Self-Assembling Micelles Based on an Intrinsically Disordered Protein Domain

Sarah H. Klass,[1] Matthew J. Smith,[1] Tahoe A. Fiala[1] Jessica P. Lee,[1] Anthony O. Omole, [1] Bong-Gyoon Han,[3] Kenneth H. Downing[3] and Matthew B. Francis[1,2,*]

[1]Department of Chemistry, University of California, Berkeley, [2]Structural Biology and Imaging Department, Life Sciences Division, Lawrence Berkeley National Laboratory, and [3]Materials Sciences Division, Lawrence Berkeley National Laboratories, and Berkeley, California 94720.

**mbfrancis@berkeley.edu*

SUPPORTING INFORMATION

Table of Contents

S7. Supplementary Figures

SI 1. General methods

Unless otherwise noted, all chemicals were obtained from commercial sources and used without further purification. Oligonucleotides were purchased from IDT DNA (USA) with standard desalting and used without further purification.

MilliQ water was purified to a resistivity of 18.2 MΩ·cm (at 25 °C) using a Milli-Q Gradient ultrapure water purification system (Millipore)

High purity grade bovine thrombin protease was obtained from MP Biomedicals(cas: 9002-04-4)

Buffer pH 7.4; 1xPBS, pH 7.4 (ThermoFisher Scientific): 1.06 mM KH_2PO_4, 155.2 mM NaCl, 3.0 mM $Na_2HPO_4•7H_2O$

Buffer pH 5.7: 100 mM phosphate buffer, measured pH: 5.7

Buffer pH 5.3: 100 mM phosphate buffer, adjust to pH 5.3 with 1 M HCl

SI 2. Construction of Plasmids

(a) MBP-IDP plasmid: The inherent repetitive sequence of IDP made the synthesis of one contiguous gene block infeasible. Instead, two gene blocks (gBlocks: IDT Technologies) were synthesized (IDP-1 and IDP-2) with a 32 bp consensus sequence that allowed for Gibson assembly. A 100 ng sample of each gBlock and 10 µL of 2x Gibson master mix was adjusted with water to a volume of 20 µL and was incubated at 50 °C for 60 min. After DNA clean up with QiaQuick (Qiagen), the assembly product was PCR amplified (VENT polymerase from NEB, Tm=61 °C) with forward and reverse primers: 5'- ATA ATA GCT AGC TTA GTT CCT CGT GCC TGG CGT G -3' and 5'- TAT TAT CTC GAG CTA TTA GGC ACA CCA GTA CGG AGA TTT C -3'. The IDT insert contained NheI and XhoI restriction sites, which were double digested, heat inactivated at 80 °C for 5 min and ligated (QuickLigase, NEB) with a 5'-terminal MBP pSKB3 vector. Plating on kanamycin agar plates yielded individual colonies that were cultured, DNA purified (NucleoSpin, MacheryNagel) and sequenced (Quintara BioSciences).

gBlock IDP-1:

ATAATAGCTAGCTTAGTTCCTCGTGCCTGGCGTGGCTCCCCGTGGGCAGAGGCCAAGAGT
CCAGCGGAAGCTAAGTCGCCAGCCGAAGTCAAGTCGCCCGCCGTCGCGAAAAGCCCCGC
AGAGGTGAAATCCCCGGCCGAAGTCAAATCGCCGGCAGAAGCGAAATCCCCGGCAGAAG
CAAAAAGTCCTGCTGAGGTCAAATCGCCAGCAACCGTCAAATCCCCTGGAGAGGCAAAAT
CTCCGGCAGAAGCCAAGTCCCCTGCCGAAGTGAAGTCAC

gBlock IDP-2:

AGAAGCCAAGTCCCCTGCCGAAGTGAAGTCACCTGTCGAAGCCAAGTCGCCGGCCGAAG
CGAAGAGCCCAGCGAGCGTGAAAAGTCCTGGTGAGGCTAAGTCCCCGGCGGAAGCGAAA
TCTCCAGCGGAAGTAAAGAGTCCGGCCACCGTTAAATCCCCGGTAGAGGCCAAAAGCCCT
GCGGAAGTTAAATCGCCGGTGACGGTCAAATCACCCGCGGAAGCGAAGTCCCCGGTGGA
GGTGAAATCTCCGTACTGGTGTGCCTAATAGCTCGAGATAATA

Underlined regions: Consensus sequences used for Gibson Assembly

(b) MBP-IDP-2Y plasmid: Overhang PCR was performed on the MBP-IDP plasmid constructed in (a) above. The forward primer extended the sequence with a Bsa1 cut site while the reverse primer extended the sequence with the desired hydrophobic portion and a Bsa1 cut site to allow for incorporation into a plasmid by Golden Gate assembly. The amplified sequence was run on a 1% agarose gel and confirmed to be of the approximate length. The PCR product was extracted and purified. To perform the Golden Gate assembly, the 2Y PCR product was incubated with the MBP-IDP Golden Gate plasmid, Bsa1, NEB ligase buffer, and ligase enzyme and cycled 25 times. After ligation plasmids were transformed into chemically competent cells and plated on Kanamycin LB agar plates at 37 °C overnight. When the agar plate was exposed to UV light, white colonies were selected (with green fluorescnece indicating no excision of GFP by Bsa1) and grown in 10 mL of LB media at 37 °C overnight. Plasmid DNA was subsequently purified (NucleoSpin, MacheryNagel) and sequenced (Quintara BioSciences).

Forward primer: 5' AGG TCT CTC ATG GCC AGC AGC CAT CAT 3’

Reverse primer: 5' TGG TCT CGT TTA CAC ATA CTG CGC ATA CGC GCC ATA GGC ACA CCA GTA CGG AGA TTT CA 3’

Underlined regions: BsaI cut sites

(c) MBP-IDP-2Yx2A plasmid construction: Overhang PCR was performed on the MBP-IDP-2Y plasmid constructed in (b). The forward primer extended the sequence with a Bsa1 cut site while the reverse primer extended the sequence with a hydrophobic portion and a Bsa1 cut site to allow for incorporation into our plasmid by Golden Gate assembly. Following the same procedure as in (b) the MBP-IDP-2Yx2A plasmid was purified and sequenced.

Forward primer: 5' AGG TCT CTC ATG GCC AGC AGC CAT CAT 3’

Reverse primer: 5' TGG TCT CGT TTA AAT ATA AGC ATA CAG ATA CCA ATA CGC ATA AAT ATA CAC ATA CTG CG 3’

Underlined rgions: BsaI cut sites

(d) MBP-IDP-2Yx3A plasmid construction: Overhang PCR was performed on the MBP-IDP-2Yx2A plasmid constructed in (c). The forward primer extended the sequence with a Bsa1 cut site, while the reverse primer extended the sequence with a hydrophobic portion and a Bsa1 cut site to allow for incorporation into our plasmid by Golden Gate assembly. Following the same procedure as in (c) the MBP-IDP-2Yx3A plasmid was purified and sequenced.

Forward primer: 5' AGG TCT CTC ATG GCC AGC AGC CAT CAT 3’

Reverse primer: 5' TGG TCT CGT TTA AAT ATA AGC ATA CAG ATA CCA ATA CGC ATA AAT ATA CAC ATA CTG CG 3’

Underlined regions: BsaI cut sites

(e) MBP-IDP-2Yx4A plasmid construction: Overhang PCR containing was performed on MBP-IDP-2Yx3A plasmid constructed in (c). The forward primer extended the sequence with a Bsa1 cut site while the reverse primer extended the sequence with a hydrophobic portion and a Bsa1 cut site to allow for incorporation into our plasmid by Golden Gate assembly. Following the same procedure as in (c) the MBP-IDP-2Yx4A plasmid was purified and sequenced.

