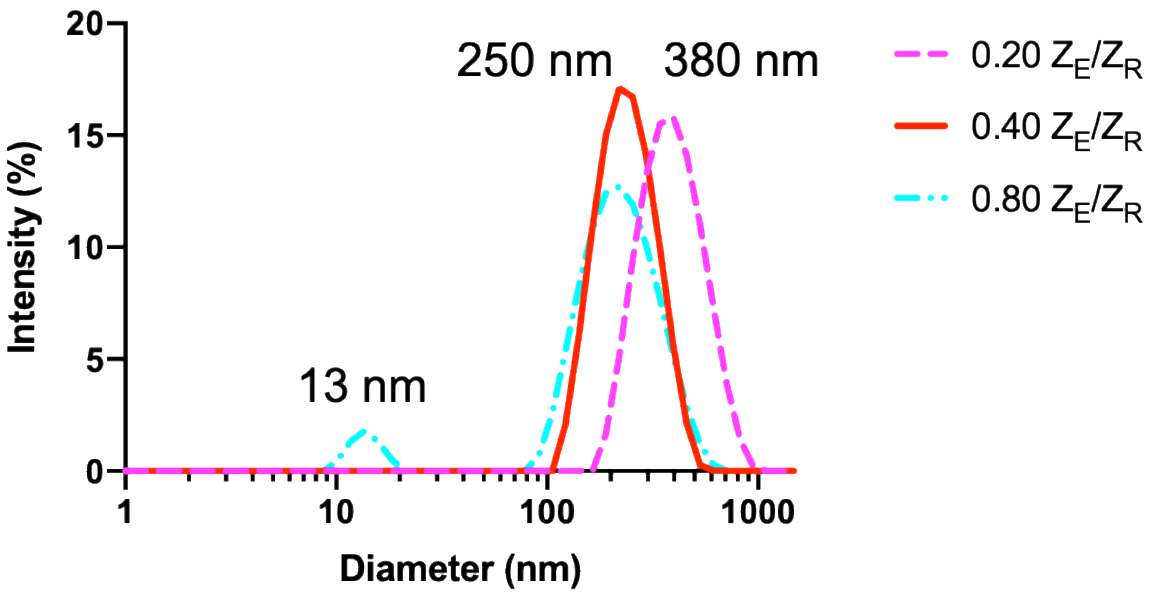

Figure S16. DLS histogram of MSG-Z_E vesicles with high Z_E/Z_R ratio formed with 10 µM Z_R-ELP. Increasing the Z_E/Z_R ratio decreases the size of vesicles until at least 0.40 Z_E/Z_R. Above 0.80 Z_E/Z_R, smaller structures are not formed, instead there is a soluble peak and a vesicle peak.

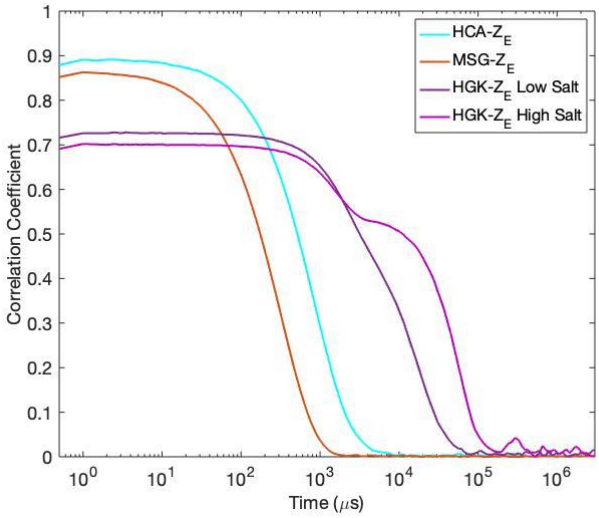

Figure S17- DLS correlograms for solutions containing 10 µM Z_R-ELP and 0.10 Z_E/Z_R incubated at 25 °C for 1 hour. Good quality correlograms are shown for HCA-Z_E and MSG-Z_E in 1.0 M PBS. Poor quality correlograms are shown for HGK-Z_E at 0.5 and 1.0 M PBS. The good quality correlograms have a high correlation coefficient, a steep slope, and a correlation coefficient of zero at high correlation times, which are indicative of low polydispersity and little sedimentation of stable vesicles. The correlograms for poor quality DLS results from HGK-Z_E have a gentler slope that spikes up and down at high correlation times, indicative of high polydispersity and sedimentation.

