Application of Repetitive Protein Polypeptides with an Upper Critical Solution Temperature at Various Length Scales

by

Michael James Dzuricky

Department of Biomedical Engineering
Duke University

Date:_______________________
Approved:

__________________________________
Ashutosh Chilkoti, Advisor

__________________________________
Brenton Hoffman

__________________________________
Joel Collier

__________________________________
Stefan Zauscher

__________________________________
Michael Rubinstein

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biomedical Engineering in the Graduate School of Duke University

2019
Application of Repetitive Protein Polypeptides with an Upper Critical Solution Temperature at Various Length Scales

by

Michael James Dzuricky

Department of Biomedical Engineering
Duke University

Date: _______________________

Approved:

___________________________
Ashutosh Chilkoti, Advisor

___________________________
Brenton Hoffman

___________________________
Joel Collier

___________________________
Stefan Zauscher

___________________________
Michael Rubinstein

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biomedical Engineering in the Graduate School of Duke University

2019
Abstract

Phase separation of macromolecules is a critical phenomenon for the human condition. This phenomenon has also been exploited for biotechnological development to improve human morbidity and mortality. However, there is still much more to learn regarding how this behavior is encoded within a protein sequence. Thus, this thesis seeks to 1) further explore the sequence space to understand how phase separation is encoded, with an emphasis on polypeptides with upper critical solution temperature (UCST) transitions and 2) use this phase separation to control availability of macromolecules at various length scales.

Using traditional molecular biology techniques, we will recombinantly express and purify a large number of polypeptides with variable sequence composition and sequence architecture. Then, using traditional polymer science and material science techniques combined with microscopic techniques that span the macro-scale and nano-scale, we will characterize their phase separation behavior and the interaction of these materials with biological systems.

We developed a practical mutation strategy that allows for complete control of the UCST binodal line in physiologic conditions that is useful for de novo design of artificial IDPs with UCST phase behavior. We evaluated the interaction of these polypeptides and their phase separation in the presence of bacterial, eukaryotic cells and in mice demonstrating how this binodal line fused to biological active partners can control biologic functions.

In bacteria, we made artificial phase separated puncta, akin to naturally occurring phase separated droplets, that have non-canonical function, demonstrating how primary
features of the polypeptide chain affect enzymatic function. We created block co-
polypeptides comprised of UCST and LCST protein sequences that exhibit remarkably
tunable and robust nanoscale self-assembly into spherical micelles, worm-like micelles
and vesicular structures capable of displaying large targeting domains on their surface.
In the presence of eukaryotic cells, these nanomaterials can dramatically increase
polypeptide uptake, increasing the avidity of the targeting molecule by over 1000-fold.
Finally, we demonstrated that phase separated polypeptides can sequester an active
peptide GLP-1 from systemic circulation, controlling the peptide’s bioactivity through
control of the phase diagram. Taken together, we demonstrate the universal power of
the phase diagram, across many length scales, where the transducing agent for
controlling biological activity is an engineered, repetitive polypeptide sequence.
Dedication

I would like to dedicate this work to my parents, Kathie and Michael Dzuricky for always nurturing my many curiosities. This work is an extrapolation of their faith in me to explore the world around us.
## Contents

Abstract .......................................................................................................................... iv  
List of Tables ................................................................................................................. xii  
List of Figures ............................................................................................................... xiii  
Acknowledgements ....................................................................................................... xx  
1. Introduction .......................................................................................................... 1  
   1.1 Historical context of polypeptides and phase separation ......................... 1  
   1.1.1 Order and chaos ......................................................................................... 1  
   1.1.2 Basics of phase separation ......................................................................... 2  
   1.1.3 A brief history of phase separation, polymer science and membrane-less  
        organelles ........................................................................................................... 5  
   1.2 Intrinsically disordered proteins ................................................................. 8  
   1.2.1 Intrinsically disordered proteins in nature .................................................. 8  
   1.2.2 Engineered IDPs ........................................................................................ 11  
   1.2.2.1 Elastin like polypeptides (ELPs) ............................................................ 11  
   1.2.2.2 Other notable proteins with repetitive domains and phase behavior ....... 14  
   1.3 Motivation ..................................................................................................... 17  
   1.4 Specific aims .................................................................................................. 19  
2. An engineered library of intrinsically disordered proteins with variable UCST behavior. ................................................................. 23  
   2.1 Introduction .................................................................................................... 23  
   2.1.2 Approach .................................................................................................. 26  
   2.2 Methods......................................................................................................... 30  
   2.2.1 Gene synthesis .......................................................................................... 30
2.2.2 Protein purification ........................................................................................31
2.2.3 Characterization of phase separation ............................................................33
  2.2.3.1 Temperature dependent UV-vis spectrophotometry ...............................33
  2.2.3.2 Temperature depending dynamic light scattering .................................33
  2.2.3.3 Temperature gradient microfluidics .........................................................34
  2.2.3.4 Fitting of phase diagram binodal ............................................................35
  2.2.3.5 Creation of water in oil droplets with chip microfluidics ...........................37
  2.2.3.6 Temperature controlled fluorescent microscopy of protocell droplets .......37
  2.2.3.7 Differential scanning calorimetry ............................................................38
2.3 Results and Discussion ......................................................................................38
  2.3.1 Repetitive IDPs exhibit robust and reversible UCST phase behavior in an aqueous environment ............................................................................................ 38
  2.3.2 Arginine composition, aromatic to aliphatic ratio, charge balance and molecular weight define UCST cloud point ..............................................................46
  2.3.3 RIDPs create dense phase separated droplets at saturation concentrations mediated by amino acid composition .................................................................63
2.4 Conclusions .......................................................................................................74

3. Designer membrane-less organelles using engineered intrinsically disordered proteins with non-canonical functionality. ....................................................................................76

3.1 Introduction ........................................................................................................76
  3.1.1 Literature review ...........................................................................................76
  3.1.2 Approach ......................................................................................................77
3.2 Methods .............................................................................................................78
  3.2.1 Gene synthesis .............................................................................................78
  3.2.2 Protein purification .................................................................................... 79
  3.2.3 Characterization of phase separation ............................................................80
3.2.3.1 Temperature dependent UV-vis spectrophotometry ........................................80
3.2.3.2 Whole cell fluorescent intensity measurements ...........................................81
3.2.3.3 Temperature controlled fluorescent microscopy of protocell droplets of E. Coli. ........................................................................................................................................81
3.2.3.4 Transient Transfection of [S]-20-sfGFP in HEK293 Cells ................................82
3.2.3.5 Confocal imaging of RIDP-sfGFP fusions for puncta formation and colocalization .............................................................................................................................82
3.2.3.6 Spinning disc confocal imaging of Lac-Z alpha-peptide-RIDP gene fusions for localization and quantification of enzymatic activity .........................................84
3.2.3.7 Fluorescent spectroscopy for determining $K_m$, $V_{max}$ and $k_{cat}$ ................84
3.2.3.8 Image quantification and statistical analysis ....................................................85

3.3 Results and Discussion .........................................................................................88
3.3.1 RIDPs have controlled $C_{sat}$ in eukaryotic and prokaryotic cell lines ........88
3.3.2 RIDPs exhibit reversible UCST droplet formation in E. Coli .............................96
3.3.3 Alpha-peptide-RIDP fusion proteins form functional intracellular droplets with enzymatic efficiency linked to RIDP molecular weight .............................................105

3.4 Conclusions .......................................................................................................113

4. Creation of UCST-LCST and UCST-UCST micelles capable of multivalent display of functional moieties .................................................................................................117
4.1 Introduction ......................................................................................................117
4.1.1 Literature review ........................................................................................120
4.1.2 Approach ...................................................................................................124
4.2 Methods ..........................................................................................................126
4.2.1 Gene synthesis ..........................................................................................126
4.2.2 Protein purification ....................................................................................126
4.2.3 Characterization of phase separation .........................................................127
List of Tables

Table 1: Theoretic examples of contributions to enthalpic and entropic terms in Gibbs free energy balance for temperature responsive polypeptides with UCST phase behavior ..........................................................23

Table 2: Repeat sequences of other resilin inspired polypeptides in the literature ........26

Table 3: Number of phase separated domains per E. Coli as a function of induction time (n = 3 images) ...............................................................................................................92

Table 4: Michaelis-Menten Enzyme Kinetics Parameters (error is standard error of the mean, n = 3) .................................................................................................................112

Table 5: Light scattering data of RLPXX-ELP80 co-polypeptides..............................134

Table 6: Micelle core and inter-core spacing of RLPXX-ELP80 block co-polypeptides (mean +/- standard deviation) ..........................................................................................137

Table 7: Light scattering data of RLPXX-ELPYY co-polypeptides..............................138

Table 8: Micelle core and inter-core spacing of RLPXX-ELPYY block co-polypeptides (mean +/- standard deviation) ..........................................................................................140

Table 9: Light scattering data of RLPXX-ELPS80 and RLPXX-ELPV80 co-polypeptides ..........................................................................................................................141

Table 10: Micelle core and inter-core spacing of RLPXX-ELPSYY and RLPXX-ELPVYY block co-polypeptides (mean +/- standard deviation) ......................................................................143

Table 11: Light scattering data of Fn3 block co-polypeptides ................................155

Table 12: Cryo-TEM core diameter and spacing of RLPXX-ELP80-Fn3s ..................159

Table 13: Light scattering characterization of [S]-XX-[QHN]-40 & [S]-XX-[NHQ]-40 co-polypeptides ................................................................................................................173
List of Figures

Figure 1: Example derivation of Upper Critical Solution Temperature (UCST) phase diagram from free energy landscape as a function of chi (χ) and volume fraction (ϕ). ........................................... 5

Figure 2: Significant events in the history of protein polymers and membrane less organelles ...................................................................................................................... 7

Figure 3: One of the first schematic drawing of a polypeptide backbone by Richardson. ................................................................. 9

Figure 4: Plot of glycine and proline chain content of various engineered IDPs against naturally occurring IDP sequences. ............................................................................. 16

Figure 5: Difference in number of journal article publications that mention LCST or UCST phase behavior in the title or abstract according to PubMed as of July 12, 2019. .......... 18

Figure 6: Previous work in the Chilkoti lab with UCST polypeptides. ................................................................................................. 25

Figure 7: Proteomic analysis of IDPs that form biomolecular condensates in various biological systems ................................................................................................................. 27

Figure 8: Synthetic polymers with UCST phase behavior. ......................................................................................................................... 28

Figure 9: Histogram plotting the length of the disordered regions analyzed in Figure 7. Bars indicate 25th-75th percentiles and whiskers indicate 10th-90th percentiles. ................. 29

Figure 10: UCST polypeptides form dense, exclusionary phases even in complex medium. ................................................................................................................................. 39

Figure 11: Example SDS-PAGE gels of purified proteins used in this study. ................................................................. 40

Figure 12: Visualization of UCST phase separation of [Q]-20 in water-in-oil droplets with fluorescent microscopy ........................................................................................................... 41

Figure 13: Wide-field fluorescent microscope images of fluorescently labeled [Q]-20 inside water-in-oil compartments related to Figure 12 ........................................................................... 43

Figure 14: Dynamic light scattering data of [Q]-20 demonstrating the change in hydrodynamic radius upon cooling. ................................................................................................. 44

Figure 15: UCST cloud points are affected by polypeptide volume fraction in solution. ................................................................................................................................. 45

Figure 16: A. Schematic describing the methodology for blending polypeptide A & B. ................................................................. 47
Figure 17: Substitution of aromatic Tyr residues with aliphatic Val amino acids dramatically reduces UCST cloud point. ..........................................................48

Figure 18: Substitution of Arg residues with Lys amino acids dramatically reduces UCST cloud point. ....................................................................................49

Figure 19: Phase behavior of polypeptide blends containing three unique repeating motifs...........................................................................................................50

Figure 20: Analysis of secondary structure with circular dichroism (CD) spectroscopy related. ..............................................................................................51

Figure 21: Effect of positive and negative charge bias on UCST cloud point. ..........52

Figure 22: The overall molecular weight of the polypeptide affects the UCST cloud point. ........................................................................................................53

Figure 23: Primary amino acid sequence and syntax influence observed phase behavior in Tyr containing RIDPs.................................................................54

Figure 24: Inconsistent ordering in the hydrophobicity of non-charged polar residues in Tyr containing RIDPs.................................................................55

Figure 25: Simulations of short motifs of [QN] and [NQ]. ............................................56

Figure 26: Effect of urea concentration on the UCST transition temperature of RIDPs of various molecular weight. .................................................................59

Figure 27: Salt concentration affects but does not eliminate UCST phase behavior .....60

Figure 28: Decreasing pH of solution results in dramatic increase in Tt of [H]-60........62

Figure 29: Partial binodal phase boundaries of agnostically non-repetitive but compositionally identical versions of [S]-20. .......................................................63

Figure 30: Collection of phase diagrams using a temperature gradient device. ...........64

Figure 31: Binodal phase lines of [S]-20 and [Q]-20 using multiple data points from temperature gradient device. .................................................................65

Figure 32: Kohlrausch-Williams-Watts (KWW) stretched exponential decay of a temperature gradient line scan to determine rate constants of phase separation....67

Figure 33: Low energy of activation (Ea) suggests growth via coalescence rather than Ostwald ripening. Data collected with [Q]-20 @ 25 mg ml⁻¹. ...............................68
Figure 34: Normalized hydropathy of RIDPs (R² = 0.23).

Figure 35: Lowering UCST cloud point of [N]-40 upon repeated cycles of interrogation despite an apparent increase in the molecular weight by SDS-PAGE.

Figure 36: Idealized chain conformation of [N] and [NQ] repeating unit.

Figure 37: Quantification of dextran uptake during phase separation of RIDPs.

Figure 38: Schematic describing the use of main effectors of UCST phase behavior to control intracellular droplet formation.

Figure 39: The chemical composition is capable of affecting the saturation concentration by 2 orders of magnitude at constant molecular weight.

Figure 40: At constant chemical composition it is possible to affect Cₛₐₜ by over four orders of magnitude by changing molecular weight of the chain.

Figure 41: Comparison of partial binodal phase diagrams of RIDP and RIDP-sfGFP fusions.

Figure 42: Confocal microscopy images of HEK293 cells with transiently transfected [S]-20-sfGFP.

Figure 43: Measurement of total cellular fluorescence at different times post induction.

Figure 44: Size of intracellular droplets grows with induction time.

Figure 45: Measurement of the cellular fluorescence at different locations within the cell.

Figure 46: Confocal images of UCST polypeptides as a function of induction time and molecular weight.

Figure 47: Reducing the aromatic content increases the saturation concentration in a dose dependent manner. All scale bars are 5 µm.

Figure 48: Intracellular droplets comprised of [S]-20-sfGFP can be formed and dissolved reversibly via alternating cooling and heating cycles.

Figure 49: Normalized UCST cloud point does not change significantly over four heating and cooling cycles.

Figure 50: [S]-20-sfGFP exhibits phase separation memory upon multiple cycles of heating and cooling.
Figure 70: Cryo-TEM images of RLP40-ELP80 at increasing volume fractions does not appear to affect observed assembly state.

Figure 71: Cryo-TEM images of RLPXX-ELPYY block co-polypeptides.

Figure 72: Cryo-TEM images of RLPXX-ELPSYY and RLPXX-ELPVYY block co-polypeptides.

Figure 73: Cryo-TEM images of RLP-ELP block co-polypeptides with varying core hydrophobicity.

Figure 74: Cryo-TEM images of (GRGDSP[Y:V]S)80-ELP80 block co-polypeptide in 140 mM PBS and distilled H2O.

Figure 75: Fluorescent measurements of pyrene peaks (I1 and I3) at various concentrations of RLPXX-ELP80 block co-polypeptides in 140 mM PBS at 20°C.

Figure 76: Cryo-TEM images of RLP40-ELP80 block co-polypeptides where the core sequence contains varying amount of Ser and Glu.

Figure 77: Schematic of paclitaxel loading of RLP40-ELP80 micelles and analytic procedure.

Figure 78: The relative molar ratio of paclitaxel (PTX) to RLP-ELP as determined by analytical high-performance liquid chromatography.

Figure 79: SDS-PAGE of RLPXX-ELP80-Fn3.

Figure 80: Thermal stability of RLP-ELP-Fn3 micelles.

Figure 81: Thermal stability of block co-polypeptide micelles.

Figure 82: Static and dynamic light scattering raw data for RLP-ELP block co-polypeptides.

Figure 83: Cryo-TEM micrographs of RLPXX−ELP80 and RLPXX−ELP80-Fn3.

Figure 84: Histogram of Observed Aspect Ratios of RLPXX-ELPYY-Fn3s.

Figure 85: Shape dependent avidity of RLPXX-ELP80-Fn3.

Figure 86: Intracellular uptake of undecorated block co-polypeptides.

Figure 87: Representative images of cellular uptake of RLPXX-ELP block co-polypeptides labeled with Alexa488 fluorophore overlaid with DIC images.
Figure 88: Cellular uptake of RLPXX-ELPYY-Fn3 Polypeptides in αvβ3 Negative K562 Cell Line.

Figure 89: Flow cytometry data of naïve cells, LM609 antibody, RLP40-ELP80-Fn3 - spherical micelles and RLP80-ELP80-Fn3 - worm-like micelles.

Figure 90: Quantification of cellular uptake by flow cytometry. *** = p < 0.001. Box indicates 25th and 75th percentile and bars indicate 10th and 90th percentile.

Figure 91: Cellular uptake of RLPXX-ELPYY-Fn3 polypeptides with variable aspect ratio in αvβ3 transfected cell line.

Figure 92: Cryo-TEM Characterization of “shape control” RLPXX-ELPYY-Fn3s.

Figure 93: Cellular uptake of block co-polypeptides over time.

Figure 94: Quantification of cellular uptake with image analysis.

Figure 95: Cryo-tem images of RLP-RLP block co-polypeptides.

Figure 96: UCST phase behavior of [S]-40-[QHN]-40 and [S]-80-[QHN]-40 block co-polypeptides as determined by UV-Vis spectrophotometry.

Figure 97: Critical micelle concentrations (CMCs) of [S]-40-[QHN]-40 and [S]-80-[QHN]-40.

Figure 98: UV-Vis spectrophotometry and dynamic light scattering of [S]-40-[V]-40 and [S]-40-[Y:3V]-40 block co-polypeptides.

Figure 99: Cryo-TEM images of [S]-40-[V]-20, [S]-40-[V]-40, [S]-40-[V]-60. Data collected at 15°C in 140 mM PBS. Scale bar = 500 nm.

Figure 100: Cryo-TEM images of [S]-40-[A]-40, [S]-40-[V]-40, [S]-40-[I]-40. Data collected at 15°C in 140 mM PBS. Scale bar = 500 nm.

Figure 101: Injection strategies for forming subcutaneous depots in vivo.

Figure 102: Fluorescence molecular tomography of (GRGDSYPY)40 labeled with a near infrared fluorescent dye after injection in the presence of 2M urea + PBS.

Figure 103: Fluorescence molecular tomography of (GRGDSYPY)40 labeled with a near infrared fluorescent dye after injection in dehydrated state.

Figure 104: Binodal phase boundary of GLP-1-RIDPs of lower molecular weight.
Figure 105: Blood glucose of GLP-1-RIDP fusion proteins of ~20 kDa in size and variable C_{sat}..............................................................199

Figure 106: Body weight change of mice with sub-cutaneous GLP-1-RIDP depots of ~20 kDa in size and variable C_{sat}..............................................................200

Figure 107: Binodal phase boundary of GLP-1-RIDPs of higher molecular weight........201

Figure 108: Blood glucose of GLP-1-RIDP fusion proteins of ~35 kDa in size and variable C_{sat}..............................................................202

Figure 109: Blood glucose of GLP-1-RIDP fusion proteins of different molecular weights but similar C_{sat}..............................................................204

Figure 110: Body weight change of mice with sub-cutaneous GLP-1-RIDP depots of ~35 kDa in size and variable C_{sat}..............................................................205
Acknowledgements

First, I must thank my advisor, Ashutosh Chilkoti for the opportunity to work in his lab for the past six years. He consistently found the balance between his high expectations and my developmental progress that consistently challenged me without the risk of paralysis. Our relationship is responsible for my feelings of unwavering confidence as a scientist and the personal pride I feel for this thesis – feelings that will last for my entire lifetime. For this, I will always be grateful.

Throughout my work I have had the pleasure of working with several scientists directly, helping to craft their own scientific projects. Davoud Mozhdehi, Samagya Banskota, Simone Costa, Joseph Simon, Isaac Weitzhandler and Sinan Xiong challenged my technical skill set and provided much needed diversity in intellectual thought that either helped directly to the chain of ideas that became chapters in this thesis or a much-needed distraction that allowed me to engage with a new challenge. I cherish our time working together and the relationships that we cultivated as a result.

There are many others who assisted intellectually to the ideas in these projects. This thesis would not have been possible without the groundbreaking work of Felipe Garcia Quiroz, Jon McDaniel, Sarah MacEwan and Kelli Luginbuhl. These individual’s thesis set the precedent for scientific excellence of my work and laid the groundwork for my experiments. I would particularly like to thank Felipe who also was responsible training me in the essential technical skills and setting me on a path of protein polymer curiosity that has not subsided. I would also like to thank Jeff Schaal, Nicholas Tang, Ryan Hill and Wafa Hassouneh who were critical resources to talk through the various technical challenges I encountered throughout my thesis.
I would like to specially thank Stefan Roberts. We never collaborated directly but you were always the first person to hear of my successes and failures. Whether it was a light-hearted chat over a “heart-attack” bowl or a distractingly intense session of posterball, you made sure that lab was never a dull place.

Finally, I would like to thank Jane Cooper. No one has heard more about my experience in graduate school nor supported me more. My best friend and partner for over five years, I don’t think I can adequately describe in words the contribution you have made to my life, but I appreciate you more than any accomplishment or experience I have had to date. I am eagerly looking forward to our next adventure together.
1. Introduction

1.1 Historical context of polypeptides and phase separation

1.1.1 Order and chaos

This thesis rests at the intersection of multiple fields. Thus, it is important to understand the context of how each of these fields arrived at this point to establish the central line of inquiry of this thesis – how do we control phase separation to suit the human condition? For these big questions, ones that have fascinated many generations of scientists, I think it is important to understand the why. What is so compelling about a simple observation that hundreds of scientists have dedicated their lives to learning more about it? In doing so, I can contextualize in history what I sought to accomplish and what I sought to contribute to the scientific community.