Forward primer: 5' AGG TCT CTC ATG GCC AGC AGC CAT CAT 3'

Reverse primer: 5' TGG TCT CGT TTA CAG CGC CAC CGC AAT ATA AGC ATA CAG ATA CCA ATA CGC ATA AAT 3'

Underlined regions: BsaI cut sites

(f) MBP-IDP-2Yx2A-C>S plasmid construction: Quick Change PCR was performed on the MBP-IDP-2Yx2A plasmid constructed in (c) above to change the 17th amino acid of the c-terminus (160th residue of IDP-2Yx2A) from a cysteine to a serine residue. After 20 cycles with an annealing temperature of 60°C, the sample was digested with Dpn1 to remove the template plasmid. The plasmid was then transformed into chemically competent cells and plated on Kanamycin LB agar plates at 37 °C overnight. Single colonies were picked and grown in 10 mL of LB media at 37 °C overnight. Plasmid DNA was subsequently purified (NucleoSpin, MacheryNagel) and sequenced (Quintara BioSciences).

Forward primer: 5' AAT CTC CGT ACT GGA GTG CCT ATG GCG 3'

Reverse primer: 5' CGC GCC ATA GGC ACT CCA GTA CGG AGA TT 3'

SI 3. DNA sequences

MBP-IDP:
ATGGCCAGCAGCCATCATCATCATCATCACGATTACGATATCCCAACGACCGAAAA
CCTTTACTTCCAGGGATCCAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACG
GCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACC
GGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGT
TGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGT
GGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGG
ACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCT
TACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAA
CCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAA
GGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGA
TTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAA
GACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGAC
CTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGC
CTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAAC
ATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCA
ACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCG
AACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCT
GGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAG
GAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACTATGGAAAACGCCCAGAAAG
GTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACT

GCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGAC
GCGCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAACAACAACCTCGG
GGCTAGCTTAGTTCCTCGTGCCTGGCGTGGCTCCCCGTGGGCAGAGGCCAAGAG
TCCAGCGGAAGCTAAGTCGCCAGCCGAAGTCAAGTCGCCCGCCGTCGCGAAAAG
CCCCGCAGAGGTGAAATCCCCGGCCGAAGTCAAATCGCCGGCAGAAGCGAAATC
CCCGGCAGAAGCAAAAAGTCCTGCTGAGGTCAAATCGCCAGCAACCGTCAAATCC
CCTGGAGAGGCAAAATCTCCGGCAGAAGCCAAGTCCCCTGCCGAAGTGAAGTCAC
CTGTCGAAGCCAAGTCGCCGGCCGAAGCGAAGAGCCCAGCGAGCGTGAAAAGTC
CTGGTGAGGCTAAGTCCCCGGCGGAAGCGAAATCTCCAGCGGAAGTAAAGAGTC
CGGCCACCGTTAAATCCCCGGTAGAGGCCAAAAGCCCTGCGGAAGTTAAATCGCC
GGTGACGGTCAAATCACCCGCGGAAGCGAAGTCCCCGGTGGAGGTGAAATCTCC
GTACTGGTGTGCCTAA

MBP-IDP-2Yx2A:
ATGGCCAGCAGCCATCATCATCATCATCACGATTACGATATCCCAACGACCGAAAA
CCTTTACTTCCAGGGATCCAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACG
GCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACC
GGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGT
TGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGT
GGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGG
ACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCT
TACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAA
CCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAA
GGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGA
TTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAA
GACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGAC
CTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGC
CTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAAC
ATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCA
ACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCG
AACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCT
GGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAG
GAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACTATGGAAAACGCCCAGAAAG
GTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACT
GCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGAC
GCGCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAACAACAACCTCGG
GGCTAGCTTAGTTCCTCGTGCCTGGCGTGGCTCCCCGTGGGCAGAGGCCAAGAG
TCCAGCGGAAGCTAAGTCGCCAGCCGAAGTCAAGTCGCCCGCCGTCGCGAAAAG
CCCCGCAGAGGTGAAATCCCCGGCCGAAGTCAAATCGCCGGCAGAAGCGAAATC
CCCGGCAGAAGCAAAAAGTCCTGCTGAGGTCAAATCGCCAGCAACCGTCAAATCC
CCTGGAGAGGCAAAATCTCCGGCAGAAGCCAAGTCCCCTGCCGAAGTGAAGTCAC
CTGTCGAAGCCAAGTCGCCGGCCGAAGCGAAGAGCCCAGCGAGCGTGAAAAGTC
CTGGTGAGGCTAAGTCCCCGGCGGAAGCGAAATCTCCAGCGGAAGTAAAGAGTC
CGGCCACCGTTAAATCCCCGGTAGAGGCCAAAAGCCCTGCGGAAGTTAAATCGCC
GGTGACGGTCAAATCACCCGCGGAAGCGAAGTCCCCGGTGGAGGTGAAATCTCC
GTACTGGTGTGCCTATGGCGCGTATGCGCAGTATGTGTATATTTATGCGTATTGGTA
TCTGTAA

MBP-IDP-2Yx3A:
ATGGCCAGCAGCCATCATCATCATCATCACGATTACGATATCCCAACGACCGAAAA
CCTTTACTTCCAGGGATCCAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACG

GCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACC
GGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGT
TGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGT
GGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGG
ACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCT
TACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAA
CCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAA
GGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGA
TTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAA
GACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGAC
CTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGC
CTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAAC
ATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCA
ACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCG
AACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCT
GGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAG
GAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACTATGGAAAACGCCCAGAAAG
GTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACT
GCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGAC
GCGCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAACAACAACCTCGG
GGCTAGCTTAGTTCCTCGTGCCTGGCGTGGCTCCCCGTGGGCAGAGGCCAAGAG
TCCAGCGGAAGCTAAGTCGCCAGCCGAAGTCAAGTCGCCCGCCGTCGCGAAAAG
CCCCGCAGAGGTGAAATCCCCGGCCGAAGTCAAATCGCCGGCAGAAGCGAAATC
CCCGGCAGAAGCAAAAAGTCCTGCTGAGGTCAAATCGCCAGCAACCGTCAAATCC
CCTGGAGAGGCAAAATCTCCGGCAGAAGCCAAGTCCCCTGCCGAAGTGAAGTCAC
CTGTCGAAGCCAAGTCGCCGGCCGAAGCGAAGAGCCCAGCGAGCGTGAAAAGTC
CTGGTGAGGCTAAGTCCCCGGCGGAAGCGAAATCTCCAGCGGAAGTAAAGAGTC
CGGCCACCGTTAAATCCCCGGTAGAGGCCAAAAGCCCTGCGGAAGTTAAATCGCC
GGTGACGGTCAAATCACCCGCGGAAGCGAAGTCCCCGGTGGAGGTGAAATCTCC
GTACTGGTGTGCCTATGGCGCGTATGCGCAGTATGTGTATATTTATGCGTATTGGTA
TCTGTATGCTTATATTTAA