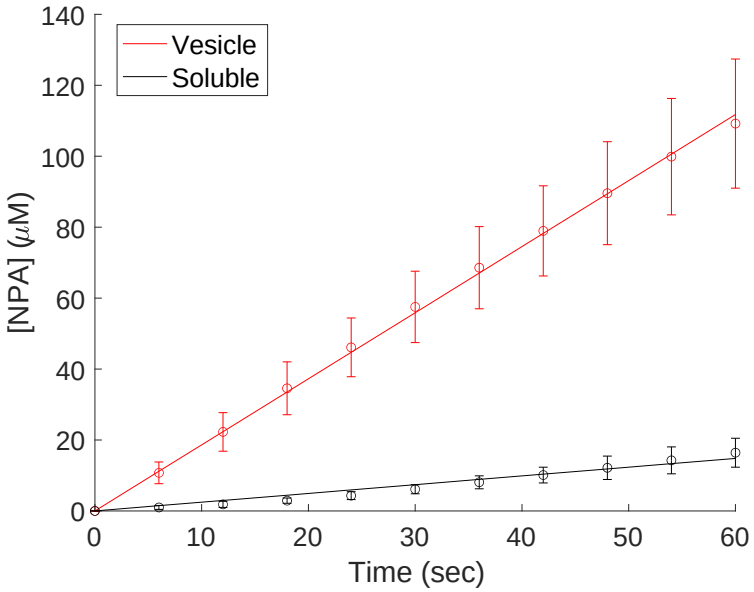

Figure S18. The activity of HCA-Z_E was assayed at 30 °C in the presence of 10 mM $ZnSO_4$, and 1 mM 4-nitrophenyl acetate. The specific activity of HCA-Z_E under these conditions was 0.34 ± 0.01 and 2.23 ± 0.38 µmol product/min/mg enzyme for the soluble enzyme and vesicles, respectively.

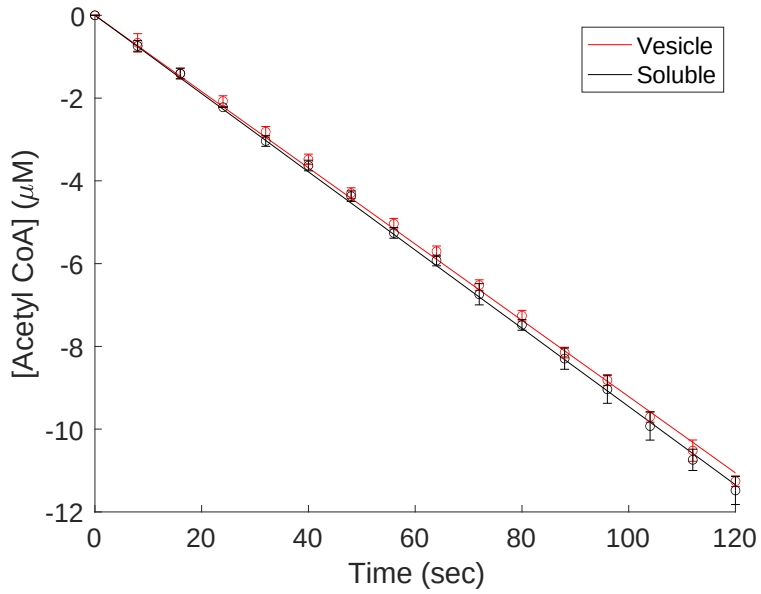

Figure S19. Activity of MSG-Z_E at 30 °C in the presence of 5 mM $MgCl_2$, 0.25 mM acetyl CoA, and 1 mM glyoxylate. The specific activity of MSG-Z_E under these conditions was 4.53 ± 0.13 and 4.42 ± 0.06 µmol product/min/mg enzyme for the soluble enzyme and vesicles, respectively.

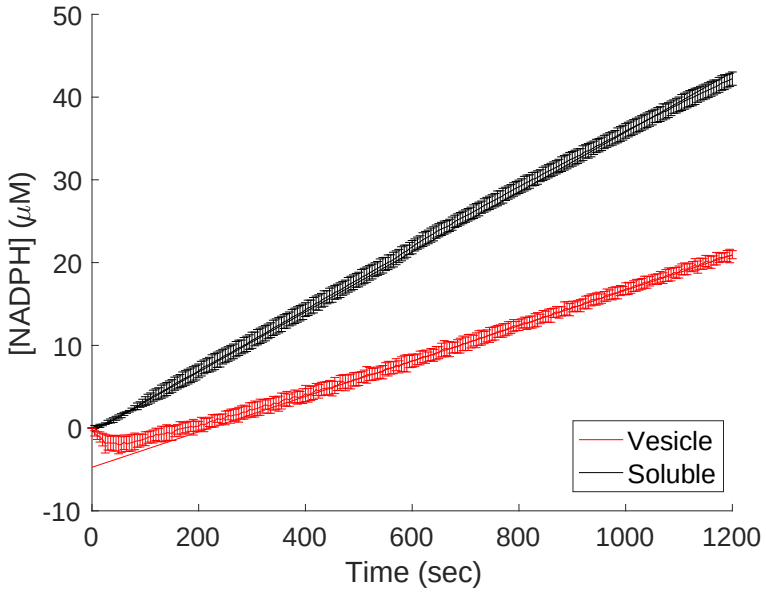

Figure S20. Activity of HGK-Z_E at 30 °C in the presence of 50 mM glucose, 1 mM ATP, and 5 mM $MgCl_2$. To detect HGK-Z_E activity, two units of glucose-6-phosphate dehydrogenase and 0.5 mM NADPH were added to the solution to obtain an absorbance reading for NADPH. The specific activity of HGK-Z_E under these conditions was 0.10 ± 0.002 and 0.06 ± 0.0006 μmol product/min/mg enzyme for the soluble enzyme and vesicles, respectively.