Philosophically, observations of physical systems can be defined on axes of order and chaos that characterize the presence (order) or absence (chaos) of symmetry. These “forces” are conceptually in balance with one another. Various phenomenon throughout our universe and can be classified as ordered and/or chaotic, from the colliding of galaxies to the organization of electrons around a nucleus. Given human's cognitive disposition for pattern recognition it is perhaps inevitable that we have been fascinated by the balance of these two entities throughout history. The oldest fascination is perhaps the inner workings of the human body, an archetypical mystery of these forces of chaos and order across various length scales. We are hierarchical creatures, formed by millions of atoms that organize themselves into the building blocks
of life – hydrocarbons – that form our fundamental unit - the cell - which further is
organized into compartments that function cooperatively with other compartments to
interact with the surrounding world.

I think my fascination with phase behavior is a derivative of this psychological
framework. Phase separating polypeptides exhibit characteristics of a chaotic and
ordered system, one that is a function of a variable within human control – temperature.
I sought to understand this behavior deeper and develop practical tools for harnessing
this natural phenomenon. Because phase separation is unaffected by the size of the
containing vessel, I also saw an opportunity to interface with biological systems at
various length scales, essentially using thermodynamic principles to transduce
information on length scales that are inaccessible by direct physical manipulation.

1.1.2 Basics of phase separation

Phase separation is a term used to describe the creation of two or more distinct
phases from a single homogeneous phase. The most abundant and obvious version of
phase separation that we encounter every day is the equilibrium of the various phases of
water, from the solid phase in our iced latte to the individual water vapor molecules in
our atmosphere. This equilibrium is often described by a phase diagram, which
describes the phase of the material within an area on a two-dimension variable axis. For
water, the most common description is the state as a function of pressure and
temperature.
Another form of phase separation occurs in mixtures where two phases are formed spontaneously where both phases display liquid like behavior (where liquid like is described as an incompressible fluid that conforms to the shape of the container). A common example you may have encountered is the separation of oil and vinegar salad dressing. This state exists because of the immiscibility of water and oil due to incompatibilities of the individual molecules that define oil and water. For this type of phase separation, the most popular description of the states of the solution is also a phase diagram but typically on an axis that relates volumetric fraction of the two species in a mixture to a parameter referred to as the Flory interaction parameter, χ (chi), which is a dimensionless parameter that describes the differences in pairwise interactions between species in a mixture. This framework, called a mean-field approach, developed by Flory and Huggins defines the free energy of mixing per lattice site for polymer solutions as

$$\Delta F_{mix} = kT \left[ \frac{\phi}{N} \ln \phi + (1 - \phi) \ln(1 - \phi) + \chi \phi(1 - \phi) \right]$$

Where χ is the Flory interaction parameter, k is Boltzman’s constant, T is temperature and N is the number of monomers of the polymer. This approach does not account for any connectivity of the chains but simply the solvent-polymer, polymer-polymer and solvent-solvent forces in the mixture.

Using minimum points of free energy at various values of the Flory parameter we construct a phase diagram that is a function of volume fraction and χ (Figure 1). This parameter is classically represented in the form $\chi \approx A + \frac{\beta}{T}$ in mean field theory where A
is the “entropic” contribution of $\chi$ and $B/T$ is the “enthalpic” contribution. Relevant for this thesis is that if $B > 0$ then the system is said to exhibit an Upper Critical Solution Temperature (UCST) where the system phase separates upon cooling ($\chi$ decreases as temperature raises) and the phase diagram appears as in Figure 1. If $B < 0$ then the system is said to a Lower Critical Solution Temperature (LCST) where the system phase separates upon heating ($\chi$ decreases as temperature decreases) and the phase diagram is inverted along the $x$-axis. What is important to recognize is in an equilibrium case, where $\chi$ is held constant, the system will equilibrate to relative species volume fractions in accordance with this mathematical description, regardless of the sign of $B$. If one is in the two-phase regime of a UCST polymer (inside the parabolic arc) due to a combination of $\chi$ and $\phi$ then the system will phase separate into a polymer-dilute phase ($\phi'$) and a polymer dense phase ($\phi''$). Conceptually it is important to understand that this is an equilibrium system and thus if one changes $\chi$ then the system wants to re-establish equilibrium to a new $\phi'$ and $\phi''$. Additionally, by changing this one parameter and the shape of the phase diagram by modifying the chemical identity of the polymer and the molecular size of the chain we can engineer systems that have different dilute phase concentrations and dense phase concentrations that change the availability of polymer to interact with its surroundings.
1.1.3 A brief history of phase separation, polymer science and membrane-less organelles

I am not the first scientist to be captivated by the order-disorder paradigm of phase separation nor the first to identify the importance of this paradigm inside living cells. However, what has been exciting to witness during the time course of my thesis is the realization in the biology community that the liquid-liquid phase separation is indeed critical for many biological functions. I have been heavily influenced by the past research
of both communities and would like to take a moment to discuss the interweaving research of these classical disciplines.

The biological community that studies these so-called "membrane less organelles" and the polymer science community have predominantly coexisted as two separate communities with brief periods of intersection[1]. Some of the most important discoveries that impacted both polymer science/protein polymer and membrane less organelles/intrinsically disordered proteins are documented in Figure 2. As early as 1826, the connection was made between the similarity of synthetic rubber to 'albumen-like' material contained within cells[2]. Naegli's observations of granules of various biologic materials is mentioned in the polymer chemistry literature, especially in regard to the separation of large polymers from solution[3]. By the time that membrane less organelles had first been robustly characterized by electron microscopy in 1962[4], polymer chemists had developed the experimental techniques to accurately determine the molecular weight of large, unstructured polymers [5, 6] and had a clear statistical understanding of their solution conformations[6-9]. This understanding enabled the rapid acceptance of these 'nuages' as disordered, high molecular weight, protein and nucleic acid condensates[5-7, 10]. Historically, the recognition that membrane less organelles contain unstructured proteins and that synthetic polymers are frequently unstructured was a moment in time when these fields could have overlapped productively, but this was not be, as the emergence of molecular and structural biology took biochemistry
Figure 2: Significant events in the history of protein polymers and membrane less organelles. Protein polymers are an offshoot of polymer science, both of which are denoted by the blue line. Membrane less organelles emerged from biochemistry and cell biology and are denoted by the red line. At various points in the 20th century, advancements made by polymer science and biochemistry and cell biology influenced one another, which is symbolically represented in the two lines crossing over one another. With time, however, both fields drifted away from each other, as graphically represented by the divergence of the parallel red and blue lines away from one another. [1-6, 8-61]

Changing research motivations in the middle of the 20th century reduced the collaborative efforts between polymer science and biologists. This was driven in part the rise of molecular biology and recombinant protein expression that provided the ability to design, synthesize and purify proteins at the single residue level, and the emergence of x-ray crystallography as the dominant structural characterization technique of proteins.
This led to the era of modern structural biology that naturally focused on the major class of known polypeptides that are folded and have a clearly defined function and are amenable to structural characterization by x-ray diffraction.[20] Recombinant protein expression and site-directed mutagenesis reinforced the structure-function paradigm by enabling study of structure-function at the single residue level, providing an unprecedented understanding of how many proteins function. An unfortunate consequence of the structure-function era, however, was the fact that disordered elements within proteins or proteins that were largely unstructured became the “dark matter” of the cell, given their lack of structure and somewhat opaque —difficult to discern— function. As their structure could not be resolved with x-ray crystallography, their role was subsumed within the dominant protein structure-function paradigm as nonfunctional connecting segments. As we now know, these unstructured—intrinsically disordered—proteins are not an evolutionary dead end, but play an important role in biology, and in human health and disease[62].

1.2 Intrinsically disordered proteins

1.2.1 Intrinsically disordered proteins in nature

Intrinsically disordered proteins (IDPs) or Intrinsically disordered protein regions (IDPRs) are proteins or segments of proteins that are lacking in a “single, well-defined equilibrium structure and exist as highly dynamic, heterogeneous ensembles of conformers resulting from their relatively flat free-energy surface – V. Uversky”[63]. Although clearly many proteins can be defined by their three-dimensional structure it is
increasingly obvious that intrinsic disorder is very common, with over 40% of human protein coding genes containing at least one disordered segment that is >30 amino acids in length and 6% of all human protein coding genes that have zero structural annotations suggesting they are completely disordered[64, 65]. These proteins and protein segments were not passed over due to ignorance but simply they were misfits in the dominant structure-function paradigm that defined biology for over 50 years in the mid-20th century. It is clear from the historical record an acknowledgement that disordered proteins and disordered protein segments existed (see Figure 3), researchers simply did not know how to fit them into the prevailing narrative at the time.

Figure 3: One of the first schematic drawing of a polypeptide backbone by Richardson[66]. Spiral ribbons represent α-helices and arrows represent strands of β sheet. Note that even in these first images that there are suggestions of unknown folded structure (denoted by floppy regions) that was not clear from X-ray crystallographs.
The discovery of the prevalence of these proteins prompted investigation into their functional importance. In the 21st century, researchers have revealed a number of functions of IDPs including prominent roles in cell signaling, regulation, recognition and disease pathogenesis[67, 68]. A feature of many of these systems is the importance of phase separation to the phenotype observed in the cellular system[60, 69-71]. Despite the increased attention to these membrane-less complexes or biological condensates, the functional significances of phase separation specifically are still unknown. It is clear that there are number of situations in which complexation with nucleic acids or another protein is a critical component of the liquid droplets[72-75] but there are other examples suggesting IDPs alone are sufficient to promote phase separation[59, 76]. As we know from the synthetic polymer literature, this phase separation may be of a UCST or LCST type transition and indeed there are examples of both in the literature with a bias towards UCST phase behavior[59, 76, 77]. However, despite the clear evolutionary importance of UCST phase behavior, very little is known about how UCST behavior is encoded at the chain level – in synthetic polymers and protein polymers.

UCST is often described as an enthalpy driven process and thus relying on key residue interactions to drive phase separation upon cooling[78-81]. Notably, the Pappu lab has posited an influential theory of a “sticker” and “spacer” model for understanding phase separation in which the strength of the stickers and their density (spacing) control the phase diagram[82]. Wang et. al. identified the importance of tyrosine and arginine residues in defining the saturation concentration of liquid droplets formed by the fused in sarcoma (FUS) protein family[83], in line with previous reports describing the importance
of RGG/RG motifs in driving phase separation with or without RNA[84] and aromatic residues similarly [85, 86]. Previous work from the Chilkoti lab uncovered the importance of charge balance, molecular weight, preference of Arg over Lys as the positive residue and the importance of aromatic residues in converting a LCST polypeptide into a UCST polypeptide[78]. Despite this growing knowledge base, the aforementioned studies are isolated examples which do not offer a wholistic understanding of how UCST phase diagram can be controlled rationally. Thus, there is an urgent need for a systematic mutation approach to controlling UCST phase behavior to more purposefully understand the molecular driving forces that underly this widespread and life-sustaining phenomenon.

1.2.2 Engineered IDPs

1.2.2.1 Elastin like polypeptides (ELPs)

In addition to the number of naturally occurring IDPs, there are a number of engineered IDPs that borrow motifs from these natural sequences but combine them in un-natural ways. Elastin-like polypeptides (ELPs) are one class of these artificial polypeptides composed of the VPGXG pentapeptide repeat unit from human tropoelastin where X can be any amino acid except proline. This repeat unit is derived from the hydrophobic domain of tropoelastin, a naturally occurring IDP that provides structural function in humans[87, 88]. These recombinant polymers display lower critical solution temperature (LCST) phase behavior that leads to the formation of an insoluble coacervate phase above the cloud point of the polymer, similar to tropoelastin[89]. The
LCST of ELPs can be tuned to respond to different stimuli such as temperature\[88, 90\], the type and concentration of salts\[91\], other co-solutes such as proteins\[92, 93\], pH\[94, 95\], and light\[96\].

Viewed from the perspective of IDPs, ELPs could be characterized as an idealized IDPs whose disorder in both their solvated and aggregated states is driven by low-complexity, tandem repeats that create disordered structures whose solvation is temperature dependent and which imparts a soluble-insoluble LCST phase behavior to these protein polymers\[97\]. The origins of the extensive biophysical study of ELPs stems from their ability to recapitulate the LCST phase behavior of tropoelastin and its native elasticity. More recently, the stimulus for the exploitation of their phase behavior has been driven by many molecular applications of this class of polymers.

The mechanical properties of ELPs have lent themselves remarkably well to functionality in biomedical and materials engineering. ELPs have been extensively developed as carriers for drug delivery, as their stimuli-responsive behavior, lack of toxicity, and tunable half-life in systemic circulation provide useful attributes for the delivery of drugs. One example is the development of highly potent chemotherapeutic-loaded ELP nanoparticles for cancer therapy. In this approach, an ELP is designed with a short peptide tag at one of its termini that contains multiple copies of a Cys residues separated by a diglycine spacer. Attachment of multiple copies of a hydrophobic cancer drug to the Cys residues then spontaneously triggers the self-assembly due to the hydrophobicity intrinsic to the drug into near-monodisperse micelles with size ranging between 40-60 nm. Dox-loaded ELP nanoparticles were shown to be highly potent as a
single injection of an ELP-based micelles containing doxorubicin (Dox) eradicated colon carcinoma in a murine model[98]. This initial demonstration of the drug-loaded ELP nanoparticles did not exploit the thermal responsiveness of ELPs. In a subsequent study, McDaniel et al showed that the ELP could be re-engineered by choice of the guest residue X in the VPGXG repeat unit and the molecular weight to a temperature that has a nanoparticle to aggregate T\textsubscript{t} of ~40°C. These nanoparticles were shown to aggregate in the vasculature of tumors that were externally heated to 42°C, a temperature that is clinically compatible. In a different design of ELPs, ELP block co-polypeptides have been designed with two distinct segments that are covalently linked, where one block has greater hydrophobicity than the other block. Below a critical micellization temperature (CMT), these ELP block co-polypeptides are soluble, but as the temperature is raised above the CMT, the more hydrophobic block selectively desolvates, thereby making the entire protein chain amphiphilic enough to self-assemble into spherical micelles. These nanoparticles have been used in vitro to trigger the formation of nanoparticles for the multivalent display of cell penetrating peptide motifs for controlled cellular uptake and in vivo for the thermal targeting of drug to tumors heated to 42°C[99, 100]. In the latter case, subsequent heating and cooling of the in vivo area of interest created a "thermal pump" capable of continually creating a diffusion gradient of the drug across the tumor vasculature.

Sustained release delivery systems have also been developed that consist of the fusion of peptide and protein drugs to soluble ELPs that undergo phase separation in vivo. This approach has been used for the controlled release of peptide drugs from sub-
cutaneous depots for the treatment of type diabetes[101]. Amiram et al genetically fused glucagon-like peptide-1 (GLP-1) to an ELP with a phase transition below physiological temperature. When injected under the skin, the polymer formed an insoluble coacervate, which was slowly released the GLP-1-ELP fusion over time, controlling glucose levels in mice for up to 5 days with a single injection.

1.2.2.2 Other notable proteins with repetitive domains and phase behavior

Elastin-derived repeat proteins are just one example of an elastomeric tandem repeat proteins. Searching the literature for proteins with similar features – largely unstructured proteins that are rich in proline and glycine and with a highly conserved sequence across species– offers a multitude of other IDPs with a vast array of natural functions. To date, they have been differentiated based on their origin and their macroscopic material properties (i.e. elastin, resilin, collagen, byssus threads et cetera). These diverse materials should instead be thought of as members of a larger family of intrinsically disordered tandem repeat proteins, where the organization and content of amino acids leads to their biological function and hierarchical material properties.

Abductin, resilin and high molecular weight (HMW) wheat gluten are three proteins that most resemble tropoelastins (and ELPs) based on their sequence[102]. Each protein contains long domains of elastomeric repeat units, interspersed with non-repetitive units.
Abductin is a naturally cross-linked elastomer that serves as the primary structural component for bivalves in mollusks, and is characterized by high resilience to compression ratio and a consensus repeat sequence of GGFGGMGGGX—where X can be any amino acid[103]. Resilin is an elastomeric protein predominantly found in Drosophila melanogaster (fruit fly), Anopheles gambiae (mosquito), and arthropod cuticles, where fast, repetitive motion and efficient energy storage is required[104]. Depending on its origin, resilin has a consensus repeat sequences of GGRPSDSYGAPGGGN or AQTPSSQYGAP[105, 106]. Lamprin (GGLGX)[107], the most important protein in lamprey cartilage, has been shown to exhibit elastomeric properties[108]. The high molecular weight (HMW) subunits of wheat gluten are seed storage proteins that store essential nutrients such as carbon, nitrogen and sulfur for growth of seedlings. Their consensus sequences are the hexapeptide repeat PGQGQQ and the nonapeptide repeat GYYPTSPQQ. Elastin, abductin and resilin utilize crosslinking domains (disulfides in abductin, dityrosines in resilin) to form mechanically elastic gels[109].
Figure 4: Plot of glycine and proline chain content of various engineered IDPs against naturally occurring IDP sequences. Dotted area roughly designates IDPs that exhibit amyloid (low Pro and low Gly) and liquid like (high Pro and high Gly) coacervation.

Evaluation of their consensus repeats by their charge/hydropathy balance suggests that abductin and lamprin fall in the molten globule region (low Pro and Gly content) with ELPs, HMW gluten and resilins are predicted to be disordered (Figure 4). This result is expected as abductin and lamprin have repeat sequences most homologous to ELPs. However, Nuclear Magnetic Resonance (NMR), Circular Dichroism (CD), and Fourier Transform Infrared Spectroscopy (FTIR) analysis of short, repeat units of the connecting elastic domains of HMW gluten, resilin and abductin indicate that they are all largely disordered with each displaying varying levels poly-
proline helix (PPII) and β-turn structures, similar to ELPs. Figure 4 shows these and other artificial repeat proteins on the traditional amyloid-elastin graph, where the properties of these proteins can be predicted —albeit to a limited extent— by their Pro and Gly content.

Although each unit (besides lamprin) is disordered, they seem to have used different combinations of amino acids to achieve disorder. For resilin and HMW wheat gluten, the chains utilize zwitterionic charge and hydrogen bonding to water to prevent aggregation. On the other hand, abductin and ELP do not utilize these principles. ELPs in particular, achieve disorder by structure breaking Pro and Gly residues that enable sampling of many possible conformation conformations. Abductin, interestingly exhibits chain disorder without regularly spaced structure breaking proline residues. These examples, when assessed as members within the larger family of IDPs, highlight the challenges that exist when it comes to predicting disorder of protein sequences containing repeated motifs.

1.3 Motivation

The overarching motivation for this work was generated by the lack of information surrounding upper critical solution temperature (UCST) transitions, in both synthetic polymers and intrinsically disordered proteins. In contrast to systems with lower critical solution temperature (LCST) transitions, where clear links between side chain chemistry and molecular weight define the entropic driving forces for phase separation, the
A combination of side chains that promote UCST phase separation are unknown.

**Figure 5**: Difference in number of journal article publications that mention LCST or UCST phase behavior in the title or abstract according to PubMed as of July 12, 2019.

This fundamental lack in understanding results in a lack of application of these materials. Compared to >1400 publication that mention LCST phase behavior in the title or abstract between 1990-present, there were a paltry 180 that mention UCST phase behavior, nearly a ten to one bias of LCST phase behavior. Thus, a major aspirational goal of this thesis was to establish principles for controlling UCST phase behavior of proteins, manifested in two overarching aims:

1. Elucidate the architecture parameters that control UCST phase separation
2. Explore applications where control of UCST phase behavior is required at various length scales.
1.4 Specific aims

Aim 1: Create an engineered library of intrinsically disordered proteins with variable UCST behavior.

Critical for the application of disordered protein in biology is the control over the phase diagram shape. The literature, although sparse on specific information of how to encode UCST phase behavior does provide a number of design considerations. First, is that UCST and LCST phase behavior can indeed coexist and thus there is the potential for significant overlap in the sequence space between LCST and UCST phase behavior. Secondly, there are few examples of polypeptides in the literature, designed or naturally occurring, that are shorter than 100 amino acids in length and have phase behavior. Thirdly, there are specific chemical identity heuristics for sequence design that appear to bias a polypeptide sequence to UCST phase separation. Namely, overall charge balance, preference of Arg over Lys as the positive residue, statistical abundance of aromatic residues, polar non-charged residues, Gly and Pro. Our working hypothesis was that it was likely that each of these parameters affect the binodal phase boundary but likely to varying degrees on percent mole basis. Theoretically, UCST is thought to arise through pairwise interactions. Thus, we hypothesize that particular pairs of interactions are more important than others. To determine the parameters of greatest interest, we created a systematic library and performed a form a type of “loss of function” or “gain of function” screen, beginning with a repetitive protein sequence that is a hybrid of previously explored repetitive sequences and naturally occurring phase separating proteins. All sequences created by this library approach will be recombinantly produced
and purified using the intrinsic phase separation of the protein sequence. Their phase separation will be evaluated under physiological conditions using a minimum of three concentrations to determine the shape of the binodal line in the dilute regime ($\phi < 0.1$).

**Aim 2: Designer membrane-less organelles using engineered intrinsically disordered proteins with non-canonical functionality.**

Interestingly, the most popular studied IDPs appear to manifest UCST-like phase separation in cells. Despite the growth of this field, there is very little information regarding the importance of sequence and its effect on the function of a membraneless organelle. Thus, we sought to develop a robust membrane-less organelle, comprised of sequences identified in Aim 1 of this proposal, that could be engineered from the bottom-up with non-canonical functions. We hypothesized that the same fundamental principles identified in Aim 1 will translate to *in vivo* design and that by via genetic fusion of our membrane forming materials of interest with a function molecule we can create domains of new activity within a cell.

**Aim 3: Creation of hybrid UCST-LCST micelles capable of multivalent display of non-peptide functional moieties.**

Nanoparticle design is an extremely well-developed field. With the development of polypeptides with tunable UCST behavior, however, there are new opportunities that present themselves to create stimuli-responsive materials created from two separately engineered systems – the LCST phase behavior of ELPs and the UCST behavior of polypeptides derived in this thesis. These materials have the chance to solve some outstanding challenges with the self-assembly and nanoparticles communities. Namely,
A significant and pervasive problem in the active and passive targeting fields is the conjugation of biomedically relevant active molecules to synthetic polymers or lipids that comprise the assembly domain and control over the size and morphology of the particle. By leveraging the opposing thermodynamic nature of UCST and LCST phase behavior we hypothesize that a block co-polyptide system may provide a more stable and versatile self-assembly platform for ligand presentation and through genetic fusion be amenable to one-pot recombinant synthesis, instead of a laborious, stepwise synthesis, self-assembly, conjugation and final purification work-flow.