MBP-IDP-2Yx2A-S/C

ATGGCCAGCAGCCATCATCATCATCATCACGATTACGATATCCCAACGACCGAAAACCTTT
ACTTCCAGGGATCCAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAGG
CTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACC
GTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGGC
CCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGG
CTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCG
TACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTA
TAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAA
AGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCAC
CTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGA
CATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGA
CCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTTT
AATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACC
AGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCGT
TCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAAG

AGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACC
GCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGC
CGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGC
TTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGA
TGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAAC
AACAACCTCGGGGCTAGCTTAGTTCCTCGTGCCTGGCGTGGCTCCCCGTGGGCAGAGGC
CAAGAGTCCAGCGGAAGCTAAGTCGCCAGCCGAAGTCAAGTCGCCCGCCGTCGCGAAAA
GCCCCGCAGAGGTGAAATCCCCGGCCGAAGTCAAATCGCCGGCAGAAGCGAAATCCCCG
GCAGAAGCAAAAAGTCCTGCTGAGGTCAAATCGCCAGCAACCGTCAAATCCCCTGGAGAG
GCAAAATCTCCGGCAGAAGCCAAGTCCCCTGCCGAAGTGAAGTCACCTGTCGAAGCCAAG
TCGCCGGCCGAAGCGAAGAGCCCAGCGAGCGTGAAAAGTCCTGGTGAGGCTAAGTCCCC
GGCGGAAGCGAAATCTCCAGCGGAAGTAAAGAGTCCGGCCACCGTTAAATCCCCGGTAGA
GGCCAAAAGCCCTGCGGAAGTTAAATCGCCGGTGACGGTCAAATCACCCGCGGAAGCGA
AGTCCCCGGTGGAGGTGAAATCTCCGTACTGGAGTGCCTATGGCGCGTATGCGCAGTATG
TGTATATTTATGCGTATTGGTATCTGTAA

SI 4. Amino Acid Sequences

MBP-IDP: ASSHHHHHHDYDIPTTENLYFQGSKIEEGKLVIWINGDKGYNGLAEVGKK
FEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLY
PFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPY
FTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNK
GETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENY
LLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAV
INAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGASLVPRAWRGSPWAEAKSPAEAKSPA
EVKSPAVAKSPAEVKSPAEVKSPAEAKSPAEAKSPAEVKSPATVKSPGEAKSPAEAKSPAEVK
SPVEAKSPAEAKSPASVKSPGEAKSPAEAKSPAEVKSPATVKSPVEAKSPAEVKSPVTVKSPA
EAKSPVEVKSPYWCA

(MBP portion in blue)

Molecular weight (after N-terminal met cleavage): 61683.50 Da

IDP: AWRGSPWAEAKSPAEAKSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPAEAK
SPAEVKSPATVKSPGEAKSPAEAKSPAEVKSPVEAKSPAEAKSPASVKSPGEAKSPAEAKSPA
EVKSPATVKSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPYWCA

Molecular weight (after N-terminal met cleavage): 16238.37 Da

MBP-IDP-2Yx2A: ASSHHHHHHDYDIPTTENLYFQGSKIEEGKLVIWINGDKGYNGLAEV
GKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQD
KLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQ
EPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAA
FNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFL
ENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVR
TAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGASLVPRAWRGSPWAEAKSPAEAKS
PAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPAEAKSPAEVKSPATVKSPGEAKSPAEAKSPAE

VKSPVEAKSPAEAKSPASVKSPGEAKSPAEAKSPAEVKSPATVKSPVEAKSPAEVKSPVTVKS
PAEAKSPVEVKSPYWCAYGAYAQYVYIYAYWYL

(MBP portion in blue, 2Yx2A in green)

Molecular weight (after N-terminal met cleavage): 63604.62 Da

IDP-2Yx2A: AWRGSPWAEAKSPAEAKSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPA
EAKSPAEVKSPATVKSPGEAKSPAEAKSPAEVKSPVEAKSPAEAKSPASVKSPGEAKSPAEAK
SPAEVKSPATVKSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPYWCAYGAYAQYVYIYAYWY
L

(2Yx2A in green)

Molecular weight: 18290.69 Da

MBP-IDP-2Yx3A: ASSHHHHHHDYDIPTTENLYFQGSKIEEGKLVIWINGDKGYNGLAEV
GKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQD
KLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQ
EPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAA
FNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFL
ENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVR
TAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGASLVPRAWRGSPWAEAKSPAEAKS
PAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPAEAKSPAEVKSPATVKSPGEAKSPAEAKSPAE
VKSPVEAKSPAEAKSPASVKSPGEAKSPAEAKSPAEVKSPATVKSPVEAKSPAEVKSPVTVKS
PAEAKSPVEVKSPYWCAYGAYAQYVYIYAYWYLYAYI

(MBP in blue, 2Yx3A in green)

Molecular weight (N-terminal met cleavage): 64115.21 Da

IDP-2Yx3A: AWRGSPWAEAKSPAEAKSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPA
EAKSPAEVKSPATVKSPGEAKSPAEAKSPAEVKSPVEAKSPAEAKSPASVKSPGEAKSPAEAK
SPAEVKSPATVKSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPYWCAYGAYAQYVYIYAYWY
LYAYI

(2Yx3A in green)

Molecular weight: 18801.28 Da

MBP-IDP-2Yx4A:
ASSHHHHHHDYDIPTTENLYFQGSKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP
DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLI
AYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAF
KYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAW
SNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKD

KPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEA
LKDAQTNSSSNNNNNNNNNNLGASLVPRAWRGSPWAEAKSPAEAKSPAEVKSPAVAKSPAE
VKSPAEVKSPAEAKSPAEAKSPAEVKSPATVKSPGEAKSPAEAKSPAEVKSPVEAKSPAEAKS
PASVKSPGEAKSPAEAKSPAEVKSPATVKSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPYW
CAYGAYAQYVYIYAYWYLYAYIAVAL

(MBP in blue, 2Yx3A in green)

Molecular weight (after N-terminal met cleavage): 64600.85 Da

IDP-2Yx4A:
AWRGSPWAEAKSPAEAKSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPAEAKSPAEVKSPA
TVKSPGEAKSPAEAKSPAEVKSPVEAKSPAEAKSPASVKSPGEAKSPAEAKSPAEVKSPATVK
SPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPYWCAYGAYAQYVYIYAYWYLYAYIAVAL

(2Yx4A in green)

Molecular weight: 19155.73 Da

MBP-IDP-2Yx2A-C>S

ASSHHHHHHDYDIPTTENLYFQGSKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP
DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLI
AYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAF
KYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAW
SNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKD
KPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEA
LKDAQTNSSSNNNNNNNNNNLGASLVPRAWRGSPWAEAKSPAEAKSPAEVKSPAVAKSPAE
VKSPAEVKSPAEAKSPAEAKSPAEVKSPATVKSPGEAKSPAEAKSPAEVKSPVEAKSPAEAKS
PASVKSPGEAKSPAEAKSPAEVKSPATVKSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPYW
SAYGAYAQYVYIYAYWYL

(MBP in blue, 2Yx2A in green, C>S mutant in red)

Molecular weight (after N-terminal met cleavage): 63588.56 Da

MBP-IDP-2Yx2A-C>S

AWRGSPWAEAKSPAEAKSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPAEAKSPAEVKSPA
TVKSPGEAKSPAEAKSPAEVKSPVEAKSPAEAKSPASVKSPGEAKSPAEAKSPAEVKSPATVK
SPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPYWSAYGAYAQYVYIYAYWYL

(2Yx2A in green, C>S mutant in red)

Molecular weight: 18274.62 Da

SI 5. Instrumentation and sample analysis

Agarose gel electrophoresis: overhang PCR products were analyzed in a 1% agarose gel in TAE buffer (40 mM Tris, 20 mM acetic acid, and 1 mM EDTA) pre-stained with SYBR Safe DNA Gel Stain (ThermoFisher Scientific). Gels were run for 30 min at 120 V. Agarose gels were imaged on a BioRad GelDoc EZ Imager. Commercially available markers (ThermoFisher Scientific) were run on each gel for the assignment of apparent DNA size. Bands of the appropriate size (~2 kb) were excised and DNA was purified using a QIAquick gel extraction kit (QIAGEN).

