Preparation of Bioderived and Biodegradable Surfactants Based on an Intrinsically Disordered Protein Sequence

Sarah H. Klass,[†] Jamie M. Gleason,[†] Anthony O. Omole,[†] Bibiana Onoa,[‡] Carlos J. Bustamante,[‡,¶,§] and Matthew B. Francis[*,†,‖]

†Department of Chemistry, University of California Berkeley, Berkeley, California 94720, United States

‡Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA

¶Institute for Quantitative Biosciences, University of California, Berkeley, CA 94720

§Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

‖Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

E-mail: mbfrancis@berkeley.edu

Abstract

Surfactants, block-copolymers, and other types of micellar systems are used in a wide variety of biomedical and industrial processes. However, most commonly used surfactants are synthetically derived and pose environmental and toxicological concerns throughout their product life cycle. Because of this, bio-derived and bio-degradable surfactants are promising alternatives. For bio-surfactants to be implemented industrially, they need to be produced on a large scale and also have tailorable properties that match those afforded by the polymerization of synthetic surfactants. In this paper, a scalable

1

and versatile production method for bio-surfactants based on a hydrophilic intrinsically disordered protein (IDP) sequence with a genetically engineered hydrophobic domain is used to study variables that impact their physicochemical and self-assembling properties. These amphiphilic sequences were found to self-assemble into micelles over a broad range of temperatures, pH values, and ionic strengths. To investigate the role of the IDP hydrophilic domain on self-assembly, variants with increased overall charges and systematically decreased IDP domain lengths were produced and examined for their sizes, morphologies, and critical micelle concentrations (CMCs). The results of these studies indicate that decreasing the length of the IDP domain and, consequently, the molecular weight and hydrophilic fraction, leads to smaller micelles. Additionally, significantly increasing the amount of charged residues in the hydrophilic IDP domain results in micelles of similar sizes, but with higher CMC values. This represents an initial step in developing a quantitative model for the future engineering of bio-surfactants based on this IDP sequence.

Introduction

Most commercially used surfactants are derived from hydrocarbon sources, and encompass both conventional surfactants consisting of simple ionic head groups attached to hydrocarbon tails as well as macromolecular block-copolymers consisting of hydrophilic and hydrophobic domains. The properties of commonly used synthetic surfactants have been heavily researched and they have well-understood physicochemical properties that can be tailored to address specific application needs.[1–4] By altering the hydrophobicities, chain lengths, charges, and geometries, the bulk self-assembly properties in a known context can be predictably modified.[5–10] Unfortunately, many synthetically derived surfactants show human or environmental toxicity due to their innate activities or those of their degradation products.[11–13] Due to the widespread use of surfactants, there have been renewed efforts to develop bio-derived and bio-degradable alternatives.[14,15] However, large-scale production of designer

bio-surfactants for medical or environmental purposes remains a challenge in terms of economic viability, compared to their synthetic counterparts.[16,17] Furthermore, the design rules that govern the ability of bio-molecules to perform as surfactants are poorly understood, which impedes the development of novel constructs with direct applications.

Intrinsically disordered proteins (IDPs) mimic the disordered nature of the polymer chains of synthetic block-copolymers, thus making them promising building blocks for bio-derived and bio-degradable alternatives. Unlike synthetic polymers, which generally comprise just one or two monomeric building blocks, IDPs can be produced from a toolbox consisting of the 20 naturally occurring amino acids, each with unique hydrophobicity, charge, and aromaticity properties. Additionally, protein-based polymers also benefit from precise monomer-level control over polymer chain length and position —a feat that is extremely difficult to achieve in synthetic polymerization reactions.[1] One of the most commonly studied types of IDPs for bio-material applications are elastin-like proteins (ELPs), which undergo temperature-dependent phase transitions from monomeric proteins to globular assemblies. The temperature at which this phase transition occurs can be altered by changing the identity of the variable amino acid X (where $X \neq P$) in the repeating ELP sequence motif $(VPGXG)_n$.[18–22] Constructs with favorable transition temperatures have been used in a wide range of applications in drug delivery and as bio-materials.[23–27] However, ELPs only constitute a small portion of all IDP sequences, which are abundant across all domains of life and display a diverse range of properties and functions.[28,29] One such IDP is a highly hydrophilic sequence derived from the neurofilament heavy arm side-chain protein found in human neurons (Figure 1a). This IDP has been shown to undergo expansion and contraction in response to charge alteration by pH, salt, and phosphorylation.[30–34] Furthermore, while the IDP protein (termed 1.0-IDP from herein) exists as a monomer at room temperature, DLS analysis over a broad range of pH and buffer conditions indicates that the protein undergoes a thermo-responsive aggregation at temperatures above 50 °C (Figure 1b). This behavior is similar to commonly used hydrophilic polymers such as poly(ethylene oxide) (PEO) and ELPs.[18–22,35] Using the 1.0-IDP sequence

3

as a hydrophilic domain, we previously added a genetically encoded hydrophobic domain (2Yx2A) at the C-terminus of the protein. This genetic fusion, 1.0-IDP-2Yx2A, was solubly expressed in *E. coli* and after purification was found to undergo controlled self-assembly at room temperature to form nanoscale micellar structures in all tested buffer and temperature conditions (Figure 1c).[36] Due to its ability to encapsulate hydrophobic cargo at low protein concentrations, the 1.0-IDP-2Yx2A sequence acts as a biological mimic of synthetic block-copolymers and possibly opens the door to a promising new class of bio-surfactants based on this IDP sequence.