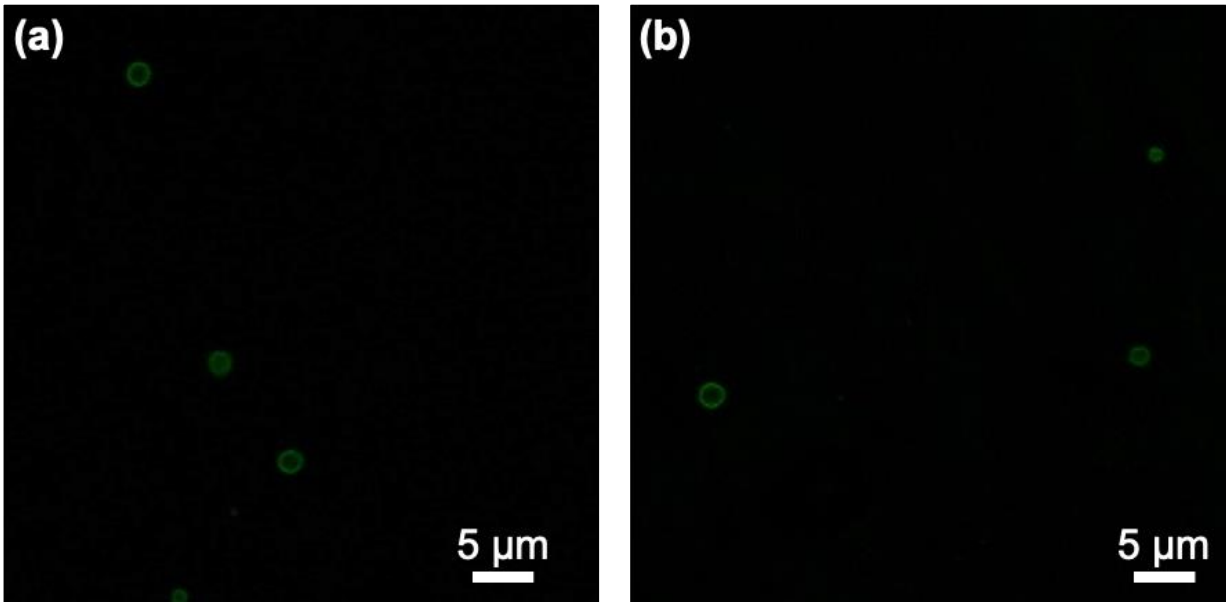

Figure S21. Effect of presence of hydrophobic substrate on HCA-Z_E vesicles. (a) HCA-Z_E vesicles incubated with 1 mM 4-nitrophenyl acetate and 10 mM $ZnSO_4$ after 1 hour of incubation under reaction conditions. (b) HCA-Z_E vesicles incubated under reaction conditions without 4-nitrophenyl acetate substrate after 1 hour.

sfGFP-Z_E Variant	-10	0	+10
Zeta Potential (mV)	-15.7 ± 1.1	-10.1 ± 0.8	9.32 ± 0.1

Table S1. Zeta potential values for soluble sfGFP-Z_E(-10), sfGFP-Z_E(0), and sfGFP-Z_E(10). The values for each variant are lower than the sfGFP alone because they are bound to a negatively charged glutamate rich leucine zipper.

DNA and amino acid sequences for all 7 proteins used in this work. The altered amino acids for sfGFP are highlighted in red for positive changes and blue for negative changes.