**Aim 4: Tunable, systemic release profile of GLP-1 from a subcutaneous implant comprised of a GLP-1-IDP fusion protein.**

Repeat polypeptides and synthetic polymers have been utilized extensively to modulate the half-life of small molecule drugs and protein therapeutics. Our hypothesis is the UCST binodal line acts similarly to a LCST binodal line, in that the left-hand side of the binodal defines the dilute phase concentration in the local injection area. Thus, we can regulate the bio-availability of a peptide drug via the UCST polypeptide chemistry. To accomplish this task, we will first genetically fuse glucagon-like polypeptide 1 (GLP-1) to repetitive protein polymers of variable hydrophobicity and molecular weight. GLP-1 is a potent, stable peptide drug that can effectively lower blood glucose but has a poor pharmacokinetic profile. Previous studies have demonstrated that an optimal design of the polypeptide has a dilute phase concentration at 35°C of between 0.1-500 µM and a molecular weight greater than 35 kilodaltons. We will determine the best method for
establishing a sub-cutaneous protein deposit and then measure the efficacy of these depots by tracking blood glucose and the body weight of the animal.
2. An engineered library of intrinsically disordered proteins with variable UCST behavior.

2.1 Introduction

There is minimal literature precedent in how one should tune the UCST/LCST behavior of synthetic polymers and the LCST of ELPs. Regarding LCST systems, the transition temperature ($T_t$) is entropically driven and therefore the most important parameters for tuning the LCST binodal are hydrophobicity and molecular weight (MW). This hydrophobic effect can be tuned by changing the co-polypeptide or amino acid in the guest residue slot to one that has different water solubility\cite{51, 110}. This simple design tool has been able to tune the LCST cloud point over an enormous range. However, UCST systems rely heavily on enthalpic factors. Therefore, to increase the UCST of a system (become more unstable in water), one could either increase the enthalpic components by increasing the value of any of the positive enthalpic terms in or decreasing the value of the negative enthalpic terms. Likewise, decreasing the value of the entropic term by increasing molecular weight (decreased entropy of mixing) or increasing the hydrophobicity should act increase the $T_t$.

Table 1: Theoretic examples of contributions to enthalpic and entropic terms in Gibbs free energy balance for temperature responsive polypeptides with UCST phase behavior

<table>
<thead>
<tr>
<th>Contribution to $\Delta H$</th>
<th>Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breaking polymer – polymer hydrogen bonds</td>
<td>+</td>
</tr>
<tr>
<td>Breaking water – water hydrogen bonds</td>
<td>+</td>
</tr>
<tr>
<td>Breaking polymer – polymer van der Waal forces</td>
<td>+</td>
</tr>
<tr>
<td>Breaking coulombic attraction between opposite charges</td>
<td>+</td>
</tr>
<tr>
<td>Formation of polymer – water hydrogen bonds</td>
<td>-</td>
</tr>
</tbody>
</table>
Previous work in the lab had sought to understand the Pro-Gly scaffold design and other key residues that could transform an LCST polypeptide to exhibit UCST phase behavior and vice versa[111]. It was shown that their transition temperature was a function of molecular weight and solution concentration (Figure 6 B,D). However, the available polymers failed to have a predictable UCST phase shift given their primary sequence based on overall hydropathy (Figure 6C), partially due to the lack of systematic control of molecular weight, repeat length and inconsistent amino acid selection (Figure 6C). This work described a few important design heuristics namely, >100 residues, charge content >10%, aromatic charge content >10% and a bias towards arginine over lysine[78].
Figure 6: Previous work in the Chilkoti lab with UCST polypeptides – A. Phase behavior of select LCST sequences as a function of the number of repeats. Data collected at 50 µM. B. Phase behavior of select UCST sequences as a function of the number of repeats. Data collected at 50 µM. C. Correlation between average hydropathy and phase behavior in B. D. Correlation between concentration and observed UCST phase behavior. E. Representative phase behavior upon cooling and subsequent heating in buffer. F. Turbidity plots of IDPPs when the charge balance is disrupted by solution pH[78].

However, these parameters could have been surmised with a simple analysis of the exons that encode for the disordered regions of rec-1 resilin. Thus, we were no closer to understanding the critical interactions that encode for UCST phase behavior nor the magnitude of these effects.

In addition to work by the Chilkoti lab, there are other reports in the literature of repetitive protein polymers that exhibit UCST phase behavior. These are summarized in
Table 2. However, none of these studies explored mutations that could influence the phase diagram.

**Table 2: Repeat sequences of other resilin inspired polypeptides in the literature**

<table>
<thead>
<tr>
<th>Title</th>
<th>Repeat Segment</th>
<th>Phase Behavior Observed</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>An16</td>
<td>(GAPAQTPSSQY)$_{18}$</td>
<td>LCST+UCST</td>
<td>[105]</td>
</tr>
<tr>
<td>Dros15</td>
<td>(GGRPSDSYGAPGGGN)$_{18}$</td>
<td>LCST+UCST</td>
<td>[112]</td>
</tr>
<tr>
<td>DroMel13</td>
<td>(GYSGGRPGGQDLG)$_{11}$</td>
<td>LCST</td>
<td>[113]</td>
</tr>
<tr>
<td>Res24</td>
<td>(AQTPSSKQFGAPAQTPSSQFGAP)$_{6}$</td>
<td>Unk</td>
<td>[114]</td>
</tr>
</tbody>
</table>

2.1.2 Approach

There is an inherent problem in protein engineering – throughput. There are no techniques that are efficient enough to cover the combinatorial landscape of even relatively short protein chains ($2^{200} = 1.36 \times 10^{130}$ possible combinations). Thus, we must often turn to nature for inspiration. We began with a proteomic search of the largest database of proteins that we could find of relevance – the IDPs that are major constituents of membrane-less organelles[60]. We were particularly interested in categories of amino acids suspected to drive phase behavior via intrachain interactions, such as charge-charge, cation-pi and hydrogen bonding via non-charged polar residues (Figure 7). The composition of these 63 proteins are remarkably similar to previously identified repetitive protein polypeptides (Table 2) that exhibit UCST phase behavior and their side groups share chemical similarity to synthetic UCST polymers (Figure 8).
Figure 7: Proteomic analysis of IDPs that form biomolecular condensates in various biological systems indicates a high percentage of charged, non-charged polar residues, Gly/Pro residues and overall charge balance (n=63). Boxes indicate 25-75th percentile and whiskers representing 10-90th percentile.
Figure 8: Synthetic polymers with UCST phase behavior. The polymers blocked in blue rely on strong hydrogen bonding interactions between polymer chains to exert UCST phase behavior. The polymers blocked in red likely rely on strong columbic interactions. These are sensitive to buffered and salt conditions. Polymers blocked in black rely strongly on hydrophobic and hydrogen bond formation for UCST behavior.

Using a combination of our own heuristic knowledge and information from this proteomic analysis, we designed a repeating octa-peptide motif that we expected to exhibit robust phase behavior under physiologically relevant solution conditions. The MW of the overall IDP was varied between 17 and 70 kDa to account for observed differences in MW in the IDPRs (Figure 9). Because this set of IDPs exhibited robust phase behavior, we then could probe the substitution of various amino acids in a “loss of function” or “gain of function” screen.
In order to manage the vast sequence space that would exist for an unrestricted octapeptide repeat, the amino acid mutations were bundled into chemically similar categories that preserves categories of intrachain interactions that could contribute to UCST phase behavior. Asn, Gln, Ser, Thr are classified as polar, uncharged amino acids and are substituted in for one another. Arg-Lys and Asp-Glu are pairs of positively charged and negatively charged amino acids. Gly and Pro are placed into a separate category given their unusual structure and importance regarding chain conformations[115, 116]. The remaining amino acids are placed in a “hydrophobic” category[117, 118].

Our reference repeating unit for these repetitive IDPs (RIDPs) is (Gly-Arg-Gly-Asp-Ser-Pro-Tyr-Ser)xx where XX refers to the number of repeats. We will use shorthand notation to refer to sequences throughout the text where the bracketed letter
represents the manipulated residue composition. All substitutions to this repeated motif reflect changes in this sequence according to their amino acid category. For example a substitution in the reference repeating unit of Ser with Gln would result in a change of notation from “[S]-XX” to “[YS:ZQ]-XX” where the Y to Z ratio represents the least common denominator ratio of Ser to Gln ratio in the final protein. [Q]-XX would therefore indicate 100% of Ser substituted with Gln. [S:Q]-40 would thus represent half of all Ser replaced with Gln. A 100% substitution of His for Tyr would be denoted as [H]-XX and so on.

2.2 Methods

2.2.1 Gene synthesis

A search of the literature provided an excellent list of intrinsically disordered proteins or protein regions are present in genes known to form membrane-less organelles [60]. Each gene was divided into disordered regions and ordered regions according to the Predictor of Natural Disordered Regions (PONDR) VSL2 algorithm which is a meta-predictor of protein disorder of various lengths [119]. Amino acid quantity was normalized to total protein length.

Each octapeptide amino acid motif inspired by our proteomic analysis was propagated twenty times in silico. This repetitive amino acid sequence was fed into an algorithm that creates an optimally non-repetitive DNA template from a repetitive protein gene[120]. This 20-mer repeat gene was then ordered from IDT with Gibson assembly overhangs for easy insertion into our in house modified pet24+ vector[120]. To increase
the number of total repeats of the gene, we performed iterative cloning steps of Recursive Directional Ligation by Plasmid Reconstruction adding an addition twenty repeats during each step. This method has been described previously[121]. Transformations were performed into the desired E. Coli cell line – BL21 (DE3) for recombinant expression and single plasmid confocal experiments and a modified BL21 (DE3) cell line termed KRX by Promega that contains a mutated LacZ gene for enzymatic experimentation.

In experiments with dual expression, genes were inserted into the pBAD33.1 vector by cutting custom pet24+ vector and pBAD33.1 cut with HindIII and XbaI. Gel purification was used to isolate the gene of interest from the housing pet24+ vector, which was then ligated into the similarly cut pBAD33.1 vector. Co-transformation was performed with ~1 ng final concentration of each plasmid on kanamycin/chloramphenicol dual selection plates.

2.2.2 Protein purification

Individual liquid cultures of BL21 E. Coli strains each harboring our gene of interest were inoculated into 5 ml of Terrific Broth (TB) medium from frozen glycerol stocks and grown to confluence overnight (16-18 hours). Cultures were then inoculated at a 1:200 dilution in 1L TB media supplemented with 45 μg ml⁻¹ kanamycin. Cells were grown at 37°C in a shaking incubator (~200 r.p.m.) for 9 hr, at which time protein expression was induced by the addition of 500 μM Isopropyl-β-D-thiogalactoside (IPTG). Cells were then incubated at 37 °C (shaking at ~200 r.p.m.) for an additional 18 hr.
Protein was then purified from the insoluble cell suspension fraction. In brief, cell pellets were isolated by centrifuging cultures at 3500 RCF and resuspending in 20 mL of milli-Q water. Cells were then lysed by sonicking the cell solutions for 2 minutes, with 10 seconds of pulsing followed by 40 seconds of rest on ice (Misonix; Farmingdale, NY)

Centrifuging each lysate suspension at 20,000 RCF for 20 minutes results in a soluble and insoluble fraction. The supernatant was discarded with the insoluble fraction resuspended in an approximately equal volume of 8M urea + 140 mM PBS (~6-8 mL). For proteins with a fluorescent fusion tag, the insoluble fraction was resuspended in 3X insoluble volume at a final concentration of urea of ~2M to prevent protein misfolding. This suspension was heated for 10 min in a 37°C water bath and then centrifuged at 20,000 RCF for 20 minutes. The supernatant was collected from this suspension and dialyzed in a 10 kDa membrane (SnakeSkin™, Thermo Fischer Scientific) against a 1:200 milli-Q water solution at 4°C. The dialysis water was changed twice over a 48-hour period. From inside the dialysis bag, both insoluble and soluble components were collected and centrifuged at 3500 RCF for 10 minutes and 4°C. The supernatant was removed and the remaining insoluble pellet containing the protein of interest was lyophilized for a minimum of three days to remove all water from the pellet.

Protein purity was characterized by 4–20% gradient tris-HCl (Biorad, Hercules, CA) sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and staining with either 0.5 M copper chloride or SimplyBlue™ SafeStain (Thermo Fischer Scientific). Protein yield was determined by weight after lyophilization.
2.2.3 Characterization of phase separation

2.2.3.1 Temperature dependent UV-vis spectrophotometry

Cloud point transition temperatures ($T_c$) were determined via temperature-controlled spectrophotometry using a Cary 300 (Agilent Technologies). Samples containing various concentrations of protein in 140 mM PBS were cooled at 1°C min$^{-1}$ while the absorbance at $\lambda=350$ nm was recorded every 1°C. Absorbance was normalized to the absorbance at the highest temperature point collected, corresponding to the more soluble point during a given experiment. The cloud point was determined as the maximum in the first derivative of the absorbance as a function of temperature. Transition temperature was calculated by the point of minimum slope. Saturation concentration was calculated using the natural logarithm fit line created from a minimum of three cloud points at different volume fractions. Error bars are standard error of the mean from three repeats of a minimum of three transition temperatures.

2.2.3.2 Temperature depending dynamic light scattering

Dynamic light scattering (DLS) measurements were performed over a temperature range of 10-80°C using a Wyatt DynaPro temperature-controlled microsampler (Wyatt Technology, Santa Barbara, CA). Samples for the DLS system were prepared in 1x PBS and filtered through 0.02 µm Whatman Anotop sterile syringe filters (GE Healthcare Life Sciences, Pittsburgh, PA) into a 12 µL quartz crystal cuvette (Wyatt Technology, Santa Barbara, CA). 5 acquisitions were taken at each temperature.
for a 5 second duration, and the results presented represent the mean radius of hydration \(R_n\) of the sample at each temperature.

### 2.2.3.3 Temperature gradient microfluidics

High concentration RIDP stock solutions \(\phi = 0.6\) were prepared by resuspending a mass of lyophilized RIDP pellets with an appropriate volume of phosphate buffer saline solution (PBS) at a solution pH of 7. The concentration was converted to mg ml\(^{-1}\) by assuming that the density of the RIDP was 1 g ml\(^{-1}\). The RIDP stock solution was heated in a water bath at 85°C for 60 minutes and mixed periodically along with sonication to ensure homogeneity. Lower concentration samples were made by mixing the initial stock solution volumetrically with PBS at a pH of 7. To prepare for temperature gradient microfluidics (TGM) measurements, the solutions were loaded into 12 mm x 1 mm x 0.1 mm rectangular borosilicate glass capillary tubes (VitroCom, Inc.), by capillary action, and sealed with wax to avoid sample evaporation and convection. The capillary tubes were held in contact with a hot plate at 85°C housed within an incubator at 65°C during the loading process. The high temperature environment ensured that the RIDP solutions were held above the critical phase transition temperature (~85°C for [S]-20). Capillary arrays were prepared by taping several capillaries together. The arrays were stored at 85°C in an oven for 10 minutes prior to subjecting them to the temperature gradient experimentation.

The temperature gradient device imposed a linear temperature gradient across the RIDP solutions. This was accomplished by placing the glass capillary array into
thermal contact with a heat source on one side and a cold sink on the other. The sample was then bathed in white light. The light that scattered by the phase separated RIDP droplets at cold temperature was imaged via dark-field microscopy. The temperature gradient was calibrated for each experiment using two reference solutions placed alongside the RIDP samples of interest. The cold temperature calibration reference contained 10 mg ml\(^{-1}\) poly(N-isopropyl acrylamide) (PNIPAM) with MW = 1.868 \(\times\) 10\(^5\) g mol\(^{-1}\) in H\(_2\)O (Polymer Source, Inc.). The hot temperature calibration reference contained 10 mg ml\(^{-1}\) poly(ethylene oxide) (PEO) with MW = 9 \(\times\) 10\(^5\) g mol\(^{-1}\) in a 1M NaCl aqueous solution (Sigma-Aldrich). The LCST of each reference solution was obtained with a melting point apparatus that measured light scattering as the temperature was increased at a rate of 0.5 K min\(^{-1}\). When placed onto the temperature gradient device, the reference solutions became cloudy at temperatures above the LCST. The pixel position of the LCST was obtained by the onset of light scattering intensity relative to the low intensity baseline on the cold side of the capillary. The temperature gradient was calculated using the pixel positions and the LCSTs of the two samples, assuming a linear relationship between position and temperature.

### 2.2.3.4 Fitting of phase diagram binodal

Fits for the roughly dilute, overlap and semi-dilute regions of our obtained phase diagrams were calculated using fitting methods for lower critical solution transition polypeptides adopted for upper critical solution transition polypeptides as described previously[122].
Briefly, for low volume fractions (φ < 0.1), RIDPs exhibit roughly a log-normal dependence on UCST cloud point with respect to volume fraction as seen with other repeat polypeptides [78, 123]. For the high density regime (φ > ~0.4) using surface tension scaling methods previously described for elastin-like polypeptides [124], we determined the coefficients of proportionality (A) and estimated theta temperature (θ) of [S]-20 and [Q]-20 to be A= -0.00092, θ = 389 K and A= -0.00092, θ = 392 K respectively.

In a poor solvent, the surface tension of a dilute phase globule γ can be written in the form \( \gamma \approx C k b^2 \phi''^2 \) where \( b \) is the polypeptide Kuhn length (\( b = 2.2 \) nm as measured by Fluegel and co-workers for other repetitive polypeptides [125]) and \( C \) an adjustable coefficient. Replacing the surface tension \( \gamma \) by \( k^2 \phi_2^2 A^*(T - \theta) \) we obtain an equation for the temperature dependence of the coacervate volume fraction

\[
\phi_2 = \left[ \frac{A^*}{C} b^2 (T - \theta) \right]^{\frac{3}{4}}
\]

Using a least-squares fit in Igor (WaveMetrics Inc. Portland, OR), we adjust the coefficient \( C \) for this temperature dependence to match the measured [S]-20 binodal and the [Q]-20 binodal points to determine \( C = 0.62 \) & 1.05, respectively.

Closer to the critical point in the so-called Ginzburg zone [79] one needs to use the critical Ising model to describe the phase behavior of polymer solutions. The phase boundary in the critical zone varies more gradually than predicted by mean field theory:

\[
\phi_2 - \phi_1 = C_c \left( \frac{T}{T_c} - 1 \right)^{0.3}
\]

Where \( T_c = 351.5 \) K for [S]-20 and \( T_c = 332.3 \) K for [Q]-20, 0.3 is the critical Ising exponent (Flory-Huggins mean field value is 0.5) for both [S]-20 and [Q]-20, and \( C_c \) is
the fitting coefficient. We calculated fitting coefficients in Igor (WaveMetrics Inc. Portland, OR) equal to 1.29 and 1.27 for [S]-20 and [Q]-20 respectively. Note that we calculated $\phi_1$ explicitly using data collected with UV-Vis spectrophotometry.

### 2.2.3.5 Creation of water in oil droplets with chip microfluidics

To create water-in-oil emulsion droplets, two liquid phases – a dispersed, aqueous phase containing protein of interest in 140 mM PBS and an organic, continuous phase comprised of 75%/5%/20% vol/vol TEGOSOFT® DEC/ABIL® EM 90/mineral oil – were injected into the microfluidic droplet generators at constant flow rates using precision syringe pumps. The flow rates of the dispersed and continuous fluids were tuned to ensure droplet formation in the dripping regime; in these experiments, the dripping regime was achieved using a constant flow rate of 500 $\mu$L hr$^{-1}$ for the organic continuous phase and 50-75 $\mu$L hr$^{-1}$ for the aqueous, dispersed phase. The production of droplets within the microfluidic device was monitored using a 5× objective on an inverted microscope (Leica) equipped with a digital microscopy camera (Lumenera Infinity 3-1 CCD).

### 2.2.3.6 Temperature controlled fluorescent microscopy of protocell droplets

Water-in-oil droplets were collected on a glass microscope slide and cooled using a precision Peltier heating and cooling stage (Linkam LTS120) equipped with a temperature control unit (Linkam PE95). The spatial distribution of Alexa Fluor 350-labeled (25% molar fraction N-terminal labeled) [Q]-20 and Alexa Fluor 594-labeled [22R:18D]-20 were characterized via fluorescence microscopy using an upright Zeiss
Axio Imager D2 microscope with a 20× objective and the appropriate filter set. Similarly, intracellular patterning of RIDP-superfolder GFP over time was characterized via fluorescence microscopy using an upright Zeiss Axio Imager D2 microscope with a 20× objective and the appropriate filter set (ex 470/40, em 525/50). Cell fluorescence was calculated using ImageJ software. Temperature ramps began at various temperatures but always were set to a constant speed of 5°C min⁻¹.

2.2.3.7 Differential scanning calorimetry

DSC experiments were performed on a differential scanning calorimeter (Nano DSC II Model 6100; Calorimetry Sciences Corp., Lindon, UT). Sample solutions (50 µm in 140 mM PBS) were cooled from 42 to 4°C at a rate of 1°C min⁻¹ and then heated to 42°C at the same rate. Data was analyzed in Nano Analyze software.

2.3 Results and Discussion

2.3.1 Repetitive IDPs exhibit robust and reversible UCST phase behavior in an aqueous environment

One advantage of repetitive protein polypeptides is their minimal interactivity with other proteins or biomolecules, given their limited frames for unique sequence epitopes. This characteristic combined with reversible aqueous two-phase separation enables column free purification by simple transition cycling between the one- and two-phase regime using alternating temperature and solution conditions. An extreme example of this purification process is shown in Figure 10, where the highly expressing RIDP
completely isolates from the cell lysate under centrifugal forces and maintains the ability to phase separate upon further cooling beyond the centrifugation temperature.

Figure 10: UCST polypeptides form dense, exclusionary phases even in complex medium. One observes almost complete separation from all other cellular proteins and debris present in the cell lysate after centrifugation, facilitating purification from the insoluble cell lysate fraction without affinity tags. Image take during [Q]-20 purification.
Subsequent removal of the protein poor supernatant and repeated transitioned cycles results in 95-99% pure protein as observed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel band analysis (Figure 11), yielding typically 25-300 mg of purified protein per liter of culture media.

![SDS-PAGE gel](image)

*Figure 11: Example SDS-PAGE gels of purified proteins used in this study.*
Repetitive genes [S]-20 and [Q]-20 exhibit UCST phase behavior in vitro. To fully classify the phase separation behavior, we employed three different techniques. First, we utilized droplet microfluidics, where monodisperse water droplets are formed in an oil solution containing our protein of interest (Figure 12).