Although suitable for initial studies, the soluble expression system was not amenable to substantial mutations to the sequence space and thus limited the analysis of factors that impact the self-assembling properties of the micelles. By switching to an insoluble expression system, the yield, stability, and purification process improved along with the tolerance to multisite mutations and deletions to the sequence space, resulting in significant changes to the protein net charge and molecular weight. The ability to alter and create a diverse sequence space is key to developing a systematic approach to understanding how alterations to the charge, relative hydrophobicity, and hydrophilic domain length affect the self-assembly behavior and physicochemical properties of these IDP-based bio-surfactants.
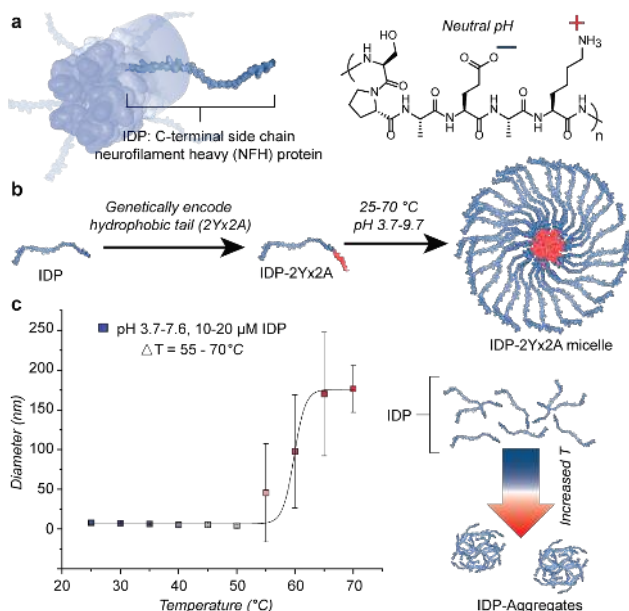
Figure 1: Generation of self-assembling IDP micelles. (a) The intrinsically disordered protein (IDP) domain derived from the human neurofilament protein protrudes from the neurofilament core. The 1.0-IDP sequence is highly hydrophilic and is largely comprised of the repetitive motif SPXEXK, where X is G, A, or V. At neutral pH, this results in two oppositely charged residues per repeat: a negatively charged glutamic acid (E) and a positively charged lysine (K). (b) When heated in aqueous solutions (pH 3.7-9.7, 0-200 mM PB), the monomeric 1.0-IDP aggregates into larger structures, as observed by DLS and pyrene excimer emergence (SI Figure 1). (c) By genetically encoding a hydrophobic 20 amino acid segment (2Yx2A) at the C-terminus of IDP, the new amphiphilic protein 1.0-IDP-2Yx2A was found to self assemble spontaneously into ordered micelle structures in a range of buffer and temperature conditions.

As this IDP sequence has not been extensively studied in an assembly context, the first variables we sought to examine were the nature of the charge and length of the 1.0-IDP sequence itself. Herein, we discuss the development of an insoluble expression and purification platform that was used to create an array of IDP-2Yx2A mutational variants with systematically decreasing IDP lengths while maintaining a neutral charge and the hydrophobic 2Yx2A assembly domain. Additionally, a "supercharged"[37] variant that maintained IDP length, but with double the overall positive charge, was also constructed. These variants were then used to investigate the role of the IDP length and charge on physicochemical properties such as size, morphology, and critical micelle concentration (CMC). From these studies, several trends have emerged.

5

Results and discussion

To gain access to a wider variety of mutations, while also increasing yield, and lowering the cost of production, DNA encoding the 1.0-IDP-2Yx2A sequence was inserted into a pET-31b(+) plasmid at the C-terminus of the inclusion body directing fusion protein KSI. A single methionine residue was included between these regions to allow cleavage with cyanogen bromide (CNBr, Figure 2a).[38] Protein expression was carried out in LB media at 18 °C for 16-24 h after induction by 0.5 mM IPTG. Due to the insoluble nature of the KSI fusion protein, protein purification was easily achieved as a series of centrifugation steps following cell lysis. This resulted in an overabundance of the desired fusion protein KSI-1.0-IDP-2Yx2A in solution when analyzed by LCMS (Figure 2b). To remove the KSI fusion tag from 1.0-IDP-2Yx2A, CNBr was used to cleave at the single Met residue between the two sequences. After cleavage, the desired 1.0-IDP-2Yx2A protein was soluble in aqueous conditions while the KSI protein remained in the insoluble fraction and could be removed via centrifugation. The level of purity typically achieved after these subsequent rounds of precipitation and centrifugation is suitable for many applications, as it still shows assembly and surfactant behavior (SI Figure 3). Importantly, the purification process uses minimal solvent and is easily performed, making it amenable to scale up with protein yields between 100-300 mg/L. For this study, where precise measurements were sought, preparatory scale HPLC was performed as additional purification step (Figure 2c). When compared with the 1.0-IDP-2Yx2A sequence produced by soluble expression, 1.0-IDP-2Yx2A produced through this insoluble scheme showed similar size, morphology, and stability properties when analyzed by DLS, AFM, and pyrene fluorescence, respectively.[36]
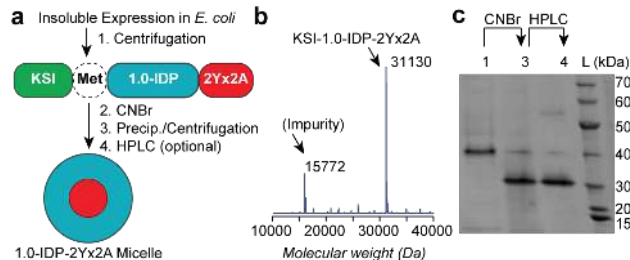
Figure 2: Insoluble expression of IDP-2Yx2A proteins. (a) The 1.0-IDP-2Yx2A is genetically encoded at the C-terminus of KSI following a single Met residue. Purification is achieved through a series of centrifugation and precipitation steps, followed by an optional HPLC purification step. (b) LCMS analysis of the KSI-1.0-IDP-2Yx2A protein recovered from inclusion bodies after centrifugal purification is shown. Expected: 31132 Da, observed: 31130 Da. (c) SDS PAGE gel analysis is shown for the purification steps outlined in (a). KSI-1.0-IDP-2Yx2A and 1.0-IDP-2Yx2A have apparent molecular weights of 40 kDa and 30 kDa, respectively.