DNA and Protein quantification: DNA and Protein concentration was determined by measuring the absorbance at A260 and A280 on a nanodrop 1000 series spectrometer (ThermoFisher Scientific). Protein concentration is then determined using calculated extinction coefficients for proteins.

LC-MS analysis: Proteins were analyzed using an Agilent 1200 series liquid chromatography (Agilent Technologies, USA) that was connected in-line with an Agilent 6224 Time-of-Flight (TOF) LC/MS system equipped with a turbo spray ion source.

SDS PAGE gel analysis: Protein samples were run on 4-12% precast linear gradient polyacrylamide gels (Bio-Rad) using a Novex apparatus. Prior to loading, protein samples were mixed with 6X loading buffer and heated for 5 min at 95 °C. Gels were run for 45 min at 200 V in MES buffer. Commercially available markers (Bio-Rad pre-stained precision plus, Novex Sharp pre-stained protein standard) were run on each gel for the assignment of apparent molecular masses. Visualization of protein bands was achieved by staining with Coomassie Brilliant Blue R-250 (Bio-Rad). ImageJ was used to quantify protein purity by gel densitometry.

SAXS: IDP-2Yx2A and the non-assembling control IDP samples were prepared at a range of concentrations from 68 μM down to 1 μM. Only samples above 16 μM provided sufficient signal to be processed. Samples were analyzed by high-throughput small angle X-ray scattering (HT-SAXS) at the SIBYLS beamline at the Advanced Light Source in Berkeley CA. Samples were exposed with a 10^{11} photon/s, 12 keV monochromatic beam in a series of exposures: 0.5, 1.0, 2.0, and 4.0 s. Samples containing only buffer was also run and used for background subtraction.[1] Data were first analyzed using the SAXS FrameSlice web app and then imported into the ScÅtter and SASview applications for further size and shape analysis.

Unstained Transmission Electron Microscopy (TEM): Samples were prepared for TEM analysis by applying the protein solution to hydrophilized Formvar carbon-coated copper grids for 2 min, followed by triple rinsing with dd-H_2O. Images were obtained at the Berkeley electron microscope lab using an FEI Tecnai 12 transmission electron microscope with 120 kV accelerating voltage. The dark contrast on the gray background indicates the presence of the heavy metal catalyst.

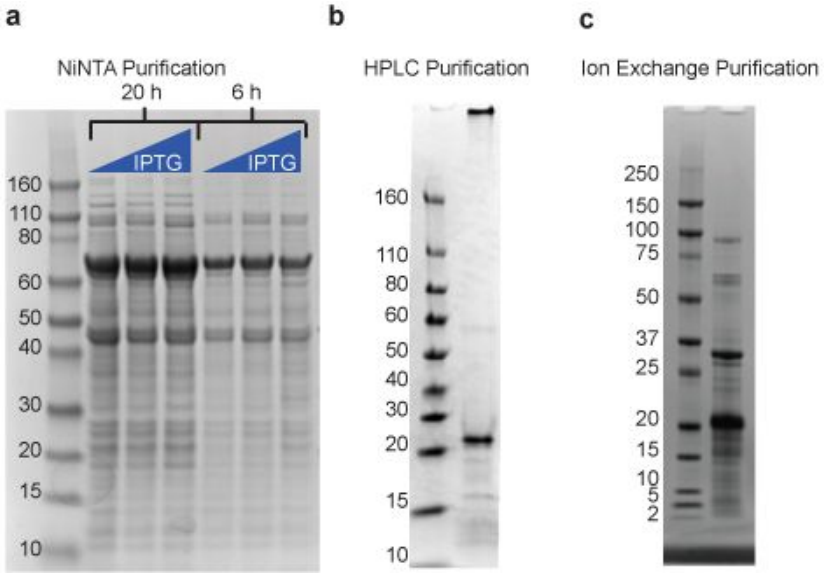
Fluorescence detection of tetraphenylporphyrin (TPP): The emission spectrum of TPP was collected on a Fluoromax-4 Spectrofluorometer (HORIBA Scientific) exciting at 420 nm with a 5 nm window and monitoring emission from 600-800 nm. In water, TPP is insoluble and non-fluorescent, when in a solution containing IDP-2Yx2A, the emission peaks of TPP 650 nm and 715 nm are observed.

HPLC analysis of pyraclostrobin and IDP-2Yx2A. HPLC analysis was performed on an Agilent 1100 Series HPLC system (Agilent Technologies, USA). Reversed phase chromatography for IDP-2Yx2A and pyraclostrobin was performed on a Poroshell 120 EC-C18 2.7 µm 4.6 x 50 mm column (Agilent Technologies, USA). Sample analysis was achieved with an inline diode array detector (DAD) and an inline fluorescence detector (FLD). Monitored wavelengths were 230 nm, 260 nm, 280 nm, and 300 nm. Samples were eluted using a gradient from 95% solvent A: 0.1% TFA in water / 5% solvent B: 0.1% TFA in acetonitrile to 5% solvent A / 95% solvent B. The areas under the curves corresponding to pyraclostrobin or IDP-2Yx2A, identified by their retention times, were compared to a calibration curve constructed from prepared standards at known concentrations.

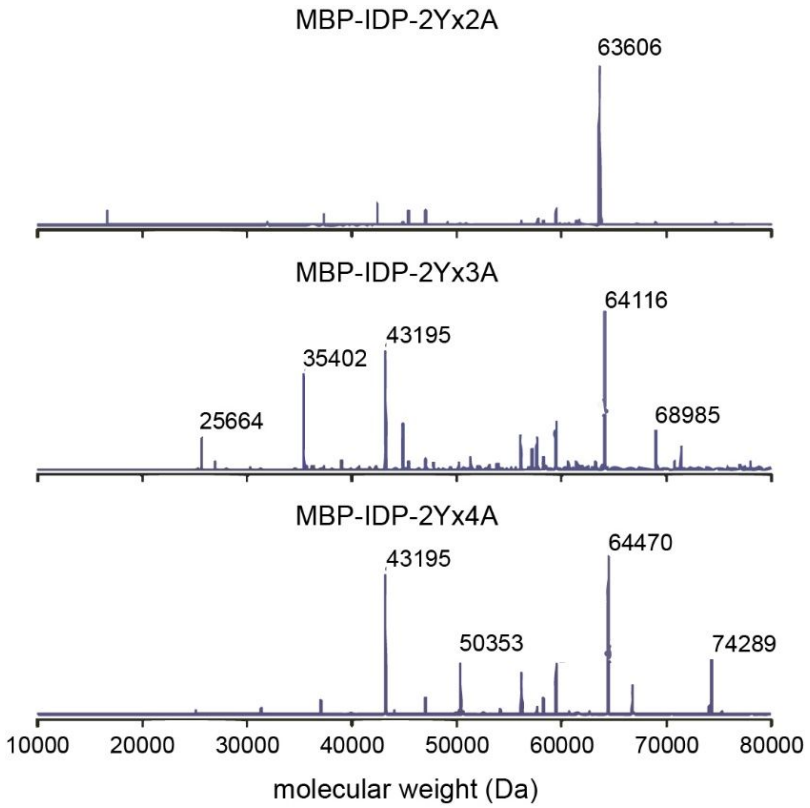
SI 6. Biodegradation analysis of IDP-2Yx2A