Figure 12: Visualization of UCST phase separation of [Q]-20 in water-in-oil droplets with fluorescent microscopy. Upon cooling, multiple nucleation sites form and slowly coalesce with one another into a single dense phase. Upon reheating, equilibrium with the surrounding dilute phase is constantly re-established, leading to a higher concentration dilute phase and smaller volume occupied by the dense phase. $\phi = 0.0018$, scale bar = 50 µm.

Here the IDP phase separation can be directly visualized in spatially limited compartments to observe the types of structures formed as one cools the surrounding medium. Proteins suspended in water droplets are minimally influenced by their surroundings, providing a closer analogue to the cellular environment. These repetitive IDPs exhibit classic liquid-liquid phase separation where, upon crossing the phase
boundary by cooling from 50°C to 10°C, multiple nucleating sites appear (Figure 12D Panel 2) that wet with one another and quickly coalesce into a single, spherical protein-dense phase that is in equilibrium with the surrounding protein-poor phase Figure 12D Panel 3 & 3-1/3-2). Upon reheating to 50°C, one observes the shrinking in size of the dense phase as the IDP re-solubilizes and the system attempts to re-establish equilibrium (Figure 12D Panel 4-1/4-2). From this experiment we conclude that repetitive polypeptides exhibit reversible UCST phase separation that matures via coalescence and growth kinetics. An idealized schematic of this behavior is depicted in Figure 12. A wider field of view that demonstrates the consistency of this transition can be found in Figure 13.
Figure 13: Wide-field fluorescent microscope images of fluorescently labeled [Q]-20 inside water-in-oil compartments related to Figure 12. [Q]-20 was labeled with AlexaFluor 350 via NHS chemistry and resuspended in 140 mM PBS at pH 7.4 to a final $\phi = 0.003$. Water-in-oil mixture was transferred to glass slide and cooled from 50°C to 10°C. Scale bar = 20 µm
Second, we employed temperature dependent dynamic light scattering (DLS) to observe the two-phase separation in bulk. In brief, polypeptides were resuspended in 140 mM PBS and passed through a 20 nm filter while heated to 80°C into a quartz cuvette. In Figure 14, a solution of this filtered [Q]-20 is heated to 80°C and then cooled to 10°C while collecting scattered light. In these experiments, we observe a transition from small molecules of approximately 4 nm in size to large aggregates larger than 1 micron. This transition is quite sharp – occurring within a 2-degree Celsius window.

Third, we employed temperature-dependent turbidity measurements at a set wavelength of 350 nm to characterize the opacity observed during a phase separation event while heating and cooling the UCST solution at a rate of 1°C min⁻¹ (Figure 15). Using this technique, we can capture partial phase diagrams of each polymer of interest as a function of many different sequence and solution parameters for direct comparison. At
dilute volume fractions of [Q]-40 we observed different UCST cloud points that increase according to a natural logarithmic scaling \( T_i = m \ln([\text{RIDP}]) + b \). We also confirmed the full reversibility of these proteins, with <1°C difference in UCST cloud point after 10 successive heating/cooling ramps (Figure 15). This was our main method of characterizing UCST cloud points in this manuscript because of its high throughput capacity.

![Figure 15: UCST cloud points are affected by polypeptide volume fraction in solution. This behavior follows a natural log dependence in the dilute regime with higher R2 fit (R2 = 0.98). Right hand side panel - cyclic cooling and heating cycles exhibit minimal hysteric behavior. Optical turbidity measured at 350 nm of repeated cooling and heating curves of [Q]-20 @ ϕ = 0.0025 between 40°C and 30°C.](image)

In summary, repetitive protein polypeptides similar in sequence to naturally occurring IDPs exhibit reversible UCST phase separation in aqueous solution. Their disordered chains can be easily purified from bacterial cell lysate and phase separate at a precise temperature dictated by their solution concentration into micron sized aggregates that exhibit coalescence and growth mechanics.
2.3.2 Arginine composition, aromatic to aliphatic ratio, charge balance and molecular weight define UCST cloud point

UCST phase behavior in synthetic polymers can be sensitive to a number of chain mutations, particularly in buffered conditions[80]. Therefore, in an effort to understand the effects of a particular substitution on UCST cloud point and to provide a reference point from which to predict phase separation behavior, we chose to create well mixed polypeptides consisting of one well characterized repeating unit (a = [S]-20) and a repeating unit of interest (b). By this method, we are essentially performing a loss of function screen for UCST cloud point upon the repeating polypeptides b (Figure 16A), akin to copolymerization studies conducted in traditional polymer synthesis experiments. Due to experimental and instrumental limitations, loss of function —or undetectable phase separation—is practically described in this thesis as a transition temperature < 4°C at volume fractions less than or equal to 0.1 in 140 mM salinity, buffered solvent. A total length of 40 repeats was chosen to approximate the median length (~320 amino acids) of IDPRs found in the proteomic analysis.
Figure 16: A. Schematic describing the methodology for blending polypeptide A & B. A reference polymer “a” with a high UCST cloud point can be blended with polypeptide “b” of interest to probe “loss of function” of UCST phase behavior. Polypeptides were mixed evenly in factors of twenty total repeats between 0 and 100% polypeptide B. Linear blending of polypeptide a (GRGDSPY5) and polypeptide b results in a linear combination of UCST phase behavior at a volume fraction of 10⁻³ (R²=0.97).

Each fusion gene was well mixed to reduce co-polypeptide ‘blockiness’ which has been shown in LCST polypeptides to lead to nanoscale self-assembly instead of the desired liquid-liquid coacervation[126]. Using two polypeptide repeats (Gly-Arg-Gly-Asp-Ser-Pro-Tyr-Ser) and (Gly-Arg-Gly-Asp-Gln-Pro-Tyr-Gln) that both exhibit detectable phase behavior in vitro, we created mixed polypeptides ranging from 0 to 100% of polypeptide (Gly-Arg-Gly-Asp-Gln-Pro-Tyr-Gln) (Figure 16B). We observed that the behavior of mixed polypeptides can be linearly interpolated between polypeptide a and b (R² = 0.97) (Figure 16B inset). This allows for the extrapolation towards homopolymers that exhibit a UCST cloud point beyond the observable range of detection, which allows for relative assessment on a single axis scale.
Implementing this strategy, we tested the effect of substituting 15 unique repeating motifs on the phase diagram of the polypeptide where a single mutation to the reference sequence was made at a site-specific location. We found that even subtle substitutions (< 12.5% change in polypeptide composition) to the repeating amino acid motif are capable of changing the Tt @ 37°C of the repeat polypeptide by over four orders of magnitude. This can be achieved most effectively (largest change in UCST cloud point compared to percent change in composition) by reducing the number of Tyr with aliphatic amino acids (I, V, A) (Figure 17) and by replacing the number of Arg residues with Lys (Figure 18).

Figure 17: A. Substitution of aromatic Tyr residues with aliphatic Val amino acids dramatically reduces UCST cloud point. B. Effect of replacing aromatic Tyr with various other hydrophobic type amino acids on Tt @ φ = 10⁻³, In 140mM PBS, pH 7.4.
Figure 18: A. Substitution of Arg residues with Lys amino acids dramatically reduces UCST cloud point. B. Effect of replacing charged amino acids with the partnering charge on $T_c$ @ $\phi = 10^{-3}$, In 140mM PBS, pH 7.4.

In both circumstances, between a 6.25% and 9.37% change to the amino acid composition eliminates phase separation in the measurable temperature range. Utilizing information about each individual repeating unit’s cloud point, we were able to make tri-polypeptide blends with substituted Val and Lys that exhibited predictable UCST phase behavior based on interpolations of the base repeating block’s UCST phase behavior (Figure 19). In each case, the observed transition temperature was less than 10% of the predicted transition temperature, suggesting that in a well-mixed blend of repeating motifs, each motif operates with independence to define the final cloud point observed.
Figure 19: Phase behavior of polypeptide blends containing three unique repeating motifs. A. Partial binodal phase boundary of tri-block polypeptides with varied ratio of aromatic:aliphatic and Arg:Lys residues. Data collected under physiologic solution conditions (140 mM PBS, pH 7.4). B. Difference in the observed UCST cloud point (transition temperature), from the predicted linear combination of each motif’s contribution.

These data present quantitative evidence for the importance of interactions between Arg and aromatic residues. When 100% of Phe, Trp and His are substituted for Tyr, we observe clear phase behavior (Figure 17). These data indicate that interactions between the cationic Arg and ringed Trp, Tyr, Phe and His groups are important thermodynamic driving forces for phase separation, although the strength of these interactions is side group dependent with Trp > Tyr > Phe >>> His. We do not believe amino acid substitutions dramatically affect chain conformations supported by various literature reports[59, 76]. A structural investigation with circular dichroism (CD) spectrophotometry of the effect of amino acid chain substitution on chain disorder indicates that the Gly & Pro rich backbone maintains chain disorder as expected [127, 128] (Figure 20).
Figure 20: Analysis of secondary structure with circular dichroism (CD) spectroscopy related. CD spectra of various RIDPs lack a defined secondary structure curve shape, characteristic of other IDP and other repetitive protein polymers. Data collected at 50°C (soluble chains) at 5 µM in 5 mM PBS, pH 7.4. Error bars indicate standard deviation of three sequential runs.

Additionally, small substitutions of Arg with Asp also eliminate phase behavior in buffered conditions while the reverse substitutions surprisingly increase the UCST cloud point at a fixed concentration (Figure 21). Increasing the positive charge of the protein dramatically increases UCST cloud point and broadens the area of the transition. Increasing negative charge, eliminates observable phase behavior under physiologic conditions. This underlines the importance of charge balance as a requirement for observable liquid-liquid phase separation. Naturally occurring IDPs have close to neutral charge balance, with over 85% of natural IDPs from our screen existing between ±5% charge balance.
In addition to chain chemistry manipulation, we also produced proteins with MWs between ~17 kDa and ~70kDa; this range covers 75% of the IDPRs in the proteomic analysis. Our results indicate that MW exhibits at least as large an effect on UCST cloud point as amino acid substitution (Figure 22A). We observed that MW scaling within this range (166-646 amino acids) can be approximated with a natural logarithm line of fit (Figure 22B).
In addition to the aforementioned variables, there are a number of parameters that exhibit a measurable effect on UCST phase behavior but with less effect than the aforementioned variables. Specifically, non-charged polar substitutions, ratio of Gly/Pro, repeating polypeptide syntax/ordering, solution salt content, pH (in the absence of His) and identity of the negatively charged amino acid (Glu vs Asp) all exhibit 1-2 orders of magnitude less change than MW, aromatic:aliphatic amino acid ratio, Arg content and charge balance.

First, let use discuss the syntax or the ordering of the repeat unit. Initially in this project, I was inspired by the diversity of synthetic polymers that display UCST phase behavior and noticed that many of them resemble the side chain architecture found in Ser, Gln, Asn and Thr – polar but non-charged side groups that would chemically...
suggest that they would ideal for H-bond formation with water or one another. Thus, I made a systematic library that took into account every possible permutation of substitutions to our design scaffold of (Gly-Arg-Gly-Asp-Ser-Pro-Tyr-Ser)-40 in the underlined positions, anticipating context dependent behavior of compositionally identical amino acids. Substitutions of these three amino acids, do not eliminate the phase behavior of the polypeptide (Figure 23). A back of the envelope calculation suggests that this particular library contains sequences that will transition at 37°C between concentrations of 1.5 mM and 5 nM.

![Graph](image)

**Figure 23:** Primary amino acid sequence and syntax influence observed phase behavior in Tyr containing RIDPs. Changing placement of amino acid sequence on ~37kDa polymers changes the observed Tt by over 50°C.

This series presented a conundrum to us. Within the context of these systematic substitutions we were observing inconsistent hydrophobicity of Asn, Ser, Gln within the context of the presented polypeptides. Specifically, if one maintains the 5th amino acid
position in the repeat to one of Asn, Ser, Gln and examine how Tt changes with substitutions to the 8th positions, a clear hydrophobic order appears where Ser is more hydrophobic than Gln which in turn is more hydrophobic than Asn (Ser > Gln > Asn) (Figure 24A). However, this result is immediately contradicted when one maintains the 8th position and changes the 5th position (Figure 24B). In this series, we observed a consistent Ser > Asn > Gln hydrophobic ordering. This inconsistency in hydrophobic nature depending on amino acid location, implicates placement in understanding overall hydration and the resulting observed phase behavior.

![Figure 24: Inconsistent ordering in the hydrophobicity of non-charged polar residues in Tyr containing RIDPs. Changing 8th position with fixed 5th position of Ser in ~37kDa polymers changes the observed Tt by ~30°C and suggests a Ser > Gln > Asn hydrophobic ordering. Changing 5th position with fixed 8th position of Ser in ~37kDa polymers changes the observed Tt by ~30°C and suggests a different hydrophobic ordering of Ser > Asn > Gln.](image)

To further understand this phenomenon, we studied polymer pairs that contain identical sequence composition (same number of each amino acid) and yet they have
different observed transition temperatures. [QN]-40 and [NQ]-40 have observed ΔTt of 20°C (Figure 25A), [SN]-40 and [NS]-40 have observed ΔTt of 15°C, and [QS]-40 and [SQ]-40 have observed ΔTt of 10°C. Graciously, Tyler Harmon of Rohit Pappu’s group at Washington University in St. Louis provided some key insight into this observed trend. The simulation was performed on 6 repeats of [QN] and [NQ] with each amino acid side group given an equation that models its water solubility as a function of temperature. A shift in Rg has been correlated with a change in overall solubility and is therefore used to indicate phase behavior of the system[129]. A sharp change in Rg would therefore be predictive of a phase transition in this system. Qualitatively, you can see that [NQ]-6, the more hydrophobic of the two, has a more Q dense core below the transition temperature of this chain (Figure 25B).

Figure 25: Simulations of short motifs of [QN] and [NQ]. Thus, aggregation propensity of this amino acid defines the overall phase behavior of the polypeptide. A. Partial phase diagram of [QN]-40 and [NQ]-40 suggest that syntax
is critical to the phase diagram of repetitive protein polypeptides. B. Coarse-grain simulations of [QN]-6 and [NQ]-6 suggest that the 8th position amino acid qualitatively forms the core of the aggregate (black arrows). C. Representative image of simulations of longer polypeptide sequences ([NQ]-20) form a “beads on a string” architecture, consistent with simulations of polyampholyte synthetic polymers.

It is well known that polyQ polymers have a more stable aggregation than polyN[130]. Therefore, our working hypothesis is that the 8th position of the repeat unit has more solvent accessibility compared to the relatively steric hindered position 5 next to Pro and therefore can dictate the overall hydration of the chain from this position. In this particular case, we observe that [NQ]-40 has a more hydrophobic nature, because the accessibility of the more hydrophobic Gln residue. This would also explain why polymers with Ser in the final position on average appear to behave more hydrophobically as it is known that the hydroxyl-hydroxyl group interaction is stronger than amine-amine[131]. However, these are just preliminary inferences based on our data and extensive simulation of longer length polypeptides and other pairs of RIDPs exhibiting similar syntax dependent phase behavior (i.e. [SN]-40 and [NS]-40).

Preliminary simulation of polypeptides of a longer chain length (20 vs 6 repeat motifs) suggest that the polypeptides are forming a “beads on a string” type architecture with patches of polypeptide associations along the length of the chain (Figure 25C). This architecture is typical of polyampholyte synthetic polymers (polymer consisting of only positive and negative charge side chains but overall neutral charge), which are similar in construction to our RIDPs.
I had great interest in the beginning of this project to understand the solvents that affect the UCST phase diagram. At the time, many UCST synthetic polymers would see their temperature sensitive phase behavior disappear in buffered or salt solutions. This lacuna prompted us to develop a UCST system that would be durable to many physiological conditions. Initially, I was focused on the effect of urea on the phase boundary of RIDPs because of its potential to dramatically improve the speed at which proteins could be purified and theoretically understand what the limitations would be on proteins that could be purified with transition cycling. We observed a linear dependence of the concentration of urea on the observed T₁ of the polypeptide which enabled us to utilize urea to solubilize the cell lysate pellet at elevated temperatures that contained our protein of interest. Subsequent dialysis into ice-cold distilled water, replaced the good solvent, urea, with water, a poor solvent for RIDPs.
Figure 26: Effect of urea concentration on the UCST transition temperature of RIDPs of various molecular weight. Red lines are linear fits to the data. Increasing the concentration of urea by 1 mole decreases the observed UCST by ~22°C. This effect is consistent for all molecular weights (20-100 repeats). RIDP with repeat sequence QYPSDGRG was used at 50µM in 140mM PBS with varying amounts of added urea.

Similarly, we characterized the effect of various amounts of salt on the transition temperature of RIDPs. We observed an inverted parabolic-like shape of transition temperature of the RIDP in the presence of increasing amounts of NaCl (Figure 27B). An interesting effect that is observed is a change in the sharpness of the phase separation behavior as additional salt is added. Specifically, the transition is broader in pure water than in >50 mM salt solution (Figure 27A). It is possible that in the absence of salt, the charge repulsion between the chains due to the zwitterionic nature of the polymer are contributing a “broadness” to the phase separation (or a retardation to the nucleation and growth kinetics) that is not seen when the charges are screened at higher salt.
concentrations. The shape of this curve suggests that there are multiple interactions occurring between chains to drive phase separation from the solvent. Our working hypothesis is that these RIDPs are durable to a wide range of salt conditions because the sequence contains multiple interactions that may drive phase separation (cation-pi interactions, charge-charge interactions, polar non-charged H-bonding) where as in the synthetic polymer examples, the side chains are typically from a single category and therefore only capable of a single type of interaction. These effects of urea and salt presented here are typical of a number of other polypeptides we screened with similar methods.

Figure 27: A. Salt concentration affects but does not eliminate UCST phase behavior. B. UCST phase behavior in H₂O decreases by 25°C upon addition of NaCl to a minimum around 500 mM. UCST cloud point then begins to increase beyond this minimum point. Data taken at constant pH 7.4 and [Q]-40 concentration of 10 µM.

One of our systematic substitutions replaced the aromatic Tyr residue with His, a pH responsive residue under physiological conditions (pKa ~ 6.2). Thus, we were
interested to observe the effect of changing the protonation state of this side group on
the observed cloud point. As seen from previous work, polypeptides that had a
substituted His for Tyr saw a large decrease in the observed Tᵢ at pH 7.4 in 140mM PBS
(60-70°C). This result can be understood from an entropic and an enthalpic perspective.
Removal of the hydroxyl group reduces the polymer-polymer interactions, decreasing
the predicted Tᵢ. Additionally, the imidazole ring is a more hydrophilic moiety in
comparison to a benzene ring, which decreases the negative entropic contribution of the
hydrophobic term, thereby decreasing the Tᵢ. However, despite this substitution the
transition curves are still quite sharp, reversible, concentration dependent and molecular
weight dependent.

Measuring the Tᵢ of two polymers of same molecular weight, concentration and
composition, save the Tyr to His mutation, revealed two interesting developments. First,
is that the decrease in pH increases the Tᵢ by over 30°C between pH 7.4-5. The
midpoint of this shift occurs directly around the predicted isoelectric point of the
imidazole group of His (pH 6.2). This result is completely unexpected. The increase in
positive charge should induce large polymer-polymer repulsion decreasing the Tᵢ.
However, we see the exact opposite effect.
Figure 28: Decreasing pH of solution results in dramatic increase in Tt of [H]-60 (normalized data to pH 7.8). A. Turbidity curves of [H]-60 55kDa polypeptides compared to non-His ([S]-60) control. B. Change in observed transition temperature of [H]-60 and [S]-60 as a function of solution pH. All data collected at 2 mg ml\(^{-1}\) via UV-Vis turbidity @ 350nm for both polypeptides.

The other observation of the His containing polypeptides is the broadening of the transition as pH is decreased. The broadening of the transition was similar to the broadening observed in water with other RIDPs tested. Therefore, one could hypothesize that this broadening is a kinetic component introduced at low pH when the competing forces of charge repulsion and some unidentified side chain interaction collide.

We also produced and tested agnostically non-repetitive, but compositionally identical versions of [S]-20 and observed minimal effects of scrambling the amino acid sequence. We approached this more granularly, randomizing progressively smaller “windows” of the polymer sequence controlling for “non-repetitiveness” of the protein using a sliding window quantification technique and observed minimal effect at
progressively smaller windows (data not shown). This result may be significant in the practical working terms of creating engineered IDPs as now the gene can be created to be non-repetitive and thus more amenable to PCR amplification and high throughput DNA synthesis.

Figure 29: Partial binodal phase boundaries of agnostically non-repetitive but compositionally identical versions of [S]-20. Data collected in 140 mM PBS @ pH 7.4.

2.3.3 RIDPs create dense phase separated droplets at saturation concentrations mediated by amino acid composition

With the understanding that $C_{sat}$ and UCST cloud point in the dilute regime can be modified drastically by amino acid substitution we were interested in investigating the dense phase of the polypeptide solution. Polypeptides [S]-20 and [Q]-20 express extraordinarily well for recombinant proteins (500 mg l$^{-1}$) and thus we were able to purify over 1 gram of material to measure the UCST cloud point behavior directly at very high-
volume fractions ($\phi > 0.1$). These experiments were performed in microcuvettes on a temperature gradient plate to assess UCST cloud point while microscopically visualizing the phase separation event (Figure 30A). Using line scans derived from this image one can determine the $T_d/T_{ph}$ of the solution (Figure 30B). The dashed black lines represent tangent lines for the high temperature baseline and increase in light scattering at lower temperatures. These two lines intersect at $T_{ph}$, as indicated by the vertical green line.

Figure 30: Collection of phase diagrams using a temperature gradient device. A. Representative dark-field image of [Q]-20 solutions on a temperature gradient device. The transition temperatures of the reference solutions (red and blue lines) and the 20 mg ml$^{-1}$ [Q]-20 solution (green line) are indicated by the horizontal colored lines. The dashed vertical magenta line along the 20 mg ml$^{-1}$ capillary tube illustrated the region of the image used to measure the line scan. B. Line scan of normalized light scattering intensity versus temperature for the 20 mg ml$^{-1}$ [Q]-20 capillary shown in A.