Protein	DNA Sequence	Amino Acid Sequence
Z_R-ELP	ATGAAAGGATCCCTGGAAATCCGTGCGGCGGCGCTGCGTC GTCGTAACACCGCGCTGCGTACCCGTGTTGCGGAACTGCGT CAGCGTGTTCAGCGTCTGCGTAACGAAGTTTCTCAGTACGA AACCCGTTACGGTCCGCTGGGTGGTGGCGGTTCTGGTGGC GGTGGATCCGGGGTGCCGGGTGTGGGCGTTCCGGGCGTG GGTGTACCGGGCTTCGGTGTCCCGGGCGTAGGTGTTCCGG GTGTCGGGGTGCCGGGTGTGGGCGTTCCGGGCGTGGGTG TACCGGGCTTCGGTGTCCCGGGCGTAGGTGTTCCGGGTGT CGGGGTGCCGGGTGTGGGCGTTCCGGGCGTGGGTGTACC GGGCTTCGGTGTCCCGGGCGTAGGTGTTCCGGGTGTCGGG GTGCCGGGTGTGGGCGTTCCGGGCGTGGGTGTACCGGGCT TCGGTGTCCCGGGCGTAGGTGTTCCGGGTGTCGGGGTGCC GGGTGTGGGCGTTCCGGGCGTGGGTGTACCGGGCTTCGGT GTCCCGGGCGTAGGTGTTCCGGGTGTCGGGGTGCCCGGGT GCTAAA	MKGSLEIRAAALRR RNTALRTRVAELRQ RVQRLRNEVSQYET RYGPLGGGGSGGG GSGVPGVGVPGVG VPGFGVPGVGVPG VGVPGVGVPGVGV PGFGVPGVGVPGV GVPGVGVPGVGVP GFGVPGVGVPGVG VPGVGVPGVGVPG FGVPGVGVPGVGV PGVGVPGVGVPGF GVPGVGVPGVGVP GC
sfGFP-Z_E(-10)	ATGGGCCACCACCACCACCACCACCGTAGCAGCAAAGGCG AGGAACTGTTCACCGGTGTGGTTCCGATCCTGGTGGAACTG GACGGCGATGTTAACGGTCACAAGTTTAGCGTTCGTGGTG AGGGCGAAGGTGACGCGACCAACGGCAAGCTGACCCTGA AATTCATTTGCACCACCGGTAAACTGCCGGTGCCGTGGCCG ACCCTGGTTACCACCCTGACCTACGGCGTTCAGTGCTTTAGC CGTTATCCGGACCACATGAAGCGTCACGATTTCTTTAAAAG CGCGATGCCGGAAGGCTACGTGCAAGAGCGTACCATCAGC TTCAAGGACGATGGTACCTATAAAACCCGTGCGGAAGTTAA ATTTGAGGGCGACACCCTGGTGAACCGTATCGAACTGAAG GGTATTGACTTCAAAGAGGATGGCAACATCCTGGGTCACA AGCTGGAATACAACTTTAACAGCCACAACGTGTATATTACC GCGGATGAACAGAAAAACGGCATCAAGGCGAACTTTAAAA TTCGTCACAACGTGGAGGACGGTAGCGTTCAACTGGCGGA TCACTACCAGCAAAACACCCCGATTGGTGATGGTCCGGTTC TGCTGCCGGATAACCACTATCTGAGCACCCAGAGCGTGCTG	MGHHHHHHHRSSK GEELFTGVVPILVEL DGDVNGHKFSVRG EGEGDATNGKLTLK FICTTGKLPVPWPT LVTTLTYGVQCFSR YPDHMKRHDFFKS AMPEGYVQERTISF KDDGTYKTRAEVKF EGDTLVNRIELKGID FKEDGNILGHKLEY NFNSHNVYITADEQ KNGIKANFKIRHNV EDGSVQLADHYQQ NTPIGDGPVLLPDN HYLSTQSVLSKDPD