A 100 µL portion of HPLC purified IDP-2Yx2A (10 µM) was left on a benchtop exposed to air. This promoted disulfide formation, and thus remaining IDP-2Yx2A observed in these samples had a molecular weight of 36580 Da. Five 20 µL time points were collected: 1 day, 2 days, 7 days, 21 days, and 50 days. A 15 µL portion of each sample was analyzed using SDS PAGE gel electrophoresis and LC-MS to determine the degree of degradation by gel densitometry using the ImageJ application. The remaining 5 µL of sample was analyzed via LC-MS, both for the presence of extra peaks in the LC trace and for the presence of the remaining starting proteins. (SI Figure 9).

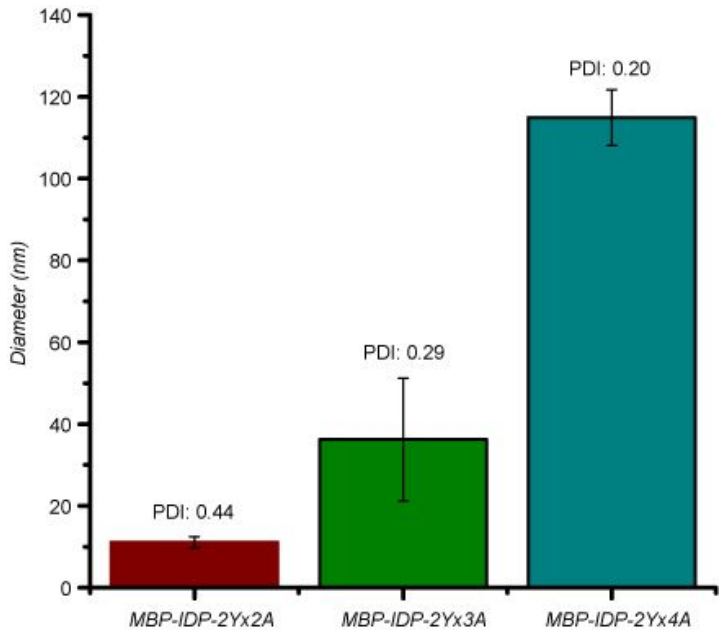
SI 7. Supplementary Figures



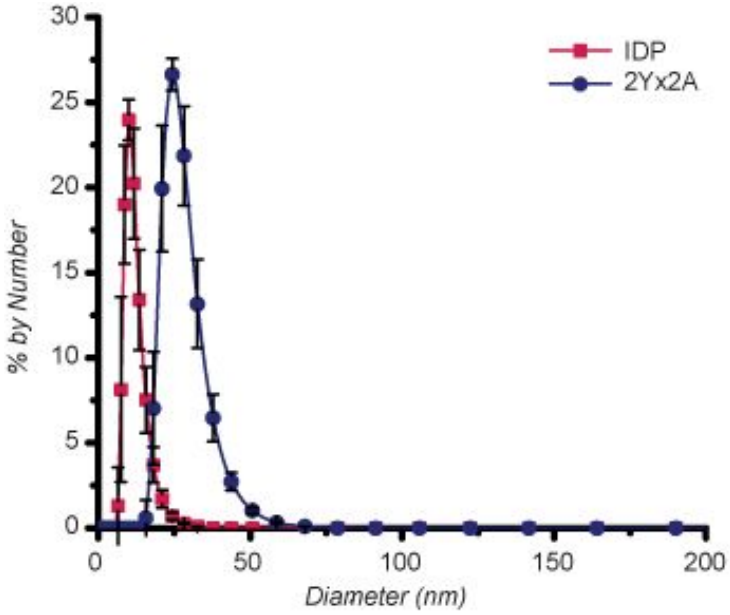
SI Figure 1: SDS PAGE analysis of MBP-IDP-2Yx2A and MBP-IDP-2Yx2A throughout the purification protocol. (a) NiNTA expression conditions for MBP-IDP-2Yx2A. Cultures were induced with 0.1, 0.2 or 0.5 mM IPTG and expressed at 16 °C for 6 h or 20 h. Samples were then purified by NiNTA chromatography. Optimal expression conditions for purity involved inducing with 0.1 mM IPTG and expressing for 6 h at 16 ºC. (b) After cleavage by thrombin the MBP could be removed by HPLC purification to yield >95% pure IDP-2Yx2A, apparent mw ~ 22kDa with the presence of intact aggregate at top of gel. (c) MBP could also be removed by anion exchange chromatography in higher throughput, but at lower purity ~75% pure IDP-2Yx2A.



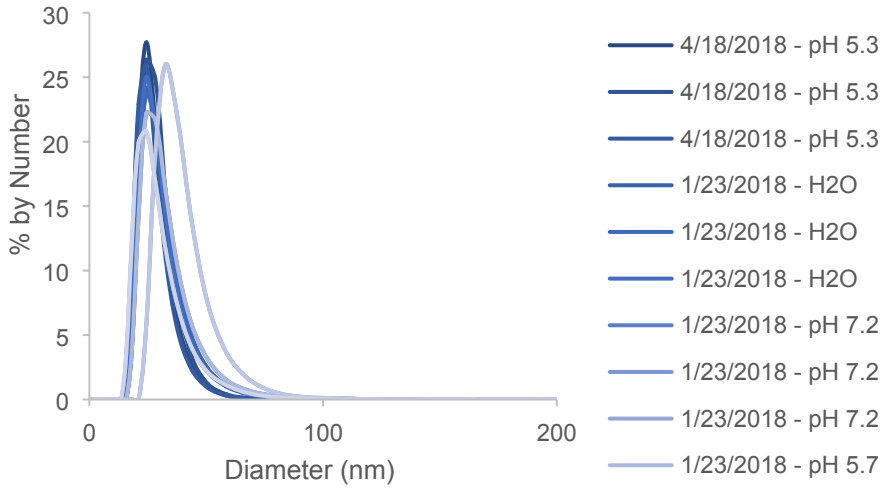
SI Figure 2: LCMS analysis of MBP-IDP-2Yx(2-4)A-MBP mutants after NiNTA purification. Only MBP-IDP-2Yx2A (mw: 63606 Da) was able to be expressed and purified in a relatively pure form without the presences of truncations. MBP-IDP-2Yx3A (mw: 64116 Da) and MBP-IDP-2Yx4A (64470 Da) are eluted from NiNTA column along with large amounts of impurities.