These experiments produce binodal phase boundaries similar to turbidity measurements and demonstrate that a $\sim$25 Celsius difference between the two binodal lines of [S]-20 and [Q]-20 is maintained over the entire concentration range. This corresponds to an increase in polypeptide volume fraction in the dense phase ($\phi''$) from
\[ \phi'' = 0.4 \text{ for } [Q]-20 \text{ to } \phi'' = 0.55 \text{ for } [S]-20 \text{ at an isotherm of 310K}. \] 

These data in combination with our ability to easily purify RIDPs with phase separation indicate that RIDPs form highly exclusionary droplets in vitro consisting of a very dense phase at physiological solution, temperature and pH conditions \( \phi'' > 0.4 \).

Figure 31: Binodal phase lines of [S]-20 and [Q]-20 using multiple data points from temperature gradient device. A three-piece fit was utilized to fit three regimes that roughly correspond to the dilute, overlap, and semi-dilute regimes of the polypeptide phase diagram.

This phase diagram also confirms our hypothesis that chemical modifications to the chain by amino acid substitution not only affect behavior in the dilute regime but change the critical point and the density of the sediment \( \phi'' \).
Temperature gradient microfluidics also confirm our suspicions surrounding the growth mechanism of the dense phase. Investigations using droplet microfluidics previous suggested that growth mechanic by coalescence as opposed to Ostwald ripening. Using kinetic analysis from the temperature gradient system we are able to confirm this hypothesis. Using a stretched exponential decay model of the change in scattering intensity over time for a given line scan (Figure 33) one can determine rate constants that occur at various temperature. From these rate constants one can determine the primary driving force of the observed phase behavior below the binodal line. In the regime of spinodal decomposition one could the rate constant to an Arrhenius model (solid black line over the blue data points) to determine the energy of activation of this even. We calculated a value that is slightly positive activation energy ($E_A \approx 2$ kcal mol$^{-1}$) that was smaller than the dissociation energy of typical water–water hydrogen bonds[132, 133]. This analysis suggests two conclusions. First, the growth mechanism of [Q]-20 appears to grow only by coalescence and grown in buffered, non-crowded conditions in the observable temperature range. Secondly, analysis on the temperature gradient device is capable of detecting a metastable regime between the binodal and spinodal decomposition regime of the phase diagram. However, in all of our experimentation we were unable to visualize the presence of this regime suggesting that this regime is either very narrow and therefore able to escape detection or that the phase separation of the RIDP in the capillary creates a thermophoretic effect where the protein travels down a temperature gradient and is therefore never “metastable” on the gradient.
Figure 32: Kohlrausch-Williams-Watts (KWW) stretched exponential decay of a temperature gradient line scan to determine rate constants of phase separation. Example data captured with [Q]-20 @ 25 mg ml\(^{-1}\).
Early on in my thesis work our assumption was that UCST phase behavior is also mediated primarily by entropic or hydrophobic forces. Thus, we were unable to fit data of UCST polypeptides available in 2013 to a single function. However, despite these failings, I attempted to develop a systematic hydrophobic library maintaining parameters that we assumed to be critical – molecular weight, importance of Arg and charge balance. This library included the repeat motifs (Gly-Arg-Gly-Asp-Ser-Ala-Tyr-Gly), (Gly-Arg-Gly-Asp-Ser-Cys-Tyr-Gly), (Gly-Arg-Gly-Asp-Ser-Ala-His-Gly), (Gly-Arg-Gly-Asp-Ser-Pro-Trp-
Gly), (Gly-Arg-Gly-Asp-Ser-Pro-Ala-Leu), and (Gly-Arg-Gly-Asp-Ser-Val-Iso-Gly) where the final three amino acids in the repeat unit were modified to produce proteins of variable hydrophobicity but without controlling for any other factor that we now know are critical to the observed cloud point. From this design class only (Gly-Arg-Gly-Asp-Ser-Ala-Tyr-Gly) and (Gly-Arg-Gly-Asp-Ser-Ala-His-Gly) expressed in E. Coli and were characterized to have similar phase behavior to proteins we had previously produced in the lab (data not shown). This failure prompted the experiments performed just by substituting the polar non-charged amino acids on the polypeptide backbone to begin to unpick the phase behavior with much smaller perturbations.

This initial hypothesis surrounding hydrophobicity was centric to elastin like polypeptide (ELP) design performed previously in our laboratory. There are many hydrophobicity scales that have been published, however, it was clear early on that these scales were not sufficient to capture the unique phase behavior of UCST polypeptides. Even a scale developed specifically in protein polymers (Urry) to describe LCST behavior does not capture half of the data contained in our new relativistic scale for UCST polypeptides (Figure 34).
Figure 34: Normalized Kyte & Doolittle linear fit to the hydropathy of RIDPs ($R^2 = 0.23$). Normalized Urry hydrophobicity scale linear fit to the hydropathy of RIDPs ($R^2 = 0.49$). Normalized Wimley and White hydrophobicity scale linear fit to the hydropathy of RLPs ($R^2 = 0.37$). Value of hydrophobicity increases moving from 0 to 1 for the X-axis (RIDP relative scale) and 1 to 0 for the other hydrophobicity scales.

One observation made during the experiments where we made many substitutions to the non-charged polar residues of our reference sequence was the appearance of unstable phase behavior in proteins that contained an Asn in the 8th position of the repeating unit. We studied this phenomenon by specifically progressing through multiple phase separation cycles while monitoring turbidity and the change in molecular weight via SDS-PAGE ensuring that upon reheating the polypeptide solution was fully resuspended and not sedimented beyond the light path of the instrument. The assumption was an enzyme had co-purified with the protein that was dramatically reducing the molecular weight of the main chain and thus reducing the observed cloud point. However, we observed just the opposite trend. Upon repeated transition cycles, [N]-40 UCST cloud point progressively dropped but its molecular weight appeared to increase. [NQ]-40, which has just a single additional side chain carbon in the 8th position,
shows <1°C difference and no change in the gel mobility. Although never officially confirmed, we suspected a level of self-catalysis caused by the repeating motif Asn-Gly in polypeptide [N], which has been shown to have autocatalytic activity in other systems [134]. This would explain the apparent increase in molecular weight if the micro-architecture of the protein was folding back upon itself into a quasi-circular shape.

Figure 35: Lowering UCST cloud point of [N]-40 upon repeated cycles of interrogation despite an apparent increase in the molecular weight by SDS-PAGE.
Figure 36: Idealized chain conformation of [N] and [NQ] repeating unit demonstrates the potential strained nature of Asn towards the backbone of the chain, particularly with a Gly, C-terminal to Asn.

Prior to our collaboration with the Cremer group at Pennsylvania State University, we struggled in alternative experiments to collect conclusive evidence of the density of the sediment or the \( \phi'' \) side of the phase boundary. One such experiment was co-incubation of various polypeptides with fluorescently tagged dextrans of a known molecular weight and thus known molecular size. We incubated the protein and dextran together for 10 min at temperatures that the RIDP would be soluble, cooled to below the \( T_\text{t} \) and mounted the solutions onto glass sides. If there was any interaction or preference for the droplet of dextran molecules or if the “mesh size” of the phase separated network
was large enough to permit entry of the dextran molecule then we would be able to calculate a partition coefficient between the dilute and dense phase of the solution. However, we never observed any measurable partitioning of the dextran regardless of RIDP chemical identity. Thus, we conclude that RIDP droplets are exclusionary compared to other naturally occurring IDPs and explains why purification from E. Coli lysate is quite easy.

Figure 37: Quantification of dextran uptake during phase separation of RIDPs. A. Fluorescent microscopy images of phase separated droplets in the presence of dextran molecules of different molecular weight (10/40kDa) labeled with AlexaFluor 488 (green) fluorophore. Inside the phase separated space (dark circles), there is very little sequestration of the dextran molecules as a function of dextran molecular weight or polypeptide sequence. Scale bar is 20 µm. B. Quantification of fluorescent signal between the area inside of phase separated droplets and outside.
2.4 Conclusions

There are number of conclusions from this first systematic exploration of UCST phase behavior in protein polypeptides. First, as suspected by the synthetic polymer literature a combination of enthalpic and entropic forces are necessary to encode UCST phase behavior at the sequence level. In my experience cloning and purifying ~100 unique protein motifs the five most important factors (in terms of % mole change) are:

1. Molecular weight of the chain – Entropic
2. Volume fraction of the solution – Entropic
3. Relative number of aromatic residues compared to aliphatic residues – Enthalpic
4. Relative amount of Arg compared to Lys – Enthalpic
5. Overall charge balance of the chain – Enthalpic/Entropic

In addition to these primary factors there were a number of other observations made throughout that anecdotally should be included into future work with RIDPs or similar repetitive protein domains.

First, this particular mutational framework is quite effective at encoding UCST phase behavior into polypeptides. In this study alone, I identified proteins (capable of being purified from bacterial lysate) that are predicted to have a dilute phase regime of their diagram occur under physiological conditions between 0.2 nM and 4-8 mM with obvious mutations that could push these limits even further with a different downstream purification scheme. This alone is a massive step forward for the field as this library consists of the largest number of individual molecules purified with partial phase diagram information.
Secondly, although these individual repeating motifs are displayed here as discrete sequences, we have demonstrated that blending of two or more of these sequences leads to phase behavior that is a linear combination of the foundational units. This suggests that these points are just that, points on a landscape of phase behavior that can be tuned continuously between each of the discrete examples presented here. This will allow for precise programming of a protein with a specific molecular weight/phase diagram information a priori.

Thirdly, I believe that there is an over-emphasis of protein syntax on the observed phase separation in vitro. The field is often obsessive over protein sequences – the order of amino acids on a chain - but in my library this effect was not dominant. I have found that although context can be important (particularly around structure breaking Pro residues), overall composition is likely a much more important driving force of UCST phase behavior than the local environment of each individual amino acid. In a repetitive protein polymer these effects are amplified but are likely less important for the largely non-repetitive IDPs found in nature.

Finally, the dense phases formed by these proteins after crossing the binodal are remarkably dense at physiologic temperatures compared to any other reported naturally occurring IDP or synthetic polymers. This appears to lead to dense phases that are also very exclusionary to other biomolecules of interest – hence our ability to rapidly purify these proteins from bacterial cell lysate.
3. Designer membrane-less organelles using engineered intrinsically disordered proteins with non-canonical functionality.

3.1 Introduction

Intrinsically disordered proteins (IDPs) are receiving significant recognition for their role in various biological (dys)functions[62, 67]. A subset of IDPs, termed biological condensates, physically separate themselves from the cytoplasm to control the accessibility of a variety of macromolecules[60, 135]. While our ability to detect protein disorder has advanced rapidly thanks to sophisticated statistical methods, the ability to predict phase separation and subsequently engineer new functions of phase behavior in cells has significantly lagged behind[136].

3.1.1 Literature review

The prediction of phase separation is non-trivial as the IDP literature cites numerous variables that influence phase separation. From Chapter 2 and various other reports they involve 1. amino acid composition and amino acid patterning of the primary protein sequence[137-139] 2. heterotypic interactions with RNA or other macromolecules[72, 140] and 3. solvent quality [59, 76, 141]. There are many studies that note the challenge of predicting IDP phase behavior, but few have sought to tackle the problem directly[83]. Most mutation strategies regarding IDPs opt for sweeping chemical substitutions or dramatic protein architecture remodeling resulting in a binary result – the presence or absence of phase behavior. These strategies linger from the era of structural protein characterization, failing to properly control for known
thermodynamic principles of phase separation[142]. Thus, this continued inability to rationally tune the phase separation of disordered proteins in vivo undermines the conclusions that link phase separation to their biological utility.

As biologists have begun to study phase separation in living cells, polymer scientists have simultaneously been studying phase separation in protein polymer in vitro[78, 123]. Although protein polymers are far more repetitive than their biologic counterparts, they have been shown to demonstrate the basic principles of phase separation inside of cells [143-146]. Armed with the repetitive IDP library from Chapter 2, we sought to apply these materials to intracellular puncta design.

3.1.2 Approach

Applying the main chain factors observed to affects saturation concentration (C_{sat}) in vitro – MW and aromatic:aliphatic ratio – we sought to affect the dynamics of droplet assembly in living cells. In order to visualize localization of the RIDPs of interest within bacterial cells, each protein was genetically fused to a super folder version of GFP (sfGFP) (Figure 38).
3.2 Methods

3.2.1 Gene synthesis

Each octapeptide amino acid motif was propagated twenty times in silico. This repetitive amino acid sequence was fed into an algorithm that creates an optimally non-repetitive DNA template from a repetitive protein gene. This 20-mer repeat gene was then ordered from IDT with Gibson assembly overhangs for easy insertion into our modified pet24+ vector[120]. To increase the number of total repeats of the gene, we performed iterative cloning steps of Recursive Directional Ligation by Plasmid Reconstruction adding an additional twenty repeats during each step. This method has been described previously[121]. Transformations were performed into the desired E. Coli cell line – BL21 (DE3) for recombinant expression and single plasmid confocal
experiments and a modified BL21 (DE3) cell line termed KRX by Promega that contains a mutated LacZ gene for enzymatic experimentation.

In experiments with dual expression, genes were inserted into the pBAD33.1 vector by cutting custom pet24+ vector and pBAD33.1 cut with HindIII and XbaI. Gel purification was used to isolate the gene of interest from the housing pet24+ vector, which was then ligated into the similarly cut pBAD33.1 vector. Co-transformation was performed with ~1 ng final concentration of each plasmid on kanamycin/chloramphenicol dual selection plates.

3.2.2 Protein purification

Individual liquid cultures of BL21 E. Coli strains each harboring our gene of interest from Table S1 or Table S2 were inoculated into 5 ml of Terrific Broth (TB) medium from frozen glycerol stocks and grown to confluence overnight (16-18 hrs). Cultures were then inoculated at a 1:200 dilution in 1L TB media supplemented with 45 μg ml⁻¹ kanamycin. Cells were grown at 37°C in a shaking incubator (~200 r.p.m.) for 9 hr, at which time protein expression was induced by the addition of 500 μM IPTG (final concentration). Cells were then incubated at 37 °C (shaking at ~200 r.p.m.) for an additional 18 hr. Protein was then purified from the insoluble cell suspension fraction. In brief, cell pellets were isolated by centrifuging cultures at 3500 RCF and resuspending in 20 mL of milli-Q water. Cells were then lysed by sonicating the cell solutions for 2 minutes, with 10 seconds of pulsing followed by 40 seconds of rest on ice (Misonix; Farmingdale, NY)
Centrifuging each lysate suspension at 20,000 RCF for 20 minutes results in a soluble and insoluble fraction. The supernatant was discarded with the insoluble fraction resuspended in an approximately equal volume of 8M urea + 140 mM PBS (~6-8 mL). For proteins with a fluorescent fusion tag, the insoluble fraction was resuspended in 3X insoluble pellet volume at a final concentration of urea of ~2M to prevent protein misfolding. This suspension was heated for 10 min in a 37°C water bath and then centrifuged at 20,000 RCF for 20 minutes. The supernatant was collected from this suspension and dialyzed in a 10 kDa membrane (SnakeSkinTM, Thermo Fischer Scientific) against a 1:200 milli-Q water solution at 4°C. The dialysis water was changed twice over a 48-hour period. From inside the dialysis bag, both insoluble and soluble components were collected and centrifuged at 3500 RCF for 10 minutes and 4°C. The supernatant was removed and the remaining insoluble pellet containing the protein of interest was lyophilized for a minimum of three days to remove all water from the pellet.

Protein purity was characterized by 4–20% gradient tris-HCl (Biorad, Hercules, CA) sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and staining with either 0.5 M copper chloride or SimplyBlueTM SafeStain (Thermo Fischer Scientific). Protein yield was determined by weight after lyophilization.

3.2.3 Characterization of phase separation

3.2.3.1 Temperature dependent UV-vis spectrophotometry

Cloud point transition temperatures (Tc) were determined via temperature-controlled spectrophotometry using a Cary 300 (Agilent Technologies). Samples
containing various concentrations of protein in 140 mM PBS were cooled at 1°C min⁻¹ while the absorbance at \( \lambda = 350 \) nm was recorded every 1°C. Absorbance was normalized to the absorbance at the highest temperature point collected, corresponding to the most soluble point during a given experiment. The cloud point was determined as the maximum in the first derivative of the absorbance as a function of temperature. Transition temperature was calculated at the point of minimum slope. Saturation concentration was calculated using the natural logarithm fit line created from a minimum of three volume fractions.

3.2.3.2 Whole cell fluorescent intensity measurements

Cells were grown overnight in 5 ml of TB media from glycerol stocks. In conjunction to fluorescent or confocal imaging, cells were analyzed for total sfGFP fluorescence and OD600. Briefly, 50 µl of cell culture at various time points was resuspended in 1 ml of 140 mM PBS. Using a combination of a UV-Vis spectrophotometry signal from a NanoDrop 1000 (Thermo Fisher Scientific, Waltham, MA) and fluorescent spectra from a NanoDrop 3300 (Thermo Fisher Scientific, Waltham, MA), we calculated the relative ratio of sfGFP fluorescence normalized to cell density. Using this information in conjunction with imaging analysis, we were able to determine the intracellular saturation concentration normalized to cell density.

3.2.3.3 Temperature controlled fluorescent microscopy of protocell droplets of E. Coli.

E. Coli were collected on a glass microscope slide and cooled using a precision Peltier heating and cooling stage (Linkam LTS120) equipped with a temperature control
unit (Linkam PE95). The spatial distribution of sfGFP was characterized via fluorescence microscopy using an upright Zeiss Axio Imager D2 microscope with a 20× objective and the appropriate filter set. Similarly, intracellular patterning of RIDP-superfolder GFP over time was characterized via fluorescence microscopy using an upright Zeiss Axio Imager D2 microscope with a 20× objective and the appropriate filter set (ex 470/40, em 525/50). Cell fluorescence was calculated using ImageJ software. Temperature ramps began at various temperatures but always were set to a constant speed of 5°C min⁻¹.

3.2.3.4 Transient Transfection of [S]-20-sfGFP in HEK293 Cells

[S]-20-sfGFP was extracted from the pet24(+) vector using polymerase chain reaction (PCR). Briefly, the forward and reverse primers were resuspended with 1 ng of pet24(+) plasmid containing [S]-20-sfGFP gene fusion. Using a PCR cycle of [98°C-1min, 65°C-30sec, 72°C-2min]x30, followed by gel purification, the gene was finally constructed with Gibson assembly. pcDNA5 vector containing [S]-20-sfGFP was transfected into HEK293 cells according to manufacturer instructions (Expi293 Expression System, Thermo Fischer Scientific, Waltham, MA). Cells were spun down at 500 RCF for 10 min at room temperature on day 5 of transient transfection and resuspended in 140 mM PBS for imaging.

3.2.3.5 Confocal imaging of RIDP-sfGFP fusions for puncta formation and colocalization

Cells were prepared as follows. A tube containing 5 ml of TB media was inoculated overnight with protein of choice from bacterial glycerol stock. After 16 hr of growth, induction with 1 mM IPTG and 2% L-rhamnose (Sigma-Aldrich, St. Louis, MO)
was added each flask of interest. Samples were collected at the indicated time points and prepared for imaging as follows. 50 µl of cell suspension was pelleted under 20,000 RCF for 1 min at room temperature. Cells were resuspended to OD600 = 0.15 at 1 cm path length. 50 µl of resuspended bacterial cells were transferred to a 384-well plated with #1.5 glass bottom (Cellvis). There was a 10 min equilibration period to the incubation chamber prior to each time point data collection.

Images were collected at different time points with a 63X oil-immersion objective on a Zeiss 710 inverted confocal with temperature-controlled incubation (Car Zeiss AG, Oberkochen, Germany). sfGFP fluorescent was detected with a 488 nm excitation laser and 488/594 emission filter. Data was primarily taken at 25°C unless otherwise noted. All fluorescent quantification and cell portioning analysis was performed in ImageJ.

In colocalization experiments, cells were grown overnight from glycerol stock in dual antibiotic media containing 45 µg/ml kanamycin and 25 µg/ml chloramphenicol (final concentration). After 16-18 hours, pet24(+) expression was induced with 1mM concentration of IPTG (final concentration). After 24 hours of IPTG induction, media was replaced with 5 ml of TB supplemented with 1mM IPTG and 2% arabinose (final concentration) (Sigma Aldrich, St. Louis, MO). After 9 hours of induction, cells were prepared for confocal imaging by spinning down 50 µl of culture at room temperature and resuspended in 140 mM PBS to OD600 = 0.15 at 1 cm path length. All imaging details remain the same except that mNeonGreen/sfGFP detection was performed with 488 nm excitation laser and 488/594 emission filter and mRuby3 detection with 561 nm excitation laser and 488/561 emission filter.
3.2.3.6 Spinning disc confocal imaging of Lac-Z alpha-peptide-RIDP gene fusions for localization and quantification of enzymatic activity

Cells were prepared as follows. A tube containing 5 ml of TB media was inoculated overnight with protein of choice from bacterial glycerol stock. After 16 hr of growth, induction with 1 mM IPTG and 2% L-rhamnose (Sigma-Aldrich, St. Louis, MO) was added each flask of interest. ~24 hours later, 50 µl of cell suspension was pelleted under 20,000 RCF for 1 min at room temperature. Cells were resuspended in 140 mM PBS to OD600 = 0.15 at 1 cm path length. 50 µl of sample were added to Culture-Insert 4 Well (1.5 coverslip, Ibidi, Madison, WI) petri dishes and allowed to incubate at room temperature for 10 minutes. After incubation, 2 µl of 1 mg ml⁻¹ FDG resuspended in 98% water, 1% DMSO and 1% EtOH was added. Imaging began immediately (within 20 seconds) and images were captured every minute for 30 minutes total.

Imaging was performed on an Andor Dragonfly Spinning Disk 500 series confocal on a LeicaDMi8 microscope stand (Oxford Instruments, Abingdon, UK) with a 63× water immersion objective and equipped with a Zyla 4.2 series camera. Converted FDG was detected with a 488 nm excitation laser and 525/50 nm emission filter and mRuby3 fluorescence with a 561 nm excitation laser and 600/50 nm emission filter.