	AGCAAGGACCCGGATGAAGAACGTGATCACATGGTGCTGC TGGAGTTCGTTACCGCGGCGGGCATCACCCACGGTATGGA TGAACTGTACAAACTGCGTGGCAGCGGTAGCCTGGAGATT GAAGCGGCGGCGCTGGAGCAGGAAAACACCGCGCTGGAG ACCGAAGTGGCGGAGCTGGAACAGGAAGTGCAACGTCTG GAGAACATTGTGAGCCAATACCGTACCCGTTATGGTCCGCT G	EERDHMVLLEFVTA AGITHGMDELYKLR GSGSLEIEAAALEQE NTALETEVAELEQE VQRLENIVSQYRTR YGPL
sfGFP-Z_E(0)	ATGGGCGGCCACCACCACCACCACCACCGTAGCAGCAAAG GCGAGGAACTGTTCACCGGTGTGGTTCCGATCCTGGTGGA ACTGGACGGCGATGTTAACGGTCACAAGTTTAGCGTTCGTG GTGAGGGCGAAGGTGACGCGACCAACGGCAAGCTGACCCT GAAATTCATTTGCACCACCGGTAAACTGCCGGTGCCGTGGC CGACCCTGGTTACCACCCTGACCTACGGCGTTCAGTGCTTTA GCCGTTATCCGGACCACATGAAGCGTCACGATTTCTTTAAA AGCGCGATGCCGGAAGGCTACGTGCAAGAGCGTACCATCA GCTTCAAGGACGATGGTACCTATAAAACCCGTGCGGAAGTT AAATTTGAGGGCGACACCCTGGTGAACCGTATCGAACTGA AGGGTATTGACTTCAAAAAAAAAGGCAACATCCTGGGTCAC AAGCTGGAATACAACTTTAACAGCCACAACGTGTATATTAC CGCGGATAAGAAAAAAAACGGCATCAAGGCGAACTTTAAA ATTCGTCACAACGTGGAGGACGGTAGCGTTCAACTGGCGG ATCACTACCAGCAAAACACCCCGATTGGTGATGGTCCGGTT CTGCTGCCGGATAACCACTATCTGAGCACCCAGAGCGTGCT GAGCAAGGACCCGAAAGAAAAACGTGATCACATGGTGCTG CTGGAGTTCGTTACCGCGGCGGGCATCACCCACGGTATGG ATGAACTGTACAAACTGCGTGGCAGCGGTAGCCTGGAGAT TGAAGCGGCGGCGCTGGAGCAGGAAAACACCGCGCTGGA GACCGAAGTGGCGGAGCTGGAACAGGAAGTGCAACGTCT GGAGAACATTGTGAGCCAATACCGTACCCGTTATGGTCCGC TGTAAAAGCTT	MGHHHHHHRSSK GEELFTGVVPILVEL DGDVNGHKFSVRG EGEGDATNGKLTLK FICTTGKLPVPWPT LVTTLTYGVQCFSR YPDHMKRHDFFKS AMPEGYVQERTISF KDDGTYKTRAEVKF EGDTLVNRIELKGID FKKKGNILGHKLEY NFNSHNVYITADKK KNGIKANFKIRHNV EDGSVQLADHYQQ NTPIGDGPVLLPDN HYLSTQSVLSKDPK EKRDHMVLLEFVTA AGITHGMDELYKLR GSGSLEIEAAALEQE NTALETEVAELEQE VQRLENIVSQYRTR YGPL
sfGFP-Z_E(10)	ATGGGCCACCACCACCACCACCACCGTAGCAGCAAGGGCG AAAAACTGTTCACCGGTGTGGTTCCGATCCTGGTGGAGCTG GACGGCGATGTTAACGGTCACAAGTTTAGCGTTCGTGGTG AGGGCGAAGGTGACGCGACCAACGGCAAGCTGACCCTGA AATTCATTTGCACCACCGGTAAACTGCCGGTGCCGTGGCCG ACCCTGGTTACCACCCTGACCTACGGCGTTCAGTGCTTTAGC CGTTATCCGAAGCACATGAAACGTCACGACTTCTTTAAGAG CGCGATGCCGGAAGGCTACGTGCAAGAGCGTACCATCAGC TTCAAGAAAGATGGTACCTATAAAACCCGTGCGGAAGTTAA ATTTGAGGGCAAAACCCTGGTGAACCGTATCGAACTGAAG GGTATTGACTTCAAGAAAAAGGGCAACATCCTGGGTCACA AACTGGAGTACAACTTTAACAGCCACAACGTGTATATTACC GCGGATAAAAAGAAAAACGGCATCAAGGCGAACTTTAAAA TTCGTCACAACGTGGAAGACGGTAGCGTTCAGCTGGCGGA	MGHHHHHHRSSK GEKLFTGVVPILVEL DGDVNGHKFSVRG EGEGDATNGKLTLK FICTTGKLPVPWPT LVTTLTYGVQCFSR YPKHMKRHDFFKS AMPEGYVQERTISF KKDGTYKTRAEVKF EGKTLVNRIELKGID FKKKGNILGHKLEY NFNSHNVYITADKK KNGIKANFKIRHNV EDGSVQLADHYQQ