Using this insoluble expression and purification protocol, five additional variant variants of 1.0-IDP-2Yx2A were produced. A positively charged version, (+)1.0-IDP-2Yx2A, was developed by genetically replacing all the glutamic acid residues (E) in the original sequence with lysine residues (K). This resulted in 45 total positively charged residues and an overall charge of +45 at neutral pH. Additionally, 3 variants with decreasing IDP lengths were produced, 0.75-IDP-2Yx2A, 0.5-IDP-2Yx2A and 0.25-IDP-2Yx2A, where the IDP domains are described as fractions of the full sequence in the naming scheme. Finally, a negative control variant (1.0-IDP), which contains the full-length IDP domain, but no assembly region was also expressed. This variant, which has no hydrophobic domain at its C-terminus, contains only one aromatic residue in the entire sequence for monitoring by UV-vis (Figure 3a). While the positively charged variant maintains a similar molecular weight and hydrophilic fraction (molecular weight hydrophilic domain/molecular weight total protein), the variants with decreasing IDP lengths have a corresponding decrease in the hydrophilic fraction and the overall molecular weight (Figure 3b). All variants were produced and isolated with a high degree of purity and analyzed by mass spectrometry (Figure 3c). When analyzed by DLS in 100 mM PB pH 5.7, the 1.0-IDP negative control variant showed a small distribution centered about 7.9 nm for its mean diameter, and no apparent self-assembly behavior at room temperature (25 °C) by DLS in any analyzed conditions (pH: 1-12, phosphate buffer [x]: 0-200 mM) (SI

Figure 2). When appended with the 2Yx2A domain, 1.0-IDP-2Yx2A showed a large increase in its apparent diameter and size distribution, indicating self-assembly (Figure 3d). DLS analysis was performed for all variant proteins at multiple pH values and concentrations. While changing the overall charge of the variant showed only a slight, not significant, increase in overall size and an increase in polydispersity, decreasing the IDP domain results in a respective decrease in average mean diameter (Figure 3e). Other DLS properties were also analyzed, including %Intensity spectra (SI Figure 6) and the averaged diffusion coefficients, Z-averages, and polydispersity indices (SI Figure 5).
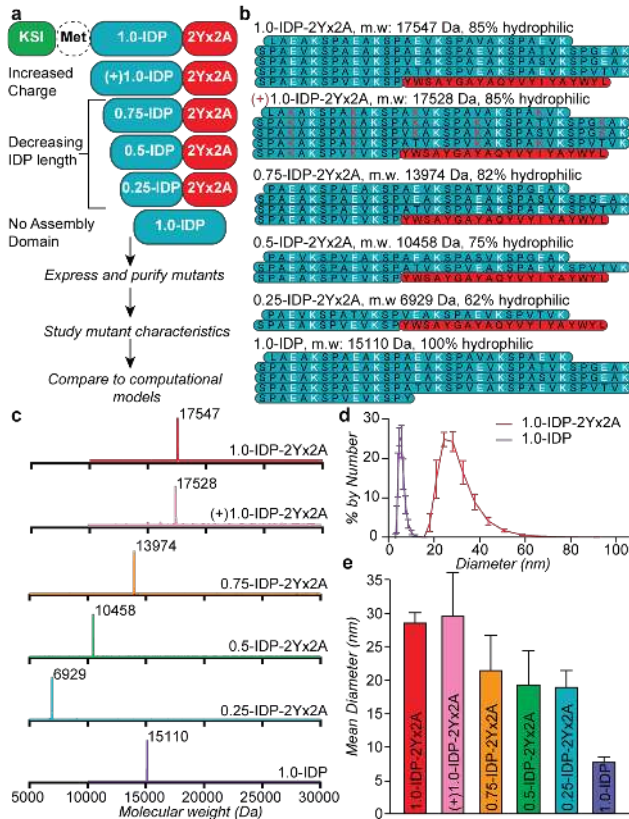
Figure 3: Generation of IDP variants with increased charges and decreasing lengths. (a) Using the insoluble expression and purification system, six protein constructs, including a non-assembling 1.0-IDP control, were expressed and purified. (b) The sequences varied in charge, length, and hydrophilic content. The parent construct, 1.0-IDP-2Yx2A, corresponds to the full length and charge-balanced 1.0-IDP domain appended with 2Yx2A. The (+)1.0-IDP-2Yx2A sequence has replaced every negatively charged E residue for a positively charged K. The 0.75-IDP-2Yx2A5, 0.5-IDP-2Yx2A, and 0.25-IDP-2Yx2A variants have decreasing fractions of the charge-balanced IDP domain and a corresponding decrease in mass fraction of the hydrophilic domain. (c) Mass spectra of the purified variants: 1.0-IDP-2Yx2A expected: 17547 Da, observed: 17547 Da; (+)1.0-IDP-2Yx2A expected: 17528 Da, observed: 17528 Da; 0.75-IDP-2Yx2A expected: 13975 Da, observed: 13974 Da; 0.5-IDP-2Yx2A expected: 10459 Da, observed: 10458 Da; 0.25-IDP-2Yx2A expected: 6929 Da, observed: 6929 Da; 1.0-IDP expected: 15110 Da, observed: 15110 Da. (d) Representative DLS spectra of 1.0-IDP-2Yx2A and the 1.0-IDP control in 100 mM PB, pH 5.7. (e) Averaged mean diameter values determined by DLS for each variant in 100 mM PB, pH 5.7: 1.0-IDP-2Yx2A: 28.9 ± 1.3 nm, (+)1.0-IDP-2Yx2A: 29.9 ± 6.2 nm, 0.75-IDP-2Yx2A: 21.7 ± 5.0 nm, 0.5-IDP-2Yx2A: 19.6 ± 4.9 nm, 0.25-IDP-2Yx2A: 19.1 ± 2.4 nm, 1.0-IDP: 7.9 ± 0.6 nm.