3.2.3.7 Fluorescent spectroscopy for determining $K_m$, $V_{max}$ and $k_{cat}$

Liquid cultures of KRX E. Coli containing plasmid of interest were grown from glycerol stocks overnight (16-18 hours). Cells were then induced with 1 mM IPTG and 2% L-rhamnose (Sigma-Aldrich, St. Louis, MO) for 24 hours. Cells were pelleted and resuspended at OD600 = ~0.15 in 140 mM PBS. Various concentrations of Fluorescein
Di-β-Galactopyranoside (FDG) were added while monitoring fluorescent intensity at 520 nm using NanoDrop 3300 (Thermo Fisher Scientific, Waltham, MA). The same instrument was used to calculate the fluorescent intensity of mRuby3 as a relativistic measure of expression level of the various alpha peptide fusions. Plotting the observed fluorescent intensity at different times provides a surrogate measure of the rate of hydrolysis at various concentrations of the substrate (V₀). These rates were then converted into typical Lineweaver-Burk conventions to determine Vₘₐₓ and Kₘ. For consistency in units, [FDG] was converted into fluorescent intensity using a fluorescein standard curve of y = 185919*[FDG in mg] + 1045. This conversion assumes that converted FDG into fluorescein has similar fluorescent intensity profile to free fluorescein dye.

3.2.3.8 Image quantification and statistical analysis

Statistical analysis was performed on experiments that determined the intracellular fluorescent intensity of RIDP-sfGFP at various points post-IPTG induction. For determining the intracellular saturation concentration, whole cell fluorescence normalized to cellular density (OD600) on three independent samples was calculated alongside of imaging of their intracellular architecture. Upon first observation of phase separation in E. Coli in more than 50% of cells within a microscopic field of view, this normalized cell density was recorded as the saturation concentration. Data is normalized to data collected for [S]-40 as a reference point. Error bars represent
propagated standard error of the mean of three separate samples from the same original cell suspension.

With the microscope images collected with confocal microscopy at various time points, we isolated the soluble and puncta fractions within the cells at various points in time via analysis in ImageJ. Puncta consistently create pixels dense enough to saturate the detector while simultaneously observing the rest of the cell. Thus, by thresholding around the upper 2% of total pixel intensities, one can easily partition this section from the remaining cell cytoplasm. Using this constant thresholding between timepoints in each experimental group, we were able to track the total size of these puncta over time with regard to the total size of the cell (puncta + soluble fraction). Error bars of these data are standard errors of the mean of normalized puncta (two-phase) area of three images of different fields of view of the sample overall cell samples.

Given the lack of automated tools for the detection of intracellular phase separation between two images, we calculated the intracellular transition temperatures manually. Similar to the detection of phase separation with UV-vis spectrophotometry, the intracellular transition temperature was determined as the midpoint between a frame that was certainly homogenous and a second frame that was certainly two phases. All transition temperatures were determined in this way, going from a point of solubility to insolubility whether the solution was being heated or cooled. Due to the level of subjectivity of this assessment, sample identifiers were blinded to the analyst and a high number of cells were analyzed in each experiment (n = 30). Data was normalized to the
initial mean fluorescence of the homogeneous cells at a consistent temperature (often 60°C unless otherwise noted). Error bars indicate standard error of the mean.

For quantification of Fluorescein Di-β-D-Galactopyranoside (FDG) relative to the different expression levels of alpha peptide, channels were split between fluorescence from FDG and mRuby3 respectively. Using the particle analysis tool from ImageJ, areas of green fluorescence were isolated from the background. If the mean fluorescence of this area was 5% greater than the background fluorescence (mean fluorescent of the area excluded by the previous particle mask), then this particular particle’s background subtracted green fluorescence was included in the analysis. Particles were excluded if their area was below 0.1 µm². Using the same particle mask, the background subtracted mean fluorescence of mRuby3 was calculated on the other fluorescent channel. We report the ratio of these two channels as a surrogate for enzymatic efficiency. Error bars are standard errors of the mean at each timepoint.

For quantification of FDG inside the cellular space versus outside the cellular space, channels were first split between fluorescence from FDG and mRuby3 respectively. Using the same particle analysis tool from ImageJ, areas of green fluorescence were isolated from the background. If the mean fluorescence of this area was 5% greater than the background fluorescence (mean fluorescent of the area excluded by the previous particle mask), then this particular particle’s background subtracted green fluorescence was included in the analysis. Particles were excluded if their area was below 0.1 µm². Ratio of fluorescent intensity inside of cells versus the extracellular space is the background corrected mean fluorescence of FDG divided by
the background fluorescence. Error bars are standard errors of the mean at each timepoint.

To quantify the amount of colocalization we used the Coloc2 plug-in available through ImageJ software. Using automated thresholding, we report the Mander’s colocalization coefficient which accounts for the intensity to the two channels of interest as described previously[147].

3.3 Results and Discussion

The intracellular droplets comprised of RIDPs proved to be surprisingly predictable, with the ability to dynamically control their cytoplasmic solubility, and therefore their interactivity, with the surrounding environment. Capitalizing on this observation, through bottom-up design, we created intracellular droplets capable of sequestering an enzyme whose efficiency can be genetically encoded into the cellular puncta by modulating the MW of the RIDP.

3.3.1 RIDPs have controlled C_{sat} in eukaryotic and prokaryotic cell lines

Our work with the RIDPs from Chapter 2 described how modifying the aromatic content, Arg content, molecular weight and charge balance can dramatically affect the UCST binodal line. Thus, to convert this information into usable information for cell experiments, we needed to calculate the saturation concentration, defined as the $\phi'$ concentration which the binodal is crossed at an isotherm of 37°C, for various polypeptide sequences. As suspected, the same variables that exhibit large shifts in the
cloud point also exhibit dramatic changes on the observed \( C_{\text{sat}} \). By changing the amino acid sequence we were able to change the \( C_{\text{sat}} \) by over three orders of magnitude (Figure 39) and by changing the molecular weight we were able to exhibit control over \( C_{\text{sat}} \) by over four logs (Figure 40). Thus, we decided to make sfGFP fusions with RIDPs that have various ratios of aromatic and aliphatic content and molecular weight.

Figure 39: The chemical composition is capable of affecting the saturation concentration by 2 orders of magnitude at constant molecular weight (\( C_{\text{sat}} \) @ 37°C = 1-400 µM). This can be visualized by normalizing to the saturation concentration of [S]-40 which is conveniently ~1 µM.
Figure 40: At constant chemical composition it is possible to affect $C_{\text{sat}}$ by over four orders of magnitude by changing molecular weight of the chain ($C_{\text{sat}} @ 37^\circ \text{C} = 1 \text{ nM} – 100 \text{ µM}$). [S]-40 has a $C_{\text{sat}}$ of 1 µM.

The genetic fusion of sfGFP did not eliminate phase behavior but did shift the phase diagram (Figure 41). These data suggests that the larger molecular weight polypeptides are less effected by sfGFP fusion as the observed difference in [S]-40 and [3Y:V]-40 is a consistent 10°C instead of ~20°C for [S]-20.
Figure 41: Comparison of partial binodal phase diagrams of RIDP and RIDP-sfGFP fusions. A. Partial binodal phase boundaries of [S]-40 and [S]-40-sfGFP. B. Partial binodal phase boundaries of [3Y:V]-40 and [3Y:V]-40-sfGFP. C. Partial binodal phase boundaries of [S]-20 and [S]-20-sfGFP. In each experiment, the addition of the sfGFP fusion, lowers the UCST binodal line.

Using confocal microscopy, we were able to observe the formation of intracellular droplets of [S]-20-sfGFP in both transfected HEK293 cells (Figure 42) and E. Coli bacteria.
Interestingly, the formation of microdroplets in HEK293 cells differs greatly from that in protocells. Droplets in HEK293 cells are spatially distributed around the cell as opposed to a single droplet in the protocell. In contrast, typically a single droplet is formed per cell in BL21 bacteria (Table 3). This result suggests differential diffusivity of the polymer chains between prokaryotic and eukaryotic cytoplasmic environment.

Table 3: Number of phase separated domains per E. Coli as a function of induction time (n = 3 images)

<table>
<thead>
<tr>
<th>Time Post Induction (hours)</th>
<th>[S]-20-sfGFP</th>
<th>[S]-40-sfGFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>N/A</td>
<td>1.01 ± 0.02</td>
</tr>
<tr>
<td>6</td>
<td>1.03 ± 0.01</td>
<td>1.03 ± 0.03</td>
</tr>
<tr>
<td>10</td>
<td>1.13 ± 0.01</td>
<td>1.10 ± 0.09</td>
</tr>
</tbody>
</table>
Tracking the sfGFP fluorescence of the bulk cell population normalized to the absorbance at 600 nm (OD600) indicates that intracellular concentration of all proteins is increasing with increased protein induction time (Figure 43).

**Figure 43**: Measurement of total cellular fluorescence at different times post induction. Measurements of the optical turbidity and fluorescent intensity of sfGFP were made and plotted as a function of time. Data collected at 22°C.
Figure 44: Size of intracellular droplets grow with induction time. As concentration increases inside the cell, the soluble concentration outside the droplet is not changing but the size of the intracellular droplets is growing relative to the total cell area. Error bars represent standard error of the mean.

Figure 45: Measurement of the cellular fluorescence at different locations within the cell. Digital partitions were made between the dense phase separated area of the cell and soluble cytoplasmic space using ImageJ. The mean fluorescent intensity of the total cell fluorescent intensity (solid line) and cytoplasmic fluorescent intensity (dotted line) are plotted as a function of time after IPTG induction. [S]-20-sfGFP does not present intracellular droplets until 6-hour mark. At this point the cytoplasmic line remains constant as the total fluorescent increases. [S]-40-sfGFP transitions prior to the 2-hr timepoint. Therefore, the dotted line remains constant as the solid line increases slowly over time.

At the point of phase separation, small densely fluorescent regions appear within the cell that grow over time (Figure 44). Simultaneously, the remaining dilute regime remains at a relatively constant concentration (Figure 45). These two results suggest that upon increasing concentration after phase separation, cytoplasmic concentration of the protein is buffered at the expense of increasing dense phase volume relative to the
size of the overall cell. This result mirrors our previous protocell experiments and understanding of the binodal phase separation of these UCST polypeptides *in vitro*.

As expected, MW and aromatic:aliphatic content affects droplet formation in bacteria. Doubling the MW of [S]-20-sfGFP to [S]-40-sfGFP decreases $C_{sat}$ enough to cause droplet formation even prior to IPTG induction (Figure 46). It is noticeable that [S]-40 has a lower soluble phase concentration outside the dense droplet phase compared with [S]-20.

![Figure 46: Confocal images of UCST polypeptides as a function of induction time and molecular weight. A higher intracellular concentration is required for [S]-20 versus [S]-40 to form intracellular droplets. All scale bars are 5 µm.](image)

Similarly, increasing the aliphatic content with Val at the expense of Tyr increases the $C_{sat}$ to a concentration that was not measurable in our experiments (Figure 47). Although differences in $C_{sat}$ are not as dramatic as predicted by our *in vitro* experiments with RIDP-sfGFP fusions, perhaps via an effect of intramolecular crowding within the cell, we were able to modulate the $C_{sat}$ by an order of magnitude using both MW and aromatic:aliphatic ratio.
These experiments demonstrate that the formation and dynamics of droplet assembly can be distilled into two simple design parameters that require minimal perturbation of the assembling sequence – chain MW and ratio of aromatic:aliphatic residues. These design rules faithfully translate from in vitro to intracellular phase separation allowing for the a priori prediction of intracellular droplet behavior.

3.3.2 RIDPs exhibit reversible UCST droplet formation in E. Coli

Just as one can cross a binodal line into the two-phase regime under isothermal conditions by increasing polypeptide volume fraction, this line may be crossed under constant volume fractions by decreasing solvent quality or chi (χ). Practically this is most easily accomplished by reducing the temperature of the bulk solution[79]. Just as
observed in vitro, RIDPs exhibit reversible UCST phase separation inside of cells that is repeatable over four cooling and heating cycles (Figure 48).

![Figure 48](image)

**Figure 48**: Intracellular droplets comprised of [S]-20-sfGFP can be formed and dissolved reversibly via alternating cooling and heating cycles. This process is completely reversible over four rounds of cooling and heating. Cooling rate = 5°C min⁻¹, induction time of 4 hours.

This phase separation exhibits minimal hysteresis as the difference in the transition temperature of cooling (Tt,C) and transition temperature of heating (Tt,H) varies by less than two degrees Celsius (Figure 49).

![Figure 49](image)

**Figure 49**: Normalized UCST cloud point does not change significantly over four heating and cooling cycles. Boxes indicated 25th-75th percentile. Whiskers indicate 10th-90th percentile.
Interestingly, upon multiple heating and cooling cycles we observed that E. Coli exhibit spatial phase separation memory, with puncta forming in the same location as the first cycle (Figure 50). Given this discovery, it is critical to note that the observed transition temperature was below room temperature (~15°C), suggesting that these cells are naïve to phase separation as they were incubated at 37°C and processed at room temperature.

Figure 50: [S]-20-sfGFP exhibits phase separation memory upon multiple cycles of heating and cooling. Scale bar indicated 5 µm. Cooling and heating rate were set to a constant 5°C min⁻¹.

Additionally, we observed cooling-triggered phase separation results in a higher number of puncta per cell (Figure 51). The greater number of puncta observed with higher MW species (but similar transition point that is < 37°C) indicates that the number of puncta formed per cell is a function of their diffusion coefficient, consistent with prior literature[148].
Figure 51: Image analysis of the number of puncta formed in each cell. A. Number of intracellular puncta formed in each cell containing [S]-20-sfGFP during a cooling ramp from 60°C -> 10°C (green) and imaged isothermally at 22°C. Isothermal analysis performed at 6 hours post induction, the first timepoint where intracellular puncta were observed. Cooling ramp performed at 4 hours post induction, where transition temperature ($T_t$) was between 22°C and 37°C. B. Number of intracellular puncta formed in each cell containing [S]-40-sfGFP during a cooling ramp from 60°C -> 10°C (green) and imaged isothermally at 22°C. Isothermal analysis performed at 4 hours post induction, the first timepoint where intracellular puncta were observed. Cooling ramp performed at 4 hours post induction, but the transition observed was >37°C indicating the possibility of memory. C. Number of intracellular puncta formed in each cell containing [S]-40-sfGFP during a cooling ramp from 60°C -> 10°C (green) and imaged isothermally at 22°C. Isothermal analysis performed at 4 hours post induction, the first timepoint where intracellular puncta were observed. Cooling ramp performed at 4 hours post induction, where $T_t$ was between 22°C and 37°C.

Increasing the MW of the RIDP increases the cloud point temperature observed inside of bacteria (Figure 52). Predictably, just as observed in vitro, reducing the aromatic content while increasing aliphatic content dramatically reduces the observed droplet formation temperature in the E. Coli cells (Figure 53).
Figure 52: UCST cloud point is a function of polypeptide molecular weight. Cooling ramp = 60°C -> 10°C. Cooling rate = 5°C min\(^{-1}\), induction time of 8 hours.

Figure 53: UCST cloud point is a function of polypeptide aromatic:aliphatic content. Cooling ramp = 60°C -> 10°C. Cooling rate = 5°C min\(^{-1}\), induction time of 8 hours.
By manipulating the aromatic:aliphatic ratio while keeping MW constant and observing transition temperature within individual bacterium at various times post-induction (varying concentration), we were able to create partial intracellular phase diagrams as a function of UCST cloud point and intracellular fluorescence (Figure 54). This result is important because it ties the observed behavior upon cooling to a specific concentration for a given construct, essentially normalizing the observed cloud point for differing overall levels of protein expression. Again, with increasing concentration we see an increase in UCST cloud point, although the rate of increase upon increasing concentration does not appear to follow log normal dependence (Figure 54).

Figure 54: Intracellular binodal lines of various RIDP-sfGFP fusion. UCST cloud point increases as a function of cellular fluorescence and polypeptide aromatic content. Data analyzed at 2, 4, 8 hours for [S]-40-sfGFP and [3Y:V]-10 and 4, 8, 24 hours for [Y:V]-40 (n = 30). Error bars indicate standard error of the mean.

To demonstrate the functional utility of these dynamic intracellular structures we designed a droplet capture experiment based on split green fluorescent protein (GFP).
On an araBad regulated plasmid, the short peptide (GFP-11) was genetically fused to [3Y:V]-40 and expressed at concentrations that produce intracellular droplets (Figure 5E). The other plasmid regulated by LacZ, containing the other half of GFP (GFP-1-10), was expressed in conjunction with continued production of GFP-11-[3Y:V]-40. As both proteins are induced simultaneously, GFP-11-[3Y:V]-40 recruits both halves of the GFP into the intracellular droplets (Figure 5E).

![Image](image1.png)

**Figure 55:** GFP-11-RIDP fusion proteins are able to recruit GFP-1-10 from the surrounding cytoplasm into intracellular droplets. Upon formation of intracellular droplets after 24 of IPTG induction, co-expression of both GFP-11-[3Y:V]-40 and GFP-1-10 enables recruitment of fully functional sfGFP into the existing intracellular dense phase.

This experiment when performed with a complementary fluorescent molecule (mRuby3) attached to the RIDP, confirms the colocalization of the GFP-1-11 and the RIDP puncta (Figure 56).
Figure 56: Confocal microscope images of split GFP recruitment into intracellular droplets. A. GFP-11-[3Y:V]-40-mRuby3 co-expressed in the presence of GFP-1-10 creates fluorescently active GFP only in the interior of the droplet. B. In the absence of GFP-1-10 induction, there is little green fluorescent inside the intracellular droplets. Data taken at 22°C. Scale bar = 5 µm.

Once a fully functional GFP molecule is recruited into the intracellular droplets, it is then possible to dynamically modify the intracellular solubility of the complemented molecule (Figure 57). Upon heating to 60°C, the GFP-[3Y:V]-40 droplet dissolves. Cooling back to 10°C recreates the droplet structure. This result indicates that droplets
could be designed to dynamically tune the accessibility of captured proteins inside of bacterial cells.

**Figure 57:** Upon recruitment of GFP-1-11 into the dense phase, the solubility of the entire complex can be modulated with temperature. Data was collected for 36 hours post-IPTG induction and 12 hours post-arabinose induction.

In addition to droplet capture of an endogenous protein, RIDPs could be utilized as intracellular colocalization tags. We co-transformed two plasmids into bacterium containing two different RIDPs with unique fluorescent reporting tags. In a cell line where the aromatic:aliphatic ratio of the two proteins are similar - [S]-40-mNeonGreen and [3Y:V]-40-mRuby3 – the two proteins phase separate in the same intracellular location. In a cell line where the ratio is different - [S]-40-mNeonGreen and [Y:3V]-40-mRuby3 – the high aromatic content RIDP phase separates without recruiting the soluble [Y:3V]-40-mRuby3 protein (Figure 58).
3.3.3 Alpha-peptide-RIDP fusion proteins form functional intracellular droplets with enzymatic efficiency linked to RIDP molecular weight

One proposed reason for the evolutionary development of biomolecular condensates is to increase kinetics of various biological functions, including enzymatic reactions\[58, 149-151\]. However, there is little evidence of enzymatic proteins that benefit from improved kinetics when sequestered inside biomolecular condensates. Given that RIDPs form condensates that are minimally interactive with the surrounding milieu, we sought to perform bottom-up design of an intracellular droplet system using a protein enzyme and substrate that can improve biological activity via sequestration. We genetically fused the alpha peptide (ap or αp) from the LacZ β-galactosidase enzyme to
a RIDP-mRuby3 construct with the hypothesis that the αp-RIDP-mRuby3 protein could bind and recruit mutated LacZΔM15 expressed endogenously in genetically modified E. coli (KRX, Promega) into intracellular droplets. By splitting the enzyme into two non-functional halves, we are linking enzymatic activity to the location of the RIDP-mRuby3 fusions. After protein induction and resulting droplet formation, we deliver the substrate Fluorescein Di β-Galactopyranoside (FDG) to the cell medium where it is trafficked intracellularly, hydrolyzed into green fluorescent fluorescein at the sites of active β-galactosidase, and eventually exported outside the cell (Figure 59)[152, 153]. By tracking the onset of green fluorescence with confocal microscopy we can specifically observe where and when enzymatic activity is occurring within the cell.

Figure 59: Schematic of experimental set up that would allow for RIDPs to recruit enzymes into engineered intracellular puncta.

In our control experiment – αp-mRuby3 – we observe limited fluorescent persistence within the cells. It is important to note that the alpha peptide itself is known to form inclusion bodies and therefore, even in this control experiment, we observe puncta inside of the bacterial cells. In this experiment, the converted product is rapidly
exported from the intracellular space into the surrounding medium. However, upon fusion with RIDP [S]-20, we observe that fluorescein fluorescence localizes long enough with the RIDP puncta to be observed with confocal microscopy (Figure 60). Despite this increased colocalization, the total fluorescent production over time is not statistically significantly than the αp-mRuby control.

Figure 60: Confocal microscopy images showing the fluorescent conversion of Fluorescein Di-β-D-Galactopyranoside (FDG). Increasing the molecular weight of the RIDP portion of the gene leads to increased FDG conversion at earlier timepoints and higher overall conversion after 20 min. In the case of αp-[S]-80-mRuby3, the fluorescence of the puncta was lower than could be adequately rebalanced with similar settings for the other constructs, but nonetheless is present.
When we increase the MW of the RIDP, and thus decrease $C_{sat}$, we observe an effect in the total FDG fluorescent intensity as well as colocalization with the $\alpha_p$-RIDP-mRuby3 fusion (Figure 60). $\alpha_p$-[S]-40-mRuby3 and $\alpha_p$-[S]-80-mRuby3 have 2.5X and 7.5X respectively the amount of FDG converted at 20 minutes compared with the $\alpha_p$-mRuby3 control (Figure 61).

![Graph](image)

**Figure 61: Quantified amount of converted FDG intracellularly, normalized to the amount of mRuby3 fluorescence.**

Quantification of the colocalization of green and red signals with Mander's overlap coefficient[147] indicates increased colocalization when the alpha peptide is fused to RIDPs compared to the fluorescent reporter alone (Figure 62).
Figure 62: Mander’s colocalization score between converted FDG and fluorescent reporter. Background was automatically set by ImageJ software plugin.

To quantify the observed phenomenon of colocalization, we analyzed individual cells within the image frame with green fluorescence that was above the background threshold at each timepoint. Higher MW RIDPs exhibit higher fluorescence inside the cell normalized to the background at each point in time (Figure 63). This dose response effect of MW emphasizes the mechanisms of increased persistence of the substrate molecule inside the droplets leading to more efficient green fluorescence conversion.
Figure 63: Ratio of converted FDG fluorescence inside the cell versus outside the cell. All αp-RIDP-mRuby3 fusions exhibit a higher ratio of FDG fluorescence inside the cell, indicating a greater persistence of fluorescent FDG inside the intracellular space compared to the αp-mRuby3 control. Error bars indicate standard error of the mean.

Quantification of fluorescent production at various substrate concentrations in vitro suggests that the mechanism of this effect is a statistically significant increase in the catalytic rate of the enzyme with increasing MW. We observed 1.4X, 1.6X and 4.2X increase in the catalytic rate for αp-[S]-20-mRuby3, αp-[S]-40-mRuby3, αp-[S]-80-mRuby3 compared to the αp-mRuby3 control (Figure 64 and Table 4). This result combined with non-significant changes to $K_m$, suggests that the binding constant of the enzyme and substrate is not a main factor affecting the catalytic efficiency between different αp-RIDP-mRuby3 fusions.
Figure 64: Lineweaver-Burk plots for determining $K_m$ and $V_{max}$. Lineweaver-Burk plots created with variable starting concentrations of FDG for A. αp-mRuby3, B. αp-[S]-20-mRuby3, C. αp-[S]-40-mRuby3 and D. αp-[S]-80-mRuby3. Slopes ($V_o$) were determined from fluorescent generation over the course of 20 minutes. Intercepts and slope were used in the calculation of $K_m$ and $V_{max}$. 
Table 4: Michaelis-Menten Enzyme Kinetics Parameters (error is standard error of the mean, n = 3)

<table>
<thead>
<tr>
<th>Protein</th>
<th>$V_{\text{max}} (\text{Fl}_{\text{FDG}} \text{min}^{-1})$</th>
<th>$K_{m} (\text{Fl}_{\text{FDG}})$</th>
<th>$k_{\text{cat}} \text{min}^{-1}$</th>
<th>$k_{\text{cat}}/K_{m} (\text{Fl}_{\text{FDG}} \text{min}^{-1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>αp-mRuby3</td>
<td>3708 ± 183.2</td>
<td>4972 ± 636.3</td>
<td>3.56 ± 0.10</td>
<td>7.43E-04 ± 1.02 E-04</td>
</tr>
<tr>
<td>αp-[S]-20-mRuby3</td>
<td>2165 ± 35.02</td>
<td>5896 ± 380.5</td>
<td>5.10 ± 0.43</td>
<td>8.79E-04 ± 1.16 E-04</td>
</tr>
<tr>
<td>αp-[S]-40-mRuby3</td>
<td>1549 ± 34.75</td>
<td>5988 ± 475.1</td>
<td>5.75 ± 0.13</td>
<td>9.81E-04 ± 9.96 E-05</td>
</tr>
<tr>
<td>αp-[S]-80-mRuby3</td>
<td>2922 ± 174.9</td>
<td>4504 ± 439.7</td>
<td>15.2 ± 0.68</td>
<td>3.45E-03 ± 4.58 E-04</td>
</tr>
</tbody>
</table>

We also fused the LacZ alpha peptide to RIDPs with differing levels of aromatic content at a constant MW (Figure 65A). We hypothesized that differing levels of aromatic content would affect FDG uptake into the droplets and therefore affect overall enzymatic activity. Surprisingly, we observed similar overall levels of enzymatic expression between αp-[S]-40-mRuby3, αp-[3Y:V]-40-mRuby3 and αp-[Y:V]-40-mRuby3. However, the dynamics of enzymatic activity are different, with RIDPs of greater aliphatic content allowing for faster uptake into the puncta (Figure 65B). The differences between the ratio of FDG fluorescence inside the cell and outside the cell between αp-RIDP-mRuby3 fusions with different aliphatic content were insignificant, indicating that MW is the primary driving force for fluorescein and/or FDG persistence inside intracellular droplets (Figure 65C).
Figure 65: Enzymatic droplets formed with variable ratio of aromatic to aliphatic ratio. All scale bars are 5µm. A. Confocal microscopy images observing the fluorescent conversion of Fluorescein Di-β-D-Galactopyranoside (FDG) of αp-mRuby3, αp-[S]-40-mRuby3, αp-[3Y:V]-40-mRuby3 and αp-[Y:V]-40-mRuby3.  B. Quantified amount of converted FDG intracellularly, normalized to the amount of mRuby3 fluorescence. Error bars indicate standard error of the mean. C. All αp-RIDP-mRuby3 fusions exhibit a higher ratio of FDG fluorescence inside the cell, indicating a greater persistence of fluorescent FDG inside the intracellular space compared to the αp-mRuby3 control. Error bars indicate standard error of the mean.

In summary, higher MW RIDPs more efficiently sequester substrate in the enzymatically active, intracellular puncta. This results in a higher enzymatic efficiency leading to an observed higher fluorescence.

3.4 Conclusions

From our exploration of the molecular driving forces of UCST cloud point in vivo, the same parameters that largely drive the observed behavior in vitro apply in vivo. Our work supports the growing evidence of Arg-aromatic interactions that drive phase behavior and adds additional evidence of the molecular hierarchy that exists between
Trp, Tyr, Phe and His. Although the IDP literature often ignores the importance of MW, our results suggest that MW may be more critical than composition on UCST cloud point. Using these two parameters—aromatic:aliphatic ratio and MW—we were able to produce RIDPs with $C_{\text{sat}}$ ranging from nanomolar to millimolar, a larger difference than any other study has measured to date. We anticipate that these results will dramatically shift the strategy for mutating wild-type IDPs and designing de novo IDPs.

All of the RIDPs reported herein exhibited UCST phase separation behavior. However, there are clear examples for LCST phase behavior in nature[77]. This dichotomy brings into focus the question of evolutionary utility of LCST versus UCST in an isothermal environment. Furthermore, if both LCST and UCST do exist, under what pressures would evolution select for one over the other given their distinct differences in protein sequence[78]? Since temperature is often ignored in eukaryotic IDP research, there may be an unexplored landscape of IDP behavior in organisms that do not maintain a single physiological temperature.

Our demonstration that RIDPs phase separate inside of cells according to the same principles that drive their cloud point in vitro indicates that basic thermodynamic driving forces of chain chemistry and molecular weight are capable of affecting droplet formation dynamics in isolation. We believe that these fundamental associative driving forces are the foundation upon which two- and three-body phase separation is built. Due to the simplicity of their design, RIDPs behave in vivo exactly how their phase diagrams in vitro would suggest—intracellular concentration increases to a $C_{\text{sat}}$, at which point small phase separating droplets form that continue to grow with increasing overall
concentration inside the cell. This predictable observation has been theorized by previous studies but unproven until now[154].

The extrinsic modularity of these intracellular droplets is remarkable. The ability to trigger phase separation with temperature indicates that these proteins are likely sensitive to other modulators of chi, such as post-translational modification, salinity or pH. Our experiments demonstrate that these triggers can be used to control intracellular organization with extra-droplet components, providing a useful tool for precise cellular signal engineering at the protein level.

Finally, these proteins can be used for the de novo design of enzymatically functional intracellular droplets. The vast majority of IDP studies to date have focused on the interaction between IDPs and nucleic acids. However, we have demonstrated the ability to engineer the sequestration and control the solubility of more complex cellular components. We logically designed intracellular puncta capable of binding and recruiting mutated β-galactosidase, which could modify the catalytic efficiency of the enzyme-substrate complex – a complex which has not evolved to form intracellular condensates. Improved enzymatic efficiency proved to be dose-dependent, with increasing fluorescence occurring at each incremental increase in MW. Our motivation for this demonstration was simple – to combine a phase separating gene with an enzyme that functions bio-orthogonally to phase separation to create an entirely new droplet forming system with emergent behavior.

In summary, leveraging over 60 IDPs produced from Chapter 2 that exhibit variable saturation concentrations, we built intracellular condensates through a bottom
up approach, yielding a mixture of predictable and emergent behavior. We expect that these materials will be the fundamental building blocks from which new biological condensates with non-canonical behaviors can be built to better study the functional significance of phase separation in living cells.
4. Creation of UCST-LCST and UCST-UCST micelles capable of multivalent display of functional moieties

4.1 Introduction

The last few decades have seen an explosion of interest in the development of nanoparticle carriers for drug delivery, much of it for the treatment of solid tumors. Many different types of nanoparticles have been synthesized and evaluated in preclinical models for cancer therapy, including inorganic nanoparticles[155-158] dendrimers[159-163], polymer nanoparticles[164-166] and self-assembled nanostructures —micelles and polymersomes/liposomes— of polymers[167-169] and lipids[170-172].

The general interest in this area is to achieve the fabled “magic bullet” drug – a drug that acts potently as intended without negative side effects[173]. The problem with this ideal is that most drugs must balance potency against delivery obstacles posed by human physiology[174, 175]. Nanoparticles have attracted much attention from the “nanomedicine” community because they may be the answer to this challenge, for two reasons. First, they can be loaded with a range of small molecules with diverse physiochemical properties, making them near universal carriers for small molecule drugs[176-178] and imaging agents[179-181]. Second, appropriately designed nanoparticles show good colloidal stability in blood and can circulate for extended periods of time[182-184]. Therefore, the promise of nanocarriers is to understand how material properties on the nanoscale can alter the negative properties of developed drugs and endow them with more desirable properties. Other desirable properties may include, resistance to drug
clearance/breakdown, tissue specific targeting, increased cell internalization and increased solubility in serum.

Over the past 20 years, many different drug delivery systems have arisen and some have progressed into approved therapies today[185]. The most widely used systems are liposomal formulations, where the core of the liposome is loaded with drugs and therefore endows the encapsulated cargo with the properties of the liposome formulation. Liposomes are a great delivery vehicle because of their biocompatibility, ease of synthesis and high loading capacity[186].

However, for encapsulating more hydrophobic moieties, polymeric micelle systems are more advantageous as they have a larger volume/g that is hydrophobic and more control over morphology in comparison to liposomes. However, challenges with synthesis, controlling polydispersity of assembled populations, incorporating additional functionality, burst drug release in vivo and overall biocompatibility challenges have limited their clinical potential[187-189].

Two terms often used in the literature are “active” and “passive” targeting. Passive targeting is enhanced through optimization of the shape, size and surface charge of nanoparticles to improve tumor accumulation due to the enhanced permeability and retention effect [190-203]. Recently, particles with high aspect ratios and high flexibility, referred to as filomicelles, have attracted much research interest due to their long circulation time, high tumor penetration and accumulation, and enhanced active target delivery[191, 204-208]. These particles are created via self-assembly or pattern-molding[200], which are convenient to create precise particle shapes, but are
somewhat incompatible with protein drugs or the presentation of protein targeting ligands, as the conditions employed for their synthesis may denature proteins rendering them inactive in the body.

Passive targeting is a useful approach for locoregional targeting of solid tumors, but it does not directly target tumor cells, which are the ultimate destination of the drug or imaging agent. The rationale for creating targeted nanoparticles for cancer therapy or imaging stems from the fact that many tumors have surface proteins that are either overexpressed or—in a few instances—are uniquely expressed on the surface of tumor cells compared to normal, healthy cells. Homing the nanoparticle to tumor cells by decorating it with a ligand specific to a tumor-selective or tumor-specific marker—once the carrier has accumulated to a high enough concentration in the local environment of the tumor—can provide a second stage of tumor-cell specific targeting.

Active targeting utilizes a specific binding motif and targeting structure on the cell of interest to localize particles loaded with drug or imaging agent with a biophysical signal. A common approach to synthesize targeted nanocarriers is to functionalize the surface of the nanoparticle with a peptide or protein by covalent conjugation. This approach however provides limited control of ligand valency, and typically requires an excess of ligand to drive the reaction, and is hence expensive to scale up, and quality control and product validation remains a significant challenge[209-211].
4.1.1 Literature review

In recent years, due to advances in polymer synthesis technology, the ability to investigate specific biophysical properties of nanomaterials has received much research attention. Almost every aspect of a nanoparticle’s design has been tested for improved efficacy in along some dimension (biocompatibility, pH, tissue extravasation etc.)[212]. A brief list of the parameters is highlighted in Figure 66[212] and a brief summary of optimal design specifications is included.
Size: Large particles (>1µm) are internalized by macrophages, neutrophils, dendritic cells[213]. Smaller than 1µm, they are internalized via pinocytosis or receptor mediated endocytosis[214]. Particles in the 40-50 nm range exhibit maximum uptake[215]. Particles between 10-100 nm are typical size ranges for optimization of biodistribution and clearance[216]. Smaller than 5.5 nm are rapidly cleared by the kidneys[217]. It is thought that corona chain curvature and conformation is particularly crucial for determining in vivo fate[218].

Shape: Rod like designs are more readily taken into cells than spherical counterparts[219]. Non-spherical particles appear to have longer circulation times compared to spherical counterparts[220].

Surface chemistry: Charged nanoparticles have shorter blood circulation times and highly nonspecific cellular uptake[216]. This can be readily tuned in most synthetic polymer systems but neutral charge is generally the best for most applications.

Corona Hydrophobicity: Block copolymers with increased corona hydrophobicity are more easily taken up by cells but also have higher levels of opsonization[214]. For in vivo applications, the optimal formulation will vary, especially when targeted therapies are concerned.

Core stability: Micelle half-life can be controlled via core stability as measured via pyrene I1/I3 fluorescence. Other studies have demonstrated that crosslinking polymeric micelle cores can also increase the observed half-life in vivo[221].

Particle rigidity: Deformable structures can last up to 30 times longer in circulation than rigid counterparts[222].
Targeting/stimuli responsive elements: Targeting has generally improved nanocarriers compared to a non-targeted system. However, the incorporation of targeting ligands or environmentally sensitive moieties often alters the surface charge, morphology or both. One study has examined the effect of ligand density on tumor targeting and did find that an optimal ratio exists for that particular cancer phenotype. This result indicates that an opsonization versus targeting tradeoff exists and must be accounted for when introducing targeting ligands[223].

With synthetic polymer systems, the most challenging of these parameters to control are the targeting/responsive elements and geometry. These design parameters can often not be controlled with one pot synthesis and therefore multiple step construction is required. Multistep processes almost inevitably introduce polydispersity/heterogeneity, which can cloud the conclusions made about a particular design choice. In fact, a recent study demonstrated that a 10-20 nm deviation could significantly impact nanoparticle behavior in the body[224]. Thus, despite all the advances in polymer micelle synthesis, it is still difficult to design optimal micelle carriers. In an ideal scenario, one would be able to a priori incorporate optimal design elements for a particular application at the design stage.

Controlling morphology of block copolymers: Micro-phase separation of diblock copolymers depends on three parameters: volume fraction of both blocks combined, total degree of polymerization and the Flory-Huggins parameters (χ). The chi parameter specifies the miscibility of both the blocks, or in an amphiphilic block copolymer case, the immiscibility. The chi parameter is also a function of temperature. For a system...
consisting of just the block copolymer, the chi parameter contains interaction energies between blocks A-B, A-A, B-B. Increasing the temperature or decreasing chi, compatibility between the blocks improves, combinatorial entropy increases and copolymers undergo an order to disorder transition.

\[
\chi_{AB} = \left( \frac{Z}{k_B T} \right) \left[ \bar{\epsilon}_{AB} - \frac{1}{2} \left( \bar{\epsilon}_{AA} + \bar{\epsilon}_{BB} \right) \right]
\]

Figure 67: Predicted equilibrium morphologies of AB diblock polymer in bulk. A. S and S’ = body-centered-cubic spheres, C and C’ = hexagonally packed cylinders, G and G’ = bicontinuous gyroids, and L = lamellae. B. Theoretical phase diagram of AB diblocks predicted by the self-consistent mean-field theory, depending on volume fraction (f) of the blocks and the segregation parameter, \(\chi_N\), where \(w\) is the Flory–Huggins segment–segment interaction energy and \(N\) is the degree of polymerization; CPS and CPS’ = closely packed spheres. C. Experimental phase diagram of polyisoprene-block-polystyrene copolymers, in which \(f_A\) represents the volume fraction of polyisoprene, PL = perforated lamellae. This figure was reproduced from Bates and Fredickson.
Of course, the number of chi parameters jumps to six once water is introduced into the system since each block can interact with itself, the other block, and water. However, controlling morphology in aqueous solution can be simplified to a function of three polymer primary variables – interfacial energy between the blocks (enthalpic), chain stretching of the core (entropic) and chain repulsion in the corona. A balance between repulsive corona-corona interactions and conformational entropy penalty for extending the chains determines the actual conformations of the corona chains. It is important to note that this balance is affected by the self-assembled morphology. The core block extension is also affected by morphology. Core chains are most extended in a spherical morphology and most compact in a rod-like morphology.

Upon the formation of microstructure, the blocks attempt to minimize the total interfacial energy of the system. During this process they sacrifice the entropic gains of forming single chains, to prevent from paying an even larger penalty of hydrophobic-water interactions. This lowers the total free energy of the system. Increasing the size of the core block (A) the corona volume fraction of the total length of the chain decreases. As a result, less curvature is observed at the interface of the polymer chain.

### 4.1.2 Approach

Combination of UCST and LCST thermodynamically decouples the core and corona blocks as much as possible via side chain chemistry. Using basic rules of polymer physics regarding block copolymers, we’ll seek to create self-assembled structures of a variety of shapes and sizes for presentation of various ligands.
Key parameters of interest are the hydrophilicity of the corona forming block, the hydrophilicity of the core block, the overall length of the co-polypeptide and the ratio of the two blocks. Parameters of evaluation are the size, morphology, stability and thermoresponsive behavior. All measurements will be made in 140 mM NaCl, 10 mM phosphate buffer, 3 mM KCl, pH 7.4 unless otherwise specified. Size will be evaluated by the hydrodynamic radios ($R_h$) and radius of gyration ($R_g$) by dynamic light scattering (DLS) and static light scattering (SLS) respectively. $R_g$ and $R_h$ can be combined to yield the shape factor $\rho = R_g/R_h$, which gives a rough indication of the morphology of the scatterer. A shape factor of 1.505 suggests a Gaussian polymer chain, 1.0 suggests a hollow sphere or vesicle, and 0.775 suggests a solid sphere. For an elongated scatterer, the shape factor depends upon the aspect ratio. A combination of temperature dependent turbidity and DLS will be utilized to determine the phase behavior of the block co-polypeptides. Cryogenic transmission electron microscopy (Cryo-TEM) will be utilized to evaluate the morphology and also provide crucial insight into the hydration of the core/corona chains. Stability of the assembled nanostructure will be determined by a shift in the I1/I3 fluorescent bands of pyrene as described previously.

One of the targeting domains chosen for the second portion of this study is the 10th, type III domain from human fibronectin (Fn3) that targets the human αvβ3 integrin, a receptor that is upregulated in the endothelium of many tumors [225-227] and is also overexpressed on several tumor cells such as glioblastoma[228], renal cell carcinoma[229], ovarian carcinoma[230, 231] and breast cancer metastases. We chose a Fn3 variant that binds the αvβ3 integrin with low affinity[232, 233] (K_D > 1x10^{-7} M), and
we have previously shown that the Fn3 domain can be expressed in E. coli as a fusion to repetitive polypeptides such as ELPs[234]. The low affinity of the parent Fn3 domain is important as, multivalent presentation could amplify its avidity, which may not be possible with ligands that possess intrinsically high affinity, so that we could test for the effect of self-assembly and multivalency on binding avidity and cellular uptake[234, 235].

4.2 Methods

4.2.1 Gene synthesis

Plasmid genes were available from previous studies for RLP20, RLP20-ELP80, RLP40-ELP80, RLP80-ELP80, RLP100-ELP80, and an Fn3 domain that binds the αvβ3 integrin. This gene was then subsequently fused with the gene that encodes the Fn3 domain. Similarly, genes encoding RLP20-ELP80, RLP40-ELP80, RLP80-ELP80 were cloned to the N-terminus of an Fn3 with the same directional ligation method. After successful confirmation of gene assembly by Sanger fluorescent DNA sequencing, the plasmids harboring each construct were isolated and transformed into BL21(DE3) expression strain of E. Coli. Aliquots of the cell stocks were stored at -80 °C until further use.

4.2.2 Protein purification

Each block polypeptide was expressed in BL21(DE3) E. coli using a previously published hyperexpression protocol. 5 mL bacterial cultures were grown overnight from frozen glycerol stocks and used to inoculate 1 L flasks of TB Dry, supplemented with 45 µg/mL kanamycin. The flasks were then incubated at 37°C for 24 hours and 190 rpm.
Each construct was purified using inverse transition cycling (ITC) [236]. Briefly, the cell suspension was centrifuged at 3,000 rpm for 10 min at 4°C, the cell pellet then resuspended in PBS and then lysed by sonication on ice for 2 min (10 s on, 40 s off) (Misonix S-4000; Farmingdale, NY). Polyethyleneimine (PEI) 0.7% w/v was added to the lysate to precipitate nucleic acid contaminants. The supernatant was then subjected to multiple rounds of ITC as follows: the solution was kept on ice, and 3 M NaCl was added to isothermally trigger the phase transition of the RLP-ELP block co-polypeptide. The coacervate was then centrifuged for 20 min at 14,000 g and 30 °C, the supernatant was decanted and discarded, and the pellet was resuspended in phosphate buffer. The dissolved product was cooled to 4°C, and then centrifuged for 10 min at 15,000 and 4°C to remove any insoluble contaminants. To remove excess salt from purified protein solutions, the samples were dialyzed against ddH2O at 4°C for at least 24 h using SpectrumTM Labs Spectra/PeptTM 2 12-14 Standard RC Dry Dialysis Kits (Fisher Scientific, Waltham, MA). The proteins were then lyophilized and stored at -20°C. Purity of the block polypeptides was assessed by SDS-PAGE gel with SimplyBlue staining.

4.2.3 Characterization of phase separation

4.2.3.1 Temperature dependent UV-vis spectrophotometry

Turbidity profiles were obtained for each of the constructs by recording the optical density as a function of temperature (1°C min⁻¹ ramp) on a temperature-controlled UV-vis spectrophotometer (Cary 300 Bio; Varian Instruments; Palo Alto, CA). The transition temperature (Tt) was defined as the inflection point of the turbidity profile.
Samples were measured in PBS at 10 µM. Because some of the block co-polypeptides which form larger micelles are slightly turbid when soluble, all measurements were taken after zeroing with PBS.