	TCACTACCAGCAAAACACCCCGATCGGCAAGGGTCCGGTTC TGCTGCCGGACAACCACTATCTGAGCACCCAAAGCGTGCTG AGCAAAGACCCGAAGGAAAAACGTGATCACATGGTGCTGC TGGAGTTCGTTACCGCGGCGGGCATCACCCACGGTATGGA TGAACTGTACAAGCTGCGTGGCAGCGGTAGCCTGGAGATT GAAGCGGCGGCGCTGGAGCAGGAAAACACCGCGCTGGAG ACCGAAGTGGCGGAGCTGGAACAGGAAGTGCAACGTCTG GAGAACATTGTGAGCCAATACCGTACCCGTTATGGTCCGCT G	NTPIGKGPVLLPDN HYLSTQSVLSKDPK EKRDHMVLLEFVTA AGITHGMDELYKLR GSGSLEIEAAALEQE NTALETEVAELEQE VQRLENIVSQYRTR YGPL
HCA-Z_E	ATGAGCCACCACTGGGGTTACGGCAAACACAACGGTCCGG AACACTGGCACAAAGACTTTCCGATTGCGAAGGGTGAACG TCAGAGCCCGGTTGACATTGATACCCACACCGCGAAATACG ACCCGAGCCTGAAGCCGCTGAGCGTGAGCTATGATCAGGC GACCAGCCTGCGTATCCTGAACAACGGTCACGCGTTCAACG TTGAATTTGACGATAGCCAAGACAAGGCGGTGCTGAAAGG TGGCCCGCTGGATGGTACCTATCGTCTGATTCAGTTCCACTT TCACTGGGGCAGCCTGGACGGTCAAGGCAGCGAACACACC GTTGATAAGAAAAAGTACGCGGCGGAGCTGCACCTGGTGC ACTGGAACACCAAATATGGTGACTTCGGCAAGGCGGTTCA GCAACCGGATGGTCTGGCGGTGCTGGGCATCTTTCTGAAG GTTGGTAGCGCGAAACCGGGCCTGCAGAAGGTGGTTGACG TGCTGGATAGCATTAAAACCAAGGGTAAAAGCGCGGACTT CACCAACTTTGATCCGCGTGGCCTGCTGCCGGAAAGCCTGG ACTACTGGACCTACCCGGGTAGCCTGACCACCCCGCCGCTG CTGGAGTGCGTGACCTGGATCGTTCTGAAAGAACCGATTA GCGTGAGCAGCGAGCAAGTTCTGAAGTTCCGTAAACTGAA CTTTAACGGCGAGGGCGAACCGGAGGAACTGATGGTTGAT AACTGGCGTCCGGCGCAGCCGCTGAAGAACCGTCAAATCA AGGCGAGCTTTAAACATATGGGTAGCGGCAGCCTGGAGAT TGAAGCGGCGGCGCTGGAGCAGGAAAACACCGCGCTGGA AACCGAGGTGGCGGAGCTGGAACAGGAAGTGCAACGTCT GGAGAACATCGTGAGCCAATACCGTACCCGTTATGGCCCGC TGCATCACCATCACCATCACTAA	MSHHWGYGKHNG PEHWHKDFPIAKG ERQSPVDIDTHTAK YDPSLKPLSVSYDQ ATSLRILNNGHAFN VEFDDSQDKAVLK GGPLDGTYRLIQFH FHWGSLDGQGSEH TVDKKKYAAELHLV HWNTKYGDFGKAV QQPDGLAVLGIFLK VGSAKPGLQKVVD VLDSIKTKGKSADFT NFDPRGLLPESLDY WTYPGSLTTPPLLE CVTWIVLKEPISVSS EQVLKFRKLNFNGE GEPEELMVDNWRP AQPLKNRQIKASFK HMGSGSLEIEAAAL EQENTALETEVAEL EQEVQRLENIVSQY RTRYGPLHHHHHH
HGK-Z_E	ATGCTGGACGATCGTGCGCGTATGGAAGCGGCGAAGAAA GAAAAGGTGGAGCAGATCCTGGCGGAGTTTCAGCTGCAAG AGGAAGACCTGAAGAAAGTTATGCGTCGTATGCAAAAGGA GATGGATCGTGGTCTGCGTCTGGAAACCCACGAGGAAGCG AGCGTGAAAATGCTGCCGACCTACGTTCGTAGCACCCCGGA AGGTAGCGAGGTTGGCGACTTTCTGAGCCTGGATCTGGGT GGCACCAACTTCCGTGTGATGCTGGTGAAAGTTGGTGAAG GCGAGGAAGGTCAATGGAGCGTTAAGACCAAACACCAAAT GTACAGCATTCCGGAGGATGCGATGACCGGTACCGCGGAA ATGCTGTTTGATTATATCAGCGAGTGCATTAGCGACTTCCTG GATAAGCACCAGATGAAACACAAGAAACTGCCGCTGGGTT	MLDDRARMEAAK KEKVEQILAEFQLQ EEDLKKVMRRMQK EMDRGLRLETHEEA SVKMLPTYVRSTPE GSEVGDFLSLDLGG TNFRVMLVKVGEG EEGQWSVKTKHQ MYSIPEDAMTGTA EMLFDYISECISDFL DKHQMKHKKLPLG