The decrease in size corresponding to a decrease in IDP length was also investigated by atomic force microscopy (AFM). The four IDP-2Yx2A variants with decreasing lengths and 1.0-IDP as a non-assembling control were deposited on atomically flat mica in low µM concentrations (1-5 µM) and subsequently imaged by AFM in air using soft tapping mode (TM-AFM). All the variants containing the 2Yx2A assembly domain showed the formation

of small spherical particles, while 1.0-IDP alone showed no aggregation behavior. Using a watershed segmentation method (as defined in the ISO 25178-2 standard in Mountains Software) the maximum heights (Z-max) and diameters of the particles were obtained. Hill motifs were detected after pre-processing the images using a 3x3 filter size. Further height pruning was performed by merging motifs of <5% of the maximum height. The Z-max distributions showed that with each decreasing IDP length, there was a significant decrease (p = 0.0045 - <0.0001) in particle height between each consecutive variant. This was especially significant in comparison to the 1.0-IDP control, with representative images shown in (Figure 4a). Although the Z-spatial resolution in AFM is high (0.1 nm), the measured height might not reflect the actual micelle diameter as the material is chemo-adsorbed and dried on the surface. This has resulted in a flattened structure as in other micelle systems observed by AFM.[39,40] In addition, the XY-spatial resolution depends highly on the radius of curvature of the AFM tips, which could change upon repeated scanning. Nevertheless, soft TM-AFM analysis allowed an approximate estimation of the assembly diameters. A comparison of the particle diameter distributions showed better consensus with the diameters obtained by DLS size data with only 1.0-IDP-2Yx2A being significantly larger than the other variants. Applying the same analysis parameters to the 1.0-IDP control could not be accomplished as no distinct particle features were observed (Figure 4b).This analysis both confirms the formation of micelle structures for all of the IDP-2Yx2A variants containing an assembly domain, and validates the DLS results that indicated a respective decrease in assembly size with decreasing IDP domain length.
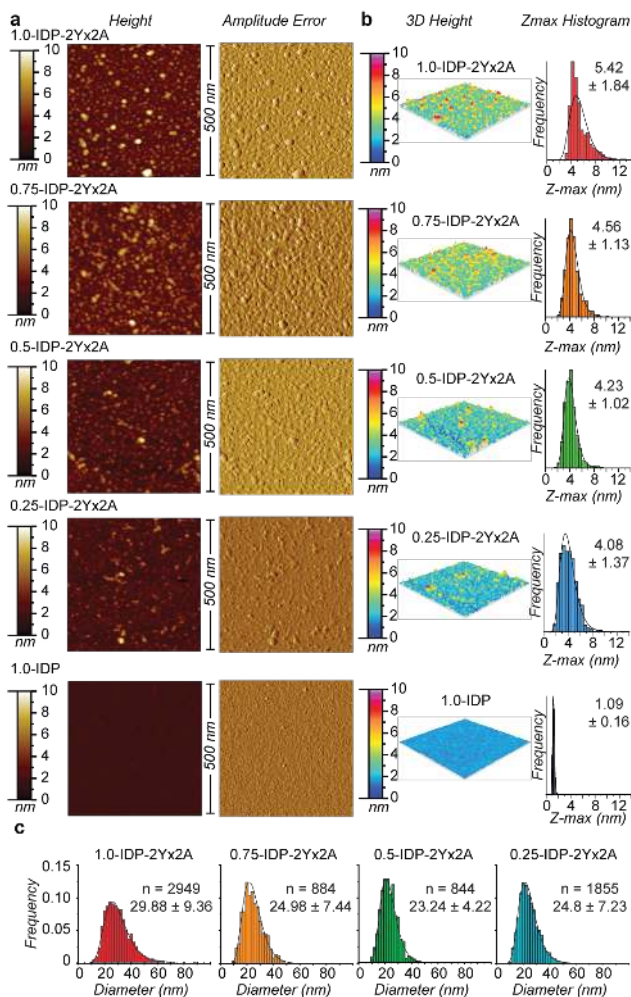
Figure 4: AFM analysis of IDP variants with decreasing lengths. (a) 2D height and amplitude error profiles showed tight packing of micelles with spherical morphologies for all variants with the 2Yx2A assembly domain. 2D amplitude error projections highlight the layering of the taller particles on top of the closely packed smaller particles. This effect is especially evident for the 1.0-IDP-2Yx2A variant. (b) Projection of height profiles in 3D along with particle analysis (watershed detection method) revealed decreasing Z-max values as IDP length decreased, $p < 0.0001$ for all variants, except for the transition between 0.5-IDP-2Yx2A to 0.25-IDP-2Yx2A where $p = 0.0045$. All variants with the 2Yx2A domain also had significantly larger z-max profiles than the IDP control ($p < 0.0001$). (c) The mean diameter in the X-Y plane for each IDP-2Yx2A variant was determined using a watershed detection method. 1.0-IDP-2Yx2A -> 0.25-IDP-2Yx2A: 29.88 ± 9.36 (N = 2939), 24.98 ± 7.44 (N = 884), 23.24 ± 4.22 (N = 844), 24.8 ± 7.23 (N = 1855). 1.0-IDP could not be evaluated for its mean diameter as no spherical particles were detected.