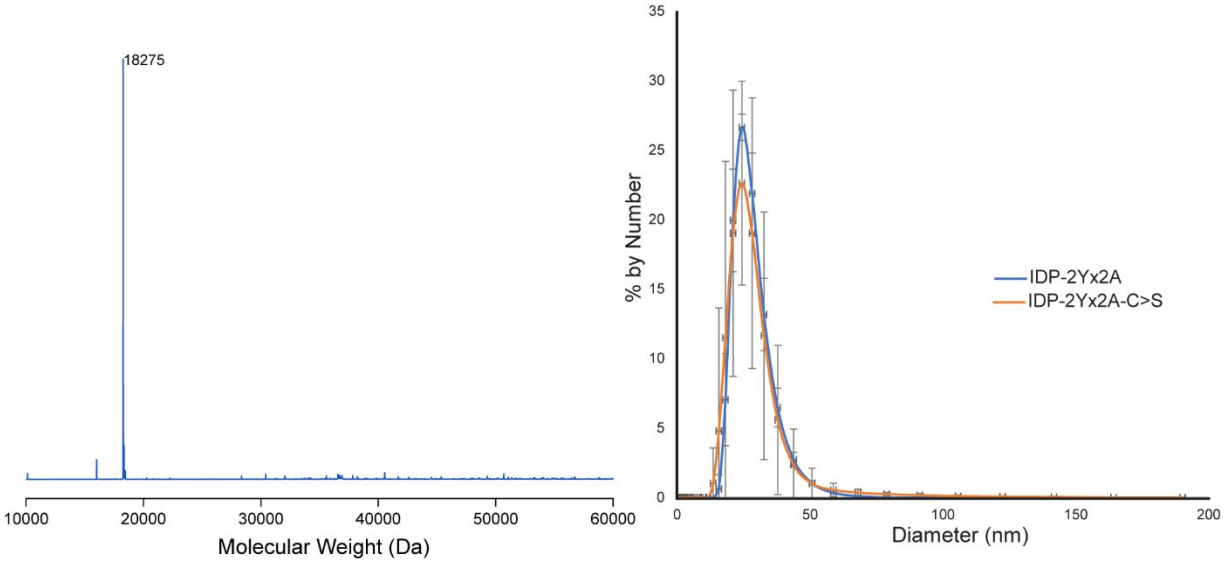
SI Figure 3: DLS analysis of MBP-IDP-2Yx(2-4)A-MBP mutants. MBP-IDP-2Yx2A shows an average diameter of 11.1 ± 1.3 nm, which is similar to that of monomeric IDP (11.3 ± 0.8 nm). MBP-IDP-2Yx3A (36.3 ± 15 nm) and MBP-IDP-2Yx4A (115 ± 6 nm) show much larger diameters, likely indicating aggregation prior to cleavage of MBP.



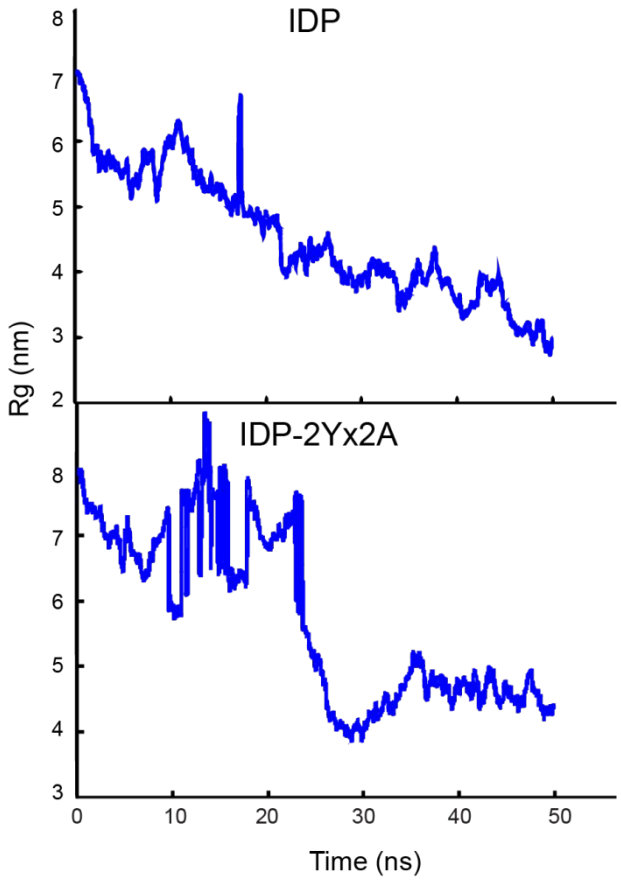
SI Figure 1: DLS curves for IDP vs. IDP-2Yx2A. Proteins diameters were monitored by "% by number". Error bars represent the standard deviation at a given diameter over the course of 3 measurements for each sample.



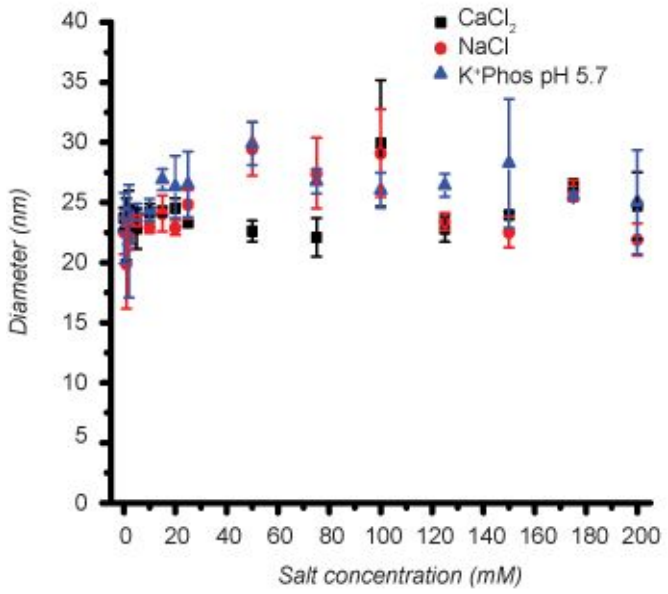
SI Figure 5: Batch to Batch reproducibility. Twelve representative DLS profiles of IDP-2Yx2A micelles taken from different expression and purification batches performed in April or January of 2018. Additionally, these traces are representative a different pH and buffer conditions.



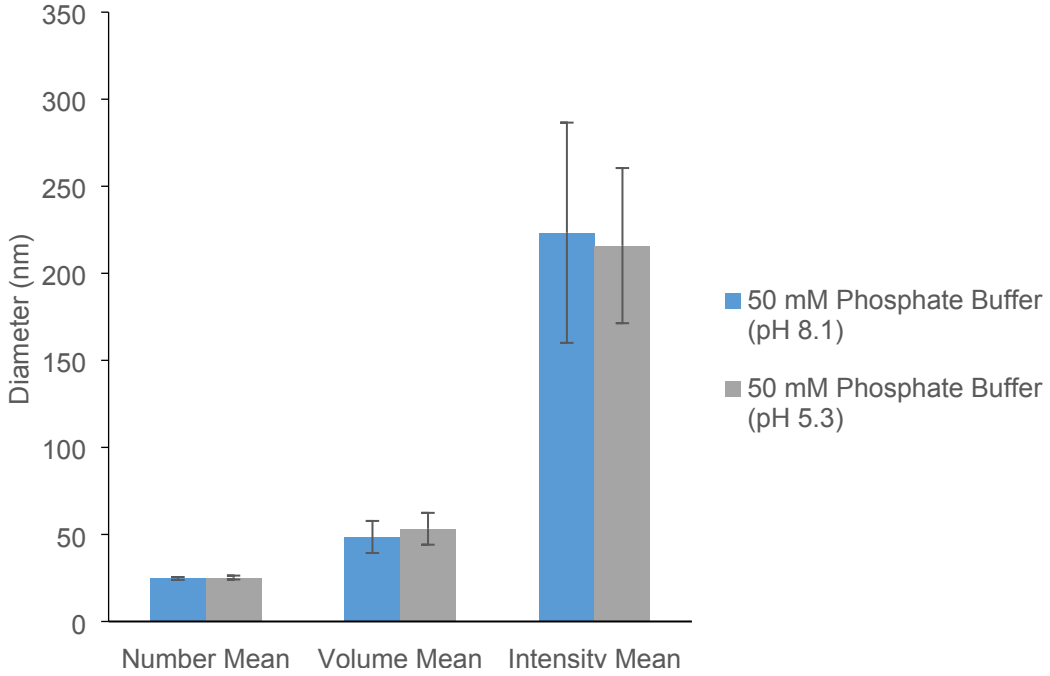
SI Figure 6: LC-MS and DLS analysis of IDP-2Yx2A-C>S mutant. LC-MS analysis of purified IDP-2Yx2A-C>S expected mw: 18274.62 Da, observed: 18275 Da. Lyophilized proteins were resuspended in MilliQ H_2O and analyzed by DLS. IDP-2Yx2A has a diameter of 26.9 ± 4.6 nm. The IDP-2YX2A-C/S mutant has a diameter of 26.5 ± 4.9 nm.