	TCACCTTTAGCTTCCCGGTGCGTCACGAGGACATCGATAAG GGTATTCTGCTGAACTGGACCAAGGGCTTTAAAGCGAGCG GCGCGGAGGGTAACAACGTGGTTGGTCTGCTGCGTGACGC GATTAAACGTCGTGGCGACTTCGAGATGGATGTGGTTGCG ATGGTGAACGACACCGTTGCGACCATGATCAGCTGCTACTA TGAGGATCACCAGTGCGAAGTGGGTATGATTGTTGGTACC GGCTGCAACGCGTGCTACATGGAGGAAATGCAAAACGTGG AACTGGTTGAGGGCGATGAAGGTCGTATGTGCGTGAACAC CGAATGGGGCGCGTTTGGTGACAGCGGTGAACTGGATGAA TTCCTGCTGGAGTATGACCGTCTGGTTGATGAAAGCAGCGC GAACCCGGGTCAGCAACTGTACGAAAAGCTGATCGGTGGC AAATATATGGGCGAGCTGGTGCGTCTGGTTCTGCTGCGTCT GGTTGACGAGAACCTGCTGTTTCACGGTGAAGCGAGCGAA CAGCTGCGTACCCGTGGCGCGTTTGAAACCCGTTTCGTGAG CCAGGTTGAGAGCGACACCGGTGATCGTAAGCAAATCTAC AACATTCTGAGCACCCTGGGCCTGCGTCCGAGCACCACCGA CTGCGATATTGTGCGTCGTGCGTGCGAAAGCGTTAGCACCC GTGCGGCGCACATGTGCAGCGCGGGTCTGGCGGGCGTGAT CAACCGTATGCGTGAGAGCCGTAGCGAAGACGTTATGCGT ATTACCGTGGGTGTTGATGGCAGCGTGTATAAGCTGCACCC GAGCTTTAAAGAGCGTTTCCACGCGAGCGTTCGTCGTCTGA CCCCGAGCTGCGAGATCACCTTCATTGAAAGCGAGGAAGG TAGCGGTCGTGGTGCGGCGCTGGTGAGCGCGGTTGCGTGC AAGAAAGCGTGCATGCTGGGTCAGCATATGGGCAGCGGTA GCCTGGAGATTGAAGCGGCGGCGCTGGAGCAAGAAAACA CCGCGCTGGAAACCGAGGTGGCGGAGCTGGAACAGGAAG TGCAACGTCTGGAGAACATTGTGAGCCAATACCGTACCCGT TATGGCCCGCTGCATCACCATCACCATCACTAA	FTFSFPVRHEDIDKG ILLNWTKGFKASGA EGNNVVGLLRDAIK RRGDFEMDVVAM VNDTVATMISCYYE DHQCEVGMIVGTG CNACYMEEMQNV ELVEGDEGRMCVN TEWGAFGDSGELD EFLLEYDRLVDESSA NPGQQLYEKLIGGK YMGELVRLVLLRLV DENLLFHGEASEQL RTRGAFETRFVSQV ESDTGDRKQIYNILS TLGLRPSTTDCDIVR RACESVSTRAAHM CSAGLAGVINRMR ESRSEDVMRITVGV DGSVYKLHPSFKER FHASVRRLTPSCEIT FIESEEGSGRGAALV SAVACKKACMLGQ HMGSGSLEIEAAAL EQENTALETEVAEL EQEVQRLENIVSQY RTRYGPLHHHHHH
MSG-Z_E	ATGAGCCAGACCATCACCCAAAGCCGTCTGCGTATTGACGC GAACTTCAAGCGTTTTGTGGATGAGGAAGTTCTGCCGGGTA CCGGCCTGGATGCGGCGGCGTTCTGGCGTAACTTTGACGA AATCGTTCACGATCTGGCGCCGGAGAACCGTCAGCTGCTG GCGGAACGTGACCGTATTCAAGCGGCGCTGGATGAGTGGC ACCGTAGCAACCCGGGCCCGGTTAAGGACAAAGCGGCGTA CAAAAGCTTCCTGCGTGAACTGGGTTATCTGGTGCCGCAGC CGGAGCGTGTGACCGTTGAAACCACCGGCATCGATAGCGA GATTACCAGCCAAGCGGGTCCGCAACTGGTGGTTCCGGCG ATGAACGCGCGTTATGCGCTGAACGCGGCGAACGCGCGTT GGGGCAGCCTGTATGATGCGCTGTATGGTAGCGATATCATT CCGCAGGAAGGTGCGATGGTTAGCGGTTACGACCCGCAGC GTGGCGAACAAGTGATCGCGTGGGTTCGTCGTTTCCTGGAT GAGAGCCTGCCGCTGGAAAACGGTAGCTATCAGGACGTGG TTGCGTTTAAGGTGGTTGATAAACAGCTGCGTATCCAACTG AAGAACGGCAAAGAAACCACCCTGCGTACCCCGGCGCAGT	MSQTITQSRLRIDA NFKRFVDEEVLPGT GLDAAAFWRNFDE IVHDLAPENRQLLA ERDRIQAALDEWH RSNPGPVKDKAAYK SFLRELGYLVPQPER VTVETTGIDSEITSQ AGPQLVVPAMNAR YALNAANARWGSL YDALYGSDIIPQEGA MVSGYDPQRGEQ VIAWVRRFLDESLP LENGSYQDVVAFKV VDKQLRIQLKNGKE TTLRTPAQFVGYRG