The most interesting and perhaps useful property of the IDP-2Yx2A variants is their ability to encapsulate hydrophobic molecules into the micelle interior at low protein concentrations. To observe the effect of IDP length and charge on these properties, the critical micelle concentrations (CMCs) of the IDP-2Yx2A length variants were determined by using a solvatochromatic pyrene assay. When encapsulated in the hydrophobic micelle interior,

pyrene exhibits a unique I3 vibronic band emission at 384 nm due to the solvatochromatic nature of the molecule. Additionally, pyrene excimer formation (ex) at approximately 450 nm is observed when a large proportion of the pyrene in solution is encapsulated in close proximity in the interior of the micelle.[41–43] While the 1.0-IDP negative control shows no characteristic excimer formation at any protein concentration analyzed (0 - 150 µM), when resuspended in the same pyrene containing buffer all of the variants with the 2Yx2A domain exhibited prominent excimer peaks at higher protein concentrations (Figure 5a). All variants with the 2Yx2A domain also showed the characteristic I3 band emergence at high concentrations and its disappearance at lower concentrations as the micelles disassembled and the hydrophobic core was disrupted. The 1.0-IDP control sample showed no I3 band emergence at concentrations up to 100 µM (SI Figure 8). CMC values for all the variants with the 2Yx2A domain were derived by plotting the ratio of the intensity of the I1 (372 nm) to the I3 (384 nm) band or the I1 to the excimer (444 nm) over a range of concentrations. This resulted in sigmoidal data plots. By fitting a logistic function to both the I1/I3 and I1/ex data sets, the CMC values could be interpreted as either the inflection points (EC_{50}) or the interceptions of the rapidly changing portion of the curve with the nearly-horizontal lower concentration regime. These values closely matched the derived EC_{20} value and are what we report.[?] In all cases the CMC values derived form the I1/ex were lower compared to the values derived from the I1/I3 band ratios. For all variants with the assembly 2Yx2A domain, the CMC values calculated using either method were in the low µM range. (SI Table 3, Table 4). Notably, when comparing the pyrene fluorescence data for 1.0-IDP-2Yx2A to those of the positively charged variant (+)1.0-IDP-2Yx2A, a significant increase in the CMC (EC_{20}, EC_{50}) value was observed when calculated by both the I1/I3 (p = 0.0029, 0.0103) or I1/ex (p = 0.00015, 0.0016, Figure 5b). While still in the µM range, the increase in CMC for (+)1.0-IDP-2Yx2A is likely due to charge repulsion on the IDP side chain.

To determine if the harbouring of hydrophobic molecules in the micelle interior disrupted micelle formation, DLS was performed on IDP-2Yx2A samples loaded with pyrene molecules.

In comparison to the unloaded protein micelles, no substantial changes to the size were observed at high protein concentrations. However, when the protein is diluted and the pyrene concentration is kept constant, an increase in micelle size is observed for all variants. Consistent with (+)1.0-IDP-2Yx2A having a larger CMC than the other variants, it also swells to form larger particles in comparison to the other variants at the same concentration (Figure 5c). This phenomenon is likely due to a swelling effect around the hydrophobic molecule core as the concentration of protein in solution decreases relative to the hydrophobic molecule.