SI Figure 7: Radius of Gyration Analysis of Molecular Dyanmic Simulations. MD simulations of IDP and 2Yx2A were carried out using the Desmond software package and run with the OPLS_2005 force field[35] available through Maestro. Both 2Yx2A and IDP were modeled in an alpha helical conformation at time zero. Due the existence of 16 extra amino acids present on the c-terminus of 2Yx2A, it has a starting Rg that is larger than that of IDP. Simulations were carried out at 300 K and a constant pressure of 1.01325 bar. Each structure was fully solvated with SPC water in a cube with an edge length of 20Å. Each simulation was run for a total of 50 ns. The Rg was tracked over time with respect to the protein back bone and side chain residues. Between 40 – 50 ns, the Rg of IDP fluctuates between 3 – 4 nm while 2Yx2A fluctuates between 4 and 5 nm.

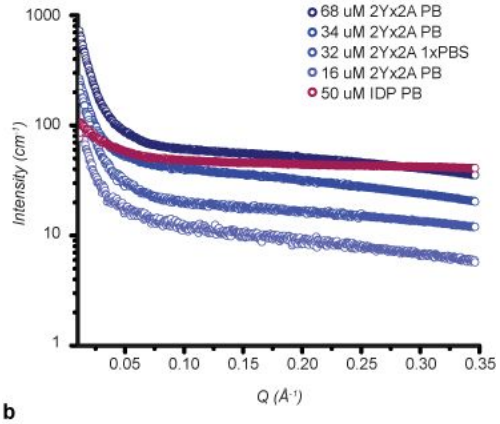


SI Figure 8: Impact of salt on IDP-2Yx2A micelle sizes. DLS measurements of IDP-2Yx2A incubated in buffer conditions containing between 0 and 200 mM $CaCl_2$, NaCl, and potassium phosphate buffer at pH 5.7. Over all conditions tested, the average diameter stayed between 20-30 nm, with no obvious trends being observed.



SI Figure 9: %Intensity, %Volume, %Number plots for IDP-2Yx2A. In our analysis we use data derived from % Number as it most accurately represents the size of the most abundant species in our sample. However, the presence of larger structures or inter-micelle interactions can be observed by the % Intensity and % Volume traces of these samples.

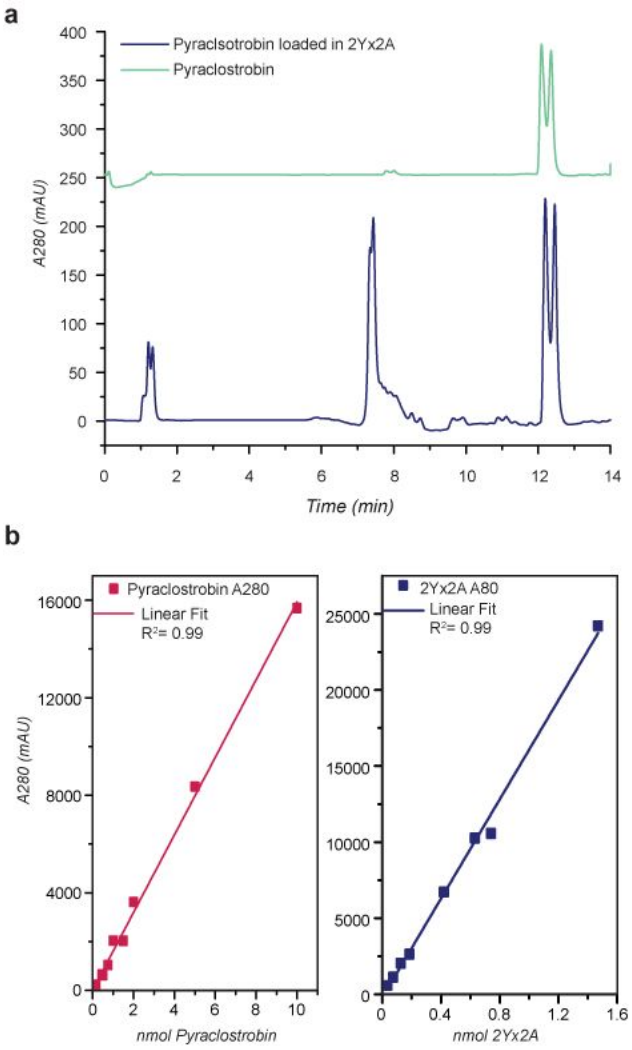
a



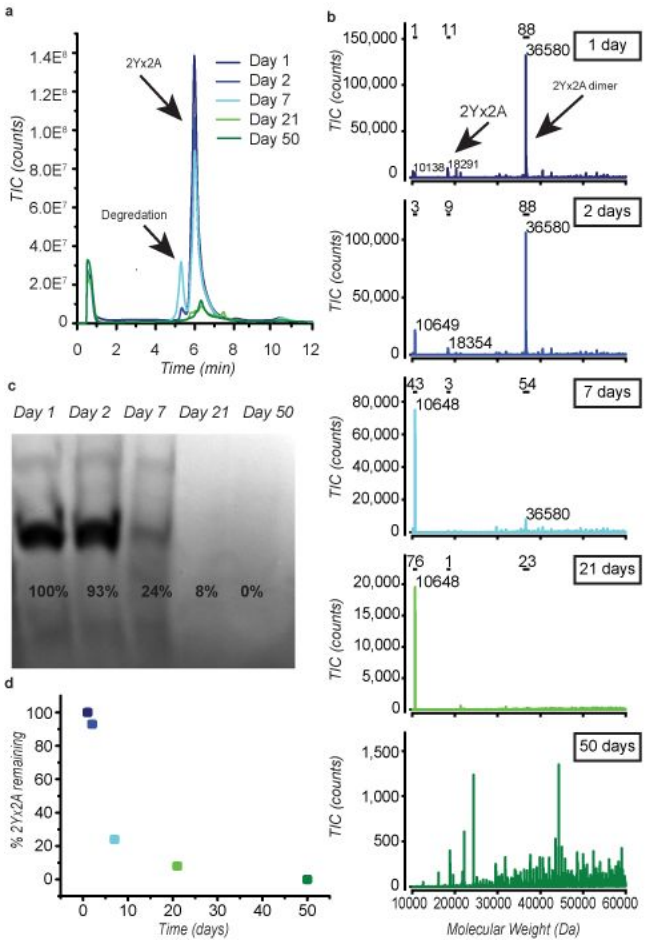
b

Sample	Rg (nm)	Rc (nm)
68 uM 2Yx2A PB	>10.3	5.2
34 uM 2Yx2A PB	>10.2	3.6
32 uM 2Yx2A 1xPBS	>9.0	4.4
16 uM 2Yx2A PB	>9.4	4.5
50 uM IDP PB	>4.0	N/A