4.2.3.2 Static and dynamic light scattering

Static and dynamic light scattering measurements (SLS/DLS) were performed using an ALV/CGS-3 goniometer system (Langen, Germany). Samples for the ALV/CGS-3 goniometer system were prepared at a concentration of 10 µM in PBS and filtered through 0.45 µm Millex-GV filters into a 10 mm disposable borosilicate glass tube (Fischer). Simultaneous SLS and DLS measurements were obtained at 15°C of the ELP for angles between 30°-150° at 5° increments, with each angle consisting of 3 runs for 15 s. SLS experiments were only conducted for self-assembling block co-polypeptides, since the molecular weight of a single block co-polypeptide chain is already known, and the R_g of a single chain is likely near or below the detection limit of the SLS instrument. The differential refractive index (dn/dc) was determined by measuring the refractive index at different concentrations using an Abbemat 500 refractometer (Anton Paar, Graz, Austria). DLS data were analyzed by fitting the autocorrelation function with a cumulant fit, using the built-in ALV software. Hydrodynamic radius (R_h) was plotted against angle and extrapolated to zero. SLS data were analyzed by partial Zimm plots using ALVSTAT software in order to determine the R_g and molecular weight (MW).
4.2.3.3 Temperature-programmed dynamic light scattering

Temperature-programmed dynamic light scattering experiments were carried out using a Dynapro plate reader (Wyatt Technology; Santa Barbara, CA) with samples filtered through 0.45 µm Millex-GV filters. Data was collected at increments of 1°C, and the cumulant fit hydrodynamic radius was taken as the radius. The T₁ was defined as the temperature at which aggregates of size hundreds of nanometers were formed.

4.2.3.4 Cryogenic transmission electron microscopy

Cryo-TEM experiments were performed at Duke University’s Shared Materials Instrumentation Facility (Durham, NC). Lacey holey carbon grids (Ted Pella, Redding, CA) were glow discharged in a PELCO EasiGlow Cleaning System (Ted Pella, Redding, CA). A 3 µl drop (10 µM RLPn-ELP80) was deposited onto the grid, blotted for 3 s with an offset of -3 mm, and vitrified in liquid ethane using the Vitrobot Mark III (FEI, Eindhoven, Netherlands). Prior to vitrification, the sample chamber was maintained at 15°C and 100% relative humidity to prevent sample evaporation. Grids were transferred to a Gatan 626 cryoholder (Gatan, Pleasanton, CA) and imaged with a FEI Tecnai G2 Twin TEM (FEI, Eindhoven, Netherlands), operating at 80 keV. Feature sizes and spacing distances were measured in ImageJ by manual measurement of at least 25 particles.

4.2.3.5 Surface plasmon resonance spectrophotometry

The surface plasmon resonance experiments were performed using Biacore T200. Purified human αvβ3 integrin (Chemicon, Temecula, CA) were immobilized on
research grade CM5 sensor chips using an amine coupling kit (BIAcore, Piscataway, NJ). The integrin was diluted in 10mM sodium acetate buffer (pH 4.5) for conjugation with a surface density of approximately 600 resonance units (RU). The measurements of binding events were performed using block co-polypeptide concentrations ranging between 2.5 and 10 µM. The block polypeptides were diluted in HBS-P buffer (10 mM HEPES, 140 mM NaCl, 0.005% Triton-X, pH 7.4) supplemented with 2 mM CaCl2, and injected into the flow cells at a flow rate of 30 µl min⁻¹ for 4 min. The complex was allowed for dissociation for 10 min. The surface was regenerated with 10 mM Glycine-HCl (pH 2.5) at a flow rate of 30 µl min⁻¹ for 45 s, followed by 10 mM Glycine-HCl (pH 2.0) at a flow rate of 30 µl min⁻¹ for 30 s. The surface was regenerated using 10 mM glycine-HCl (pH 2.0). Kinetic modeling and simulations were performed using BIAevaluation software with the heterogeneous ligand model for self-assembled proteins and with a 1:1 ligand model for the unimeric protein (RLP20-ELP80-Fn3). The equilibrium binding constants (K_D1 and K_D2) for each experiment were calculated by dividing kinetic dissociation rate (k_off) by association rate (k_on), from which the mean K_D1/2 was derived. All SPR measurements were carried out at 25°C. The SPR measurements were carried out using polypeptide concentrations ranging between 2.5 and 10 µM. Goodness-of-fit was evaluated by analyzing residual plots and residual sum of squares.

4.2.3.6 Flow cytometry

Approximately 1×10⁶ cells were harvested from either K562 or K562+ αvβ3 cell lines and resuspended into 1 ml of serum-free medium containing 10 µM of the various
Fn3-decorated and control block polypeptides. LM609 antibody was also resuspended at 10 µM in serum-free medium. Micelles were prepared from a mixture of ~10% Alexa 488 dye-labeled RLP-ELP block co-polypeptides and 90% unlabeled polypeptides on a molar basis. The cells were incubated at 37°C with the labeled micelles for a specified time, then rinsed with 1 ml of Hanks Buffered Saline Solution (HBSS), collected by centrifugation at 500 RCF for 5 min at 20°C, and resuspended in HBSS + 1% BSA. Cells were maintained on ice until they were analyzed by flow cytometry (BD Accuri C5). The cell fluorescence intensity of Alexa 488 (Green) was quantified after gating to remove cellular debris on unstained control samples.

4.2.3.7 Confocal microscopy

Approximately $1 \times 10^6$ cells were harvested from either K562 or K562+ αvβ3 cell lines and resuspended into 1 ml of serum-free medium containing 10 µM of the various decorated and undecorated block polypeptides. Cells were incubated at 37°C for various times (20-240 min). After washing with HBSS thrice, 20 µL of cell suspension was added to a 384 well plate with a #1.5 coverslip on the bottom. Cells were imaged on a Zeiss 710 inverted confocal (Oberkochen, Germany) equipped with a live cell chamber maintained at 37°C using a 40X oil immersion objective.

4.2.3.8 Confocal image analysis

For analysis of percentage of cells that show uptake of the polypeptides, the fluorescent channel and DIC channel were isolated and analyzed independently. In the fluorescent channel the lowest 10% of cell fluorescence was removed, to eliminate any
autofluorescence from naïve K562 cells. Using this cutoff, locations and area of green fluorescence were identified using the fluorescent channel only. Total number of cells were then counted using the DIC channel.

4.2.3.9 Fluorophore labeling

The N-terminus and lysine residues in the RLP-ELP-Fn3 fusions were labeled with the NHS-ester derivative of Alexa488. To bias the reaction towards N-terminal labeling, the pH of the reaction mixture was adjusted to 8.3. The RLP-ELP block copolypeptides, dissolved in 0.1 M sodium bicarbonate buffer (pH=8.3), were incubated with a molar excess of dye (with a dye-to-protein molar ratio depending on the total number of reactive groups in the proteins, which includes lysine residues and the N-terminus (e.g. dye:RLP$_{20}$ = 2:1, dye:RLP$_{20}$-Fn3 is 5:1), for 2 hours at room temperature with continuous agitation. Excess dye was removed with 3 rounds of dialysis over 3 days at 4°C with a 1:500 volume ratio of reaction mixture to milli-Q water. The samples were lyophilized and stored at -20°C.

4.3 Results and Discussion

4.3.1 UCST-LCST construction and characterization

The first area of investigation was the effect of hydrophilic weight fraction of an RLP-ELP block copolypeptide. The core block sequence was (Gln-Tyr-Pro-Ser-Asp-Gly-Arg-Gly)-XX (RLPXX) and the corona sequence was (Val-Pro-Gly-[Ala/Gly]-Gly)-YY (ELPYY) where the guest ratio was a 50/50 split between Ala and Gly. The core block
size was controlled to be 20, 40, 60 or 80 repeat units of (Gln-Tyr-Pro-Ser-Asp-Gly-Arg-Gly) and the corona block was 80 repeats of (Val-Pro-Gly-[Ala/Gly]-Gly)-YY.


We understand some detail about the assembly of these polypeptides from our scattering experiments (Table 5). First, RLP20 – ELP80 has a hydrodynamic radius of 5.5 nm, in accordance with a fully soluble ~47 kDa polymer chain, and thus does not self-assemble. Second, RLP40 – ELP80 and RLP60 – ELP80 both self-assemble into structures with \( R_h \) less than 50 nm, \( R_g \) less than 40 nm, shape factors below 1 and aggregation numbers under 250. The radii, combined with the shape factors and the aggregation numbers, indicate that both RLP40 – ELP80 and RLP60 – ELP80 likely self-assemble into spherical micelles. Third, RLP80 – ELP80 self-assembles into much larger structures, with hydrodynamic radii over 100 nm, radii of gyration above 140 nm, a
shape factor around 1.2, and aggregation numbers in the thousands of chains. These results indicate that RLP80 – ELP80 self-assembles into much larger, non-spherical structures.

**Table 5: Light scattering data of RLPXX-ELP80 co-polypeptides**

<table>
<thead>
<tr>
<th>Block co-polypeptide</th>
<th>Hydrophilic Content</th>
<th>$N_{agg}$</th>
<th>$R_g$(nm)</th>
<th>$R_h$(nm)</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLP20–ELP80</td>
<td>64.7%</td>
<td>-</td>
<td>5.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RLP40–ELP80</td>
<td>47.4%</td>
<td>68</td>
<td>29.9</td>
<td>33.3</td>
<td>0.8</td>
</tr>
<tr>
<td>RLP60–ELP80</td>
<td>37.5%</td>
<td>231</td>
<td>23.8</td>
<td>36.7</td>
<td>0.6</td>
</tr>
<tr>
<td>RLP80–ELP80</td>
<td>30.9%</td>
<td>2240</td>
<td>145.4</td>
<td>114.3</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Cryo-TEM results confirm the suspicions of both dynamic and static light scattering (Figure 69). Some important observations for all constructs are that the nanostructures are spatially very close to one another indicating that these structures are near the overlap regime. Sample preparation was at 10 µM, theoretically in the dilute regime. Therefore, the concentration of these structures is artificially increased by the vitrification process.
Figure 69: Cryo-TEM images of RLPXX-ELP80 block co-polypeptides. Overall increasing the hydrophilic weight fraction by decreasing the size of the core block will shift the self-assembly from worm-like to a spherical shape. Data collected at 10 µM in 140 mM PBS. Scale bar = 500 nm.

However, increasing the preparation concentration of RLP40 – ELP80 to 100µM and 1mM did not seem to affect the observed structure by cryo-TEM (Figure 70). Therefore, we expect that the self-assembled morphology has a wide range of concentration independence and that the observed shape in cryo-TEM is consistent with light scattering data. Finally, we were unable to visualize any of the corona chains for any of the RLP-ELPs sampled, likely due to their high water content and corresponding low contrast with water.
Figure 70: Cryo-TEM images of RLP40-ELP80 at increasing volume fractions does not appear to affect observed assembly state. Data collected at 10, 100 and 1000 µM in 140 mM PBS respectively. Scale bar = 500 nm.

RLP40 – ELP80 and RLP60 – ELP80 (Figure 69) both self-assemble into spherical micelles. This result was predicted by DLS and SLS and is confirmed by cryo-TEM. Measurements of core radii indicate that RLP40 – ELP80 and RLP60 – ELP80 cores are approximately 12.8 nm and 17.5 nm, consistent with a larger core-forming block leading to a larger micelle core. This is consistent with the larger values reported by light scattering, because R_h and R_g incorporate both the core and corona portions of the micelle, whereas only the micelle core is directly observed by cryo-TEM. The spacing was similar for both: 29.5 nm, which is consistent because both block co-polypeptides have the same morphology and same corona-forming ELP block.

Cryo-TEM reveals that RLP80 – ELP80 (Figure 69) forms a different nanostructure. These block co-polypeptides form long, cylindrical structures. Again, an apparent increase in the core block size from 17.5 nm to 19.9 nm and an increase in the
Spacing of 28.9 nm to 34.7 nm is consistent with increasing size of the core block (although the core size increase is non-significant). These overlapping structures are more consistent with lamellae formation and therefore the aspect ratio of the cylinder is not observable with cryo-TEM at this concentration.

Table 6: Micelle core and inter-core spacing of RLPXX-ELP80 block co-polypeptides (mean +/- standard deviation)

<table>
<thead>
<tr>
<th>Block polypeptide</th>
<th>Core radius (nm)</th>
<th>Spacing (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLP40–ELP80</td>
<td>12.8 +/- 1.8</td>
<td>29.5 +/- 4.8</td>
</tr>
<tr>
<td>RLP60–ELP80</td>
<td>17.5 +/- 2.7</td>
<td>28.9 +/- 5.1</td>
</tr>
<tr>
<td>RLP80–ELP80</td>
<td>19.9 +/- 2.8</td>
<td>34.7 +/- 4.2</td>
</tr>
</tbody>
</table>

In order to test if overall hydrophilic weight fraction or block length was the main driving force of micelle morphology, overall length was changed while maintaining a specific hydrophilic weight fraction. This resulted in a comparison between overall hydrophilic weight fractions of 64.7% (RLP20 – ELP80 and RLP40 – ELP-160), 47.4% (RLP20 – ELP-40, RLP40 – ELP80, RLP80 – ELP-160) and 30.9% (RLP40 – ELP-40, RLP80 – ELP80).

RLP20 – ELP80 with a hydrophilic weight of 64.7% did not assemble as mentioned previously. However, doubling the block length of the corona and the core (RLP40 – ELP-160) resulted in an assembled structure with a R_g of 70.8 nm, R_h of 92.3 nm and form factor of 0.8 indicating a spherical micelle morphology. This result was confirmed with cryo-TEM, which revealed spheres that had an average core radius of 11.0 nm and spacing of 22.3 nm between particle cores.
Table 7: Light scattering data of RLPXX-ELPYY co-polypeptides

<table>
<thead>
<tr>
<th>Block co-polypeptide</th>
<th>Hydrophilic Content</th>
<th>( N_{\text{agg}} )</th>
<th>( R_g (\text{nm}) )</th>
<th>( R_h (\text{nm}) )</th>
<th>( \rho )</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLP40–ELP160</td>
<td>64.7%</td>
<td>246</td>
<td>70.8</td>
<td>92.3</td>
<td>0.8</td>
</tr>
<tr>
<td>RLP20–ELP40</td>
<td>47.4%</td>
<td>-</td>
<td>-</td>
<td>5.3</td>
<td>-</td>
</tr>
<tr>
<td>RLP80–ELP160</td>
<td>47.4%</td>
<td>1728</td>
<td>78.2</td>
<td>93.5</td>
<td>0.8</td>
</tr>
<tr>
<td>RLP40–ELP40</td>
<td>30.9%</td>
<td>91</td>
<td>56.2</td>
<td>52.6</td>
<td>1.1</td>
</tr>
</tbody>
</table>

RLP20 – ELP40, RLP80 – ELP160, are both expected to assemble into spherical micelles due to the overall hydrophilic weight fraction of 47.4% which showed spherical micelles with RLP40 – ELP80. RLP20 – ELP40 however did not assemble and had a soluble \( R_h \) of 5.3 nm, consistent with a 32 kDa chain. This can be explained as a 32 kDa chain not having sufficient assembly domain size. RLP80 – ELP160 did assemble with \( R_g \) of 78.2 nm, \( R_h \) of 93.3 nm and form factor of 0.8 indicating spherical micelles.
Cryo-TEM imaging confirmed this result and provided interesting nanostructure information (Figure 71). The core of these micelles had a radius of 29.3 nm with an intra-core spacing of 59.0 nm. This core dimension is much larger than RLP80 – ELP80 which adopted a cylindrical micelle formation indicating that the core of the larger spheres is more expanded in a spherical micelle than in a cylindrical micelle. This result
can be directly predicted from the theory of synthetic polymer micelles as the chains of a sphere are expected to be more extended than in a rod-like conformation. Additionally, we observe a much larger intra-core spacing due to the doubling in the size of the ELP chain.

Table 8: Micelle core and inter-core spacing of RLPXX-ELPYY block copolypeptides (mean +/- standard deviation)

<table>
<thead>
<tr>
<th>Block polypeptide</th>
<th>Core radius (nm)</th>
<th>Spacing (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLP40–ELP160</td>
<td>11.0 +/- 1.6</td>
<td>22.3 +/- 5.2</td>
</tr>
<tr>
<td>RLP80–ELP160</td>
<td>29.3 +/- 4.5</td>
<td>59.0 +/- 13.2</td>
</tr>
<tr>
<td>RLP40–ELP40</td>
<td>11.8 +/- 1.6</td>
<td>19.5 +/- 4.3</td>
</tr>
</tbody>
</table>

Finally, RLP40 – ELP40 likely adopts a non-spherical geometry with a $R_g$ of 56.2 nm, $R_h$ of 52.6 nm and form factor greater than 1. This is comparable to RLP80 – ELP80 which adopted a similar conformation. Both the core radius (11.8 nm) and the spacing (19.5 nm) are smaller than the larger polymer which is consistent with the smaller core and corona chains. The core size is about the same size as the core radii of other assembled structures with different morphologies.

Another key design parameter besides the effect of hydrophilic weight fraction is the effect of corona hydrophilicity. Therefore, a more hydrophobic guest residue of Val (V) and more hydrophilic guest residue of Ser (S) were substituted into RLP40 – ELP80, RLP60 – ELP80, and RLP80 – ELP80. Our hypothesis is that the introduction of Val will reduce corona chain repulsion, leading from a more spherical to more rod-like assembly. A Ser would provide more chain repulsion and drive a rod-like to spherical transition.
The substitution of Ala/Gly for Val resulting in a spherical to worm shift in the light scattering data (Table 9). Comparing RLP40–ELP80 (spherical) and RLP40–ELPV80, RLP40–ELPV80 has a higher aggregation number and a larger $R_g$ value, resulting in a $\rho > 1$. Since these two polymers have nearly identical molecular weight and the exact same chain length, we can surmise that RLP40 – ELPV80 is forming more elongated structures. A similar increase is seen with RLP40 – ELPV80 compared to RLP40 – ELP-80. Just as with the Ala/Gly constructs you can clearly see $N_{agg}$, $R_g$, and $R_h$ increase as the core block length increases and yet the form factor remains >1, indicating that all the Val constructs are worm-like micelles.

### Table 9: Light scattering data of RLPXX-ELPS80 and RLPXX-ELPV80 copolypeptides

<table>
<thead>
<tr>
<th>Block co-polypeptide</th>
<th>Hydrophilic Content</th>
<th>$N_{agg}$</th>
<th>$R_g$ (nm)</th>
<th>$R_h$ (nm)</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLP40 – ELPS80</td>
<td>48.2%</td>
<td>80</td>
<td>27.2</td>
<td>33.0</td>
<td>0.8</td>
</tr>
<tr>
<td>RLP60 – ELPS80</td>
<td>38.3%</td>
<td>244</td>
<td>27.5</td>
<td>43.5</td>
<td>0.6</td>
</tr>
<tr>
<td>RLP80 – ELPS80</td>
<td>32.1%</td>
<td>392</td>
<td>36.0</td>
<td>49.4</td>
<td>0.7</td>
</tr>
<tr>
<td>RLP40 – ELPV80</td>
<td>48.9%</td>
<td>213</td>
<td>33.1</td>
<td>31.7</td>
<td>1.0</td>
</tr>
<tr>
<td>RLP60 – ELPV80</td>
<td>39.0%</td>
<td>565</td>
<td>41.3</td>
<td>39.4</td>
<td>1.0</td>
</tr>
<tr>
<td>RLP80 – ELPV80</td>
<td>32.8%</td>
<td>693</td>
<td>94.4</td>
<td>71.0</td>
<td>1.3</td>
</tr>
</tbody>
</table>

On the other hand, RLP40 – ELPS80 has approximately the same $N_{agg}$, $R_g$, and $R_h$ as RLP40–ELP80. This would indicate that both constructs are spherical micelles. Interestingly, it appears the substitution of serine for alanine and glycine in RLP80 – ELP80 has reduced the $N_{agg}$, $R_g$, $R_h$, so that the form factor is now <1. This would indicate that this substitution has led to a shift in morphology – from a wormlike micelle to a sphere. Just as with the Ala/Gly and Val constructs you can clearly see $N_{agg}$, $R_g$, and $R_h$. 
and $R_h$ increase as the core block length increases and yet the form factor remains $<1$, indicating that all the serine constructs are spherical.

As suspected by the light scattering data, an Ala/Gly to Val substitution resulted in spherical to worm-like micelle shift according to cryo-TEM imaging. In Figure 72 you can also see quite clearly that with increasing the hydrophilic weight fraction that the worms get progressively longer, the core radii remain approximately the same size (Table 10) and eventually become an interconnected network. The Ala/Gly substitution with Ser appears to have increased the corona repulsion so much that it can only form spherical micelles.
In addition to changes to the corona sequence to increase or decrease repulsion between chains (or perhaps changing the occupied volume of the chain), we sought to perform the same experiment, this time with the core sequence. As we know from
previous chapters, substituting Val for Tyr increases the saturation concentration and decreases the density of the dense phase. Thus, we hypothesized that this same substitution would create a similar effect in the core of a particle, increasing or decreasing the core occupied volume and perhaps changing the self-assembled structure. We thus made mutations to the RLP40-ELP80 and RLP80-ELP80 core repeat sequence, replacing (QYPSDGRG) with (GRGDP[Y]S) where the Tyr in the new repeating unit is systematically replaced with Val. What we observed for both core molecular weights that a systematic replacement of Tyr along the backbone results in a transition from more spherical particles to more elongated scattering structures (Figure 73). Specifically, with a core repeat number of 80, we observe all three transition states, spherical micelles that transition to worm like micelles or lamellae and vesicular structures. For this vesicular structure in particular, it appears to be in two phase equilibrium with worm like structures as there are additional areas of contrast throughout the various cryo-TEM imaging frames. The assembly of these vesicles is also quite broad, with various shapes, size, and multilaminar structures formed simultaneously
despite the monodispersity of the repeating unit.

Figure 73: Cryo-TEM images of RLP-ELP block co-polypeptides with varying core hydrophobicity. Scale bar = 500 nm. Data collected at 1 mg ml\(^{-1}\) in 140 mM PBS.

In addition to changes to the block architecture, one can affect the assembly of these dynamic molecules but affecting the solvent quality. A dramatic example of this affect can be seen in Figure 74, where modifying the buffer of (GRGDP\[Y:V\]S\)\(_{80}\)-ELP\(_{80}\) from 140 mM PBS to water - a poor solvent for the core and superior solvent for the corona – changes the assembly from vesicular/worm-like micelles to spherical micelles. This is further evidence to support our hypothesis that changes to the
chemical sequence affect the chain volume and thus changes in hydrophobicity also result in physical changes at the chain level.

Figure 74: Cryo-TEM images of (GRGDSP[Y:V]S)80-ELP80 block co-polypeptide in 140 mM PBS and distilled H$_2$O. Data collected at 1 mg ml$^{-1}$ and 15°C.

All of these modifications to the core and corona chains prompted us to measure the critical micelle concentration (CMC) of these micelle constructs. Thus