One major application of bio-surfactants is as their use in the delivery and solubilization of hydrophobic drugs. To test if an active drug molecules could also associated with the protein micelles, 1.0-IDP-2Yx2A (5 µM) was loaded with the hydrophobic chemotherapeutic agent SN-38[44] (40 µM) and examined by AFM to determine if the size or morphology was changed by the presence of SN-38. Indeed, spherical micelles with a similar Zmax (loaded: 5.76±1.36, unloaded: 5.42 ± 1.84) and diameter (loaded: 25.88 ± 6.11, unloaded: 29.88 ± 9.36) were observed, indicating that neither size nor morphology is appreciably altered by the presence of SN-38 (Figure 5d).
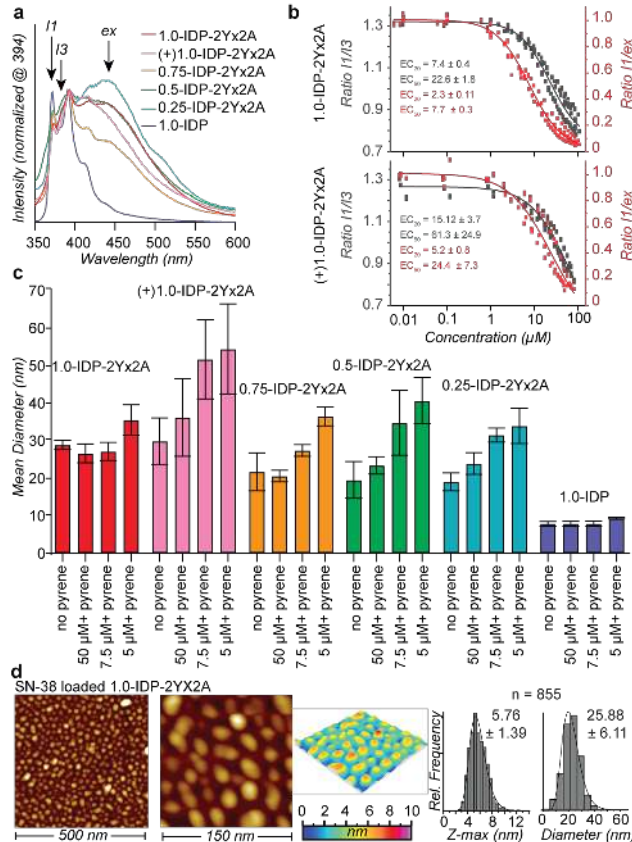
Figure 5: Encapsulation of hydrophobic molecules in the micelle core. (a) Representative fluorescent spectra of IDP-2Yx2A variants (30-50 µM) and 1.0-IDP control (100 µM) in the presence of 2 µM pyrene buffer are shown normalized for their emission at 394 nm. All variants except the 1.0-IDP control showed excimer fluorescence at 450 nm, which decreased with decreasing concentration. (b) I1/I3 and I1/ex ratios derived from the pyrene fluorescence spectra for 1.0-IDP-2Yx2A and (+)1.0-IDP-2Yx2A variants at concentrations of 0.01-100 µM. CMC values were determined by fitting a logistic function to three independently run data sets of both the I1/I3 ratio: 1.0-IDP-2Yx2A: $(EC_{20}, EC_{50}) = I1/I3$: $(7.4 \pm 0.4, 22.6 \pm 1.8$ µM), (+)1.0-IDP-2Yx2A: $(15.1 \pm 3.7, 61.3 \pm 24.9$ µM), and the I1/ex ratio: 1.0-IDP-2Yx2A: $(EC_{20}, EC_{50}) = I1/I3$: $(2.3 \pm 0.1, 7.7 \pm 0.3$ µM), (+)1.0-IDP-2Yx2A: $(5.2 \pm 0.8, 24.4 \pm 7.3$ µM). (c) DLS analysis of IDP-2Yx2A variants loaded with hydrophobic pyrene shows that size was unaffected by the presence of 2 µM pyrene at high protein concentrations. As protein concentration was decreased, an increase in size is observed for all variants. (d) AFM analysis of 1.0-IDP-2Yx2A loaded with SN38 exhibits spherical micelles with Zmax and diameter values of 5.76 ± 1.36 nm and 25.66 ± 6.11 nm, respectively.

In summary, it was found that mutations to the hydrophilic IDP domain of IDP-2Yx2A protein micelles can impact the physicochemical properties of self-assembly. By using an insoluble expression system, IDP-2Yx2A variants with high positive charge densities and variable hydrophilic domain lengths were expressed and purified with ease. In contrast to the 1.0-IDP protein alone, when the 2Yx2A hydrophobic domain was appended, self-assembly into micelle structures was observed for all variants. As the IDP domain length is decreased,

spherical micelle morphology was maintained, but the overall particle diameter decreased.The ability to harbor hydrophobic molecules in the micelle interior without disrupting morphology was further confirmed for all variants with the 2Yx2A assembly domain. Importantly, it was found that by increasing the charge of the hydrophilic IDP domain by replacing every gultamic acid residue with a lysine residue, the CMC value of the positively charged variant (+)1.0-IDP-2Yx2A was significantly higher than that of the neutral IDP sequence, indicating that this variant formed a less stable assembly likely due to charge repulsion.

This study represents a systematic approach to investigate the impact of charge and domain length on the self-assembly of protein micelles based on this IDP sequence. Overtime, data set such as these are envisioned to build a complete picture of the self-assembly behavior of these IDP-based bio-surfactants. However, further studies using this diverse insoluble expression system will aim to expand the sequence space and examine new parameters that may influence the physicochemical properties of these bio-surfactants.

Acknowledgement

Supporting Information Available

The supporting information is presented in the attached file.

References

(1) Ruzette, A. V.; Leibler, L. Block copolymers in tomorrow's plastics. *Nature Materials* **2005**, *4*, 19–31.

(2) Deschênes, L.; Bousmina, M.; Ritcey, A. M. Micellization of PEO/PS block copolymers at the air/water interface: A simple model for predicting the size and aggregation number of circular surface micelles. *Langmuir* **2008**, *24*, 3699–3708.

(3) Binder, K.; Müller, M. Monte Carlo simulation of block copolymers. *Current Opinion in Colloid and Interface Science* **2000**, *5*, 314–322.

(4) Mayes, A. M.; de la Cruz, M. O. Cylindrical versus Spherical Micelle Formation in Block Copolyrner/Homopolymer Blends. *Macromolecules* **1988**, *21*, 2543–2547.

(5) Jäger, C.; Hirsch, A.; Schade, B.; Böttcher, C.; Clark, T. Counterions Control the Self-Assembly of Structurally Persistent Micelles: Theoretical Prediction and Experimental Observation of Stabilization by Sodium Ions. *Chemistry - A European Journal* **2009**, *15*, 8586–8592.