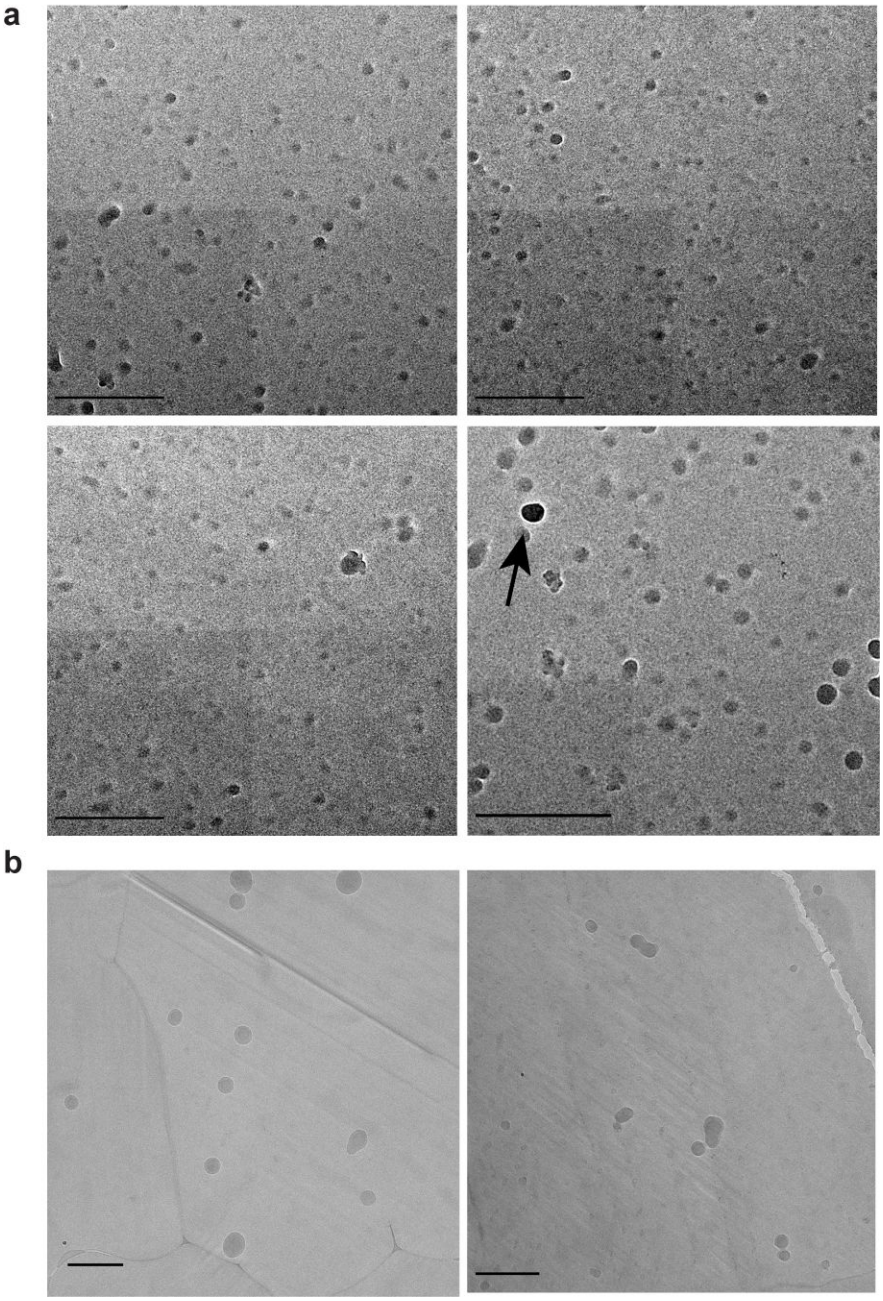
SI Figure 10: SAXS analysis of IDP-2Yx2A and IDP. Four concentrations and two buffers were analyzed for IDP-2Yx2A, all showing a large increase in the low q region in comparison to IDP itself. This indicates the presence of aggregates and larger particles in the IDP-2Yx2A samples. At the high concentrations required to run SAXS, there is likely the formation of larger particles and extended structures. This non-homogeneity makes fitting the data difficult. What can be extracted is minimal Guinier region radii of gyration (Rg) values and radii of the cross section (Rc) values. IDP-2Yx2A shows larger Rg values than IDP, and while Rc values could be obtained for all IDP-2Yx2A samples, those for IDP could not be calculated. This indicates a clear difference in the samples that can be attributed to the hydrophobic tail extension of IDP-2Yx2A.



SI Figure 11: Quantification of pyraclostrobin and IDP-2Yx2A by HPLC. (a) HPLC trace of Pyraclostrobin stock solutions (top) and pyraclostrobin loaded into IDP-2Yx2A (bottom). The peak at 1 min is the injection void volume. The peak at 7.5 min corresponds to the elution of IDP-2Yx2A, while the peak at 12.5 min corresponds to the elution of pyraclostrobin. (b) HPLC calibration curves were constructed for both IDP-2Yx2A and pyraclostrobin by injecting known concentrations of both species and integrating the peak areas. These calibration curves were then used to determine the loading of pyraclostrobin in IDP-2Yx2A micelle samples.



SI Figure 12: Biodegradation analysis of IDP-2Yx2A. (a) By LC-MS, the emergence of a new peak in the HPLC trace corresponding to cleavage products of IDP-2Yx2A was observed at a slightly earlier retention time than IDP-2Yx2A itself (degradation). By 7 days, this peak was especially prominent. (b) Deconvoluted molecular weights and integrated values are shown for IDP-2Yx2A and its degradation products. The IDP-2Yx2A dimer can be observed at the expected mass of 36580 Da. Over time the emergence of lower molecular weight species increased. By 21 days practically no remaining IDP-2Yx2A was observable. (c) SDS PAGE analysis was performed on each degradation sample. (d) Gel densitometry with the ImageJ application was used to quantify the loss of protein relative to the IDP-2Yx2A concentration present on day 1 (24 h) at room temperature.



SI Figure 13: Additional Cryo-TEM images of IDP-2Yx2A. (a) A 6.5 µM solution of IDP-2Yx2A micelles (10 mM phosphate buffer pH 5.7) was imaged on holey gold girds coated with thin carbon film. The presence of many spherical particles is observed, confirming the shape and low CMC of IDP-2Yx2A micelles. There are additional contaminates present in many of the photos that appear as darker circles such as the one identified by the dark arrow. (b) A more dilute sample (0.4 µM solution of 2Yx2A micelles in 100 mM phosphate buffer pH 5.3) was also examined embedded in vitrified ice on uncoated carbon grids. All scale bars are set to 200 nm.