	TTGTTGGTTATCGTGGCGATGCGGCGGCGCCGACCTGCATC	DAAAPTCILLKNNG
	CTGCTGAAGAACAACGGTCTGCACATCGAACTGCAAATTGA	LHIELQIDANGRIGK
	CGCGAACGGTCGTATTGGCAAAGACGATCCGGCGCACATC	DDPAHINDVIVEAA
	AACGATGTGATTGTTGAGGCGGCGATCAGCACCATTCTGG	ISTILDCEDSVAAVD
	ACTGCGAGGATAGCGTGGCGGCGGTTGATGCGGAAGATA	AEDKILLYRNLLGL
	AGATCCTGCTGTACCGTAACCTGCTGGGTCTGATGCAGGGC	MQGTLQEKMEKN
	ACCCTGCAAGAAAAGATGGAGAAAAACGGTCGTCAGATTG	GRQIVRKLNDDRHY
	TTCGTAAACTGAACGACGATCGTCACTATACCGCGGCGGAT	TAADGSEISLHGRSL
	GGCAGCGAGATCAGCCTGCACGGTCGTAGCCTGCTGTTTAT	LFIRNVGHLMTIPVI
	TCGTAACGTTGGCCACCTGATGACCATCCCGGTGATTTGGG	WDSEGNEIPEGILD
	ATAGCGAGGGTAACGAGATCCCGGAAGGTATTCTGGACGG	GVMTGAIALYDLKV
	CGTTATGACCGGTGCGATCGCGCTGTACGATCTGAAGGTGC	QKNSRTGSVYIVKP
	AGAAAAACAGCCGTACCGGCAGCGTGTATATTGTTAAGCC	KMHGPQEVAFANK
	GAAAATGCACGGTCCGCAAGAGGTGGCGTTCGCGAACAAG	LFTRIETMLGMAPN
	CTGTTTACCCGTATCGAAACCATGCTGGGCATGGCGCCGAA	TLKMGIMDEERRTS
	CACCCTGAAAATGGGTATTATGGATGAGGAACGTCGTACC	LNLRSCIAQARNRV
	AGCCTGAACCTGCGTAGCTGCATCGCGCAGGCGCGTAACC	AFINTGFLDRTGDE
	GTGTTGCGTTCATTAACACCGGCTTTCTGGACCGTACCGGT	MHSVMEAGPMLR
	GATGAGATGCACAGCGTGATGGAAGCGGGCCCGATGCTGC	KNQMKSTPWIKAY
	GTAAGAACCAGATGAAAAGCACCCCGTGGATCAAGGCGTA	ERNNVLSGLFCGLR
	CGAGCGTAACAACGTTCTGAGCGGTCTGTTCTGCGGTCTGC	GKAQIGKGMWAM
	GTGGCAAGGCGCAAATTGGTAAAGGCATGTGGGCGATGCC	PDLMADMYSQKG
	GGACCTGATGGCGGATATGTACAGCCAGAAAGGCGACCAA	DQLRAGANTAWV
	CTGCGTGCGGGTGCGAACACCGCGTGGGTGCCGAGCCCGA	PSPTAATLHALHYH
	CCGCGGCGACCCTGCATGCGCTGCACTATCACCAGACCAAC	QTNVQSVQANIAQ
	GTGCAGAGCGTTCAAGCGAACATCGCGCAAACCGAGTTCA	TEFNAEFEPLLDDLL
	ACGCGGAGTTTGAACCGCTGCTGGACGATCTGCTGACCATT	TIPVAENANWSAQ
	CCGGTGGCGGAAAACGCGAACTGGAGCGCGCAGGAGATC	EIQQELDNNVQGIL
	CAGCAAGAACTGGATAACAACGTTCAAGGTATTCTGGGCTA	GYVVRWVEQGIGC
	CGTGGTTCGTTGGGTGGAGCAGGGTATCGGCTGCAGCAAG	SKVPDIHNVALME
	GTGCCGGACATTCACAACGTTGCGCTGATGGAAGATCGTG	DRATLRISSQHIAN
	CGACCCTGCGTATCAGCAGCCAACACATTGCGAACTGGCTG	WLRHGILTKEQVQ
	CGTCACGGTATCCTGACCAAGGAGCAGGTTCAAGCGAGCC	ASLENMAKVVDQQ
	TGGAAAACATGGCGAAAGTGGTTGACCAGCAAAACGCGG	NAGDPAYRPMAG
	GCGATCCGGCGTACCGTCCGATGGCGGGTAACTTCGCGAA	NFANSCAFKAASDL
	CAGCTGCGCGTTTAAGGCGGCGAGCGACCTGATCTTTCTGG	IFLGVKQPNGYTEP
	GCGTGAAACAGCCGAACGGTTATACCGAACCGCTGCTGCA	LLHAWRLREKESHH
	TGCGTGGCGTCTGCGTGAGAAAGAAAGCCACCAtATGGGT	MGSGSLEIEAAALE
	AGCGGCAGCCTGGAGATTGAAGCGGCGGCGCTGGAGCAG	QENTALETEVAELE
	GAAAACACCGCGCTGGAAACCGAGGTGGCGGAGCTGGAA	QEVQRLENIVSQYR
	CAGGAAGTGCAACGTCTGGAGAACATTGTGAGCCAATACC	TRYGPLHHHHHH
	GTACCCGTTATGGTCCGCTGCACCACCACCACCACCACTAA	