(6) Stephenson, B. C.; Goldsipe, A.; Beers, K. J.; Blankschtein, D. Quantifying the hydrophobic effect. 1. A computer simulation-molecular- thermodynamic model for the self-assembly of hydrophobic and amphiphilic solutes in aqueous solution. *Journal of Physical Chemistry B* **2007**, *111*, 1025–1044.

(7) Vishnyakov, A.; Lee, M. T.; Neimark, A. V. Prediction of the critical micelle concentration of nonionic surfactants by dissipative particle dynamics simulations. *Journal of Physical Chemistry Letters* **2013**, *4*, 797–802.

(8) Nagarajan, R.; Ganesh, K. Block copolymer self-assembly in selective solvents: Spherical micelles with segregated cores. *The Journal of Chemical Physics* **1989**, *90*, 5843–5856.

(9) Ianiro, A.; Patterson, J.; González García, Á.; van Rijt, M. M. J.; Hendrix, M. M. R. M.; Sommerdijk, N. A. J. M.; Voets, I. K.; Esteves, A. C. C.; Tuinier, R. A roadmap for poly(ethylene oxide)- block -poly-ϵ-caprolactone self-assembly in water: Prediction, synthesis, and characterization. *Journal of Polymer Science Part B: Polymer Physics* **2018**, *56*, 330–339.

(10) Israelachvili, J. N.; Mitchell, D. J.; Ninham, B. W. Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics* **1976**, *72*, 1525–1568.

(11) Lechuga, M.; Fernández-Serrano, M.; Jurado, E.; Núñez-Olea, J.; Ríos, F. Acute toxicity of anionic and non-ionic surfactants to aquatic organisms. *Ecotoxicology and Environmental Safety* **2016**, *125*, 1–8.

(12) Cowan-Ellsberry, C.; Belanger, S.; Dorn, P.; Dyer, S.; McAvoy, D.; Sanderson, H.; Versteeg, D.; Ferrer, D.; Stanton, K. Environmental Safety of the Use of Major Surfactant Classes in North America. *Critical Reviews in Environmental Science and Technology* **2014**, *44*, 1893–1993.

(13) Fischer, D.; Li, Y.; Ahlemeyer, B.; Krieglstein, J.; Kissel, T. In vitro cytotoxicity testing of polycations: influence of polymer structure on cell viability and hemolysis. *Biomaterials* **2003**, *24*, 1121 – 1131.

(14) Gozlan, C.; Deruer, E.; Duclos, M. C.; Molinier, V.; Aubry, J. M.; Redl, A.; Duguet, N.; Lemaire, M. Preparation of amphiphilic sorbitan monoethers through hydrogenolysis of sorbitan acetals and evaluation as bio-based surfactants. *Green Chemistry* **2016**, *18*, 1994–2004.

(15) Zhang, Q. Q.; Cai, B. X.; Gang, H. Z.; Yang, S. Z.; Mu, B. Z. A family of novel bio-based zwitterionic surfactants derived from oleic acid. *RSC Advances* **2014**, *4*, 38393–38396.

(16) Foley, P.; Kermanshahi Pour, A.; Beach, E. S.; Zimmerman, J. B. Derivation and synthesis of renewable surfactants. *Chemical Society Reviews* **2012**, *41*, 1499–1518.

(17) Douglas G. Hayes, R. D. A., Daniel K. Solaiman *Biobased Surfactants: Synthesis, Properties, and Applications*.

(18) Keeley, F. W.; Bellingham, C. M.; Woodhouse, K. A. Elastin as a self-organizing biomaterial: use of recombinantly expressed human elastin polypeptides as a model for investigations of structure and self-assembly of elastin. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **2002**, *357*, 185–9.

(19) Wright, E. R.; Conticello, V. P. Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. *Advanced Drug Delivery Reviews* **2002**, *54*, 1057–1073.

(20) Dreher, M. R.; Simnick, A. J.; Fischer, K.; Smith, R. J.; Patel, A.; Schmidt, M.; Chilkoti, A. Temperature Triggered Self-Assembly of Polypeptides into Multivalent Spherical Micelles. *Journal of the American Chemical Society* **2008**, *130*, 687–694, PMID: 18085778.

(21) Ribeiro, A.; Arias, F. J.; Reguera, J.; Alonso, M.; Rodríguez-Cabello, J. C. Influence of the amino-acid sequence on the inverse temperature transition of elastin-like polymers. *Biophysical Journal* **2009**, *97*, 312–320.

(22) Li, N. K.; Roberts, S.; Quiroz, F. G.; Chilkoti, A.; Yingling, Y. G. Sequence Directionality Dramatically Affects LCST Behavior of Elastin-Like Polypeptides. *Biomacromolecules* **2018**, *19*, 2496–2505.

(23) MacEwan, S. R.; Chilkoti, A. Applications of elastin-like polypeptides in drug delivery. *Journal of Controlled Release* **2014**, *190*, 314–330.

(24) Dzuricky, M.; Roberts, S.; Chilkoti, A. Convergence of Artificial Protein Polymers and Intrinsically Disordered Proteins. *Biochemistry* **2018**, *57*, 2405–2414.

(25) Hassouneh, W.; Zhulina, E. B.; Chilkoti, A.; Rubinstein, M. Elastin-like Polypeptide Diblock Copolymers Self-Assemble into Weak Micelles. *Macromolecules* **2015**, *48*, 4183–4195.

(26) Weitzhandler, I.; Dzuricky, M.; Hoffmann, I.; Garcia Quiroz, F.; Gradzielski, M.; Chilkoti, A. Micellar Self-Assembly of Recombinant Resilin-/Elastin-Like Block Copolypeptides. *Biomacromolecules* **2017**, *18*, 2419–2426.

(27) Injectable tissue integrating networks from recombinant polypeptides with tunable order. *Nature Materials* **2018**, *17*, 1154–1163.

(28) Holehouse, A. S.; Das, R. K.; Ahad, J. N.; Richardson, M. O.; Pappu, R. V. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophysical Journal* **2017**, *112*, 16–21.

(29) Peng, Z.; Yan, J.; Fan, X.; Mizianty, M. J.; Xue, B.; Wang, K.; Hu, G.; Uversky, V. N.; Kurgan, L. Exceptionally abundant exceptions: Comprehensive characterization of intrinsic disorder in all domains of life. *Cellular and Molecular Life Sciences* **2014**, *72*, 137–151.

(30) Adiga, S. P.; Brenner, D. W. Molecular Basis for Neurofilament Heavy Chain Side Arm Structure Modulation by Phosphorylation. *The Journal of Physical Chemistry C* **2010**, *114*, 5410–5416.

(31) Ackerley, S.; Thornhill, P.; Grierson, A. J.; Brownlees, J.; Anderton, B. H.; Leigh, P. N.; Shaw, C. E.; Miller, C. C. J. Neurofilament heavy chain side arm phosphorylation regulates axonal transport of neurofilaments. *The Journal of cell biology* **2003**, *161*, 489–95.

(32) Bhagawati, M.; Rubashkin, M. G.; Lee, J. P.; Ananthanarayanan, B.; Weaver, V. M.; Kumar, S. Site-Specific Modulation of Charge Controls the Structure and Stimulus

Responsiveness of Intrinsically Disordered Peptide Brushes. *Langmuir* **2016**, *32*, 5990–5996.

(33) Chang, R.; Kwak, Y.; Gebremichael, Y. Structural Properties of Neurofilament Sidearms: Sequence-Based Modeling of Neurofilament Architecture. *Journal of Molecular Biology* **2009**, *391*, 648–660.

(34) Lees, J. F.; Shneidman, P. S.; Skuntz, S. F.; Carden, M. J.; Lazzarini, R. A. The structure and organization of the human heavy neurofilament subunit (NF-H) and the gene encoding it. *The EMBO journal* **1988**, *7*, 1947–55.

(35) Hocine, S.; Li, M.-H. Thermoresponsive self-assembled polymer colloids in water. *Soft Matter* **2013**, *9*, 5839–5861.

(36) Klass, S. H.; Smith, M. J.; Fiala, T. A.; Lee, J. P.; Omole, A. O.; Han, B.-G.; Downing, K. H.; Kumar, S.; Francis, M. B. Self-Assembling Micelles Based on an Intrinsically Disordered Protein Domain. *Journal of the American Chemical Society* **2019**, *141*, 4291–4299.

(37) Lawrence, M. S.; Phillips, K. J.; Liu, D. R. Supercharging Proteins Can Impart Unusual Resilience. *Journal of the American Chemical Society* **2007**, *129*, 10110–10112, PMID: 17665911.

(38) Kuliopulos, A.; Walsh, C. T. Production, Purification, and Cleavage of Tandem Repeats of Recombinant Peptides. *Journal of the American Chemical Society* **1994**, *116*, 4599–4607.

(39) Ouanezar, M.; Guyomarc'h, F.; Bouchoux, A. AFM imaging of milk casein micelles: Evidence for structural rearrangement upon acidification. *Langmuir* **2012**, *28*, 4915–4919.

(40) Di Maria, F.; Blasi, L.; Quarta, A.; Bergamini, G.; Barbarella, G.; Giorgini, L.; Benaglia, M. New biocompatible polymeric micelles designed for efficient intracellular uptake and delivery. *Journal of Materials Chemistry B* **2015**, *3*, 8963–8972.

(41) Guler, M. O.; Claussen, R. C.; Stupp, S. I. Encapsulation of pyrene within self-assembled peptide amphiphile nanofibers. *Journal of Materials Chemistry* **2005**, *15*, 4507.

(42) Piñeiro, L.; Novo, M.; Al-Soufi, W. Fluorescence emission of pyrene in surfactant solutions. *Advances in Colloid and Interface Science* **2015**, *215*, 1–12.

(43) Basu Ray, G.; Chakraborty, I.; Moulik, S. P. Pyrene absorption can be a convenient method for probing critical micellar concentration (cmc) and indexing micellar polarity. *Journal of Colloid and Interface Science* **2006**, *294*, 248–254.

(44) Nguyen, F.; Alferiev, I.; Guan, P.; Guerrero, D. T.; Kolla, V.; Moorthy, G. S.; Chorny, M.; Brodeur, G. M. Enhanced Intratumoral Delivery of SN38 as a Tocopherol Oxyacetate Prodrug Using Nanoparticles in a Neuroblastoma Xenograft Model. *Clinical Cancer Research* **2018**, *24*, 2585–2593.

Graphical TOC Entry

Some journals require a graphical entry for the Table of Contents. This should be laid out "print ready" so that the sizing of the text is correct.

Inside the `tocentry` environment, the font used is Helvetica 8 pt, as required by *Journal of the American Chemical Society*.

The surrounding frame is 9 cm by 3.5 cm, which is the maximum permitted for *Journal of the American Chemical Society* graphical table of content entries. The box will not resize if the content is too big: instead it will overflow the edge of the box.

This box and the associated title will always be printed on a separate page at the end of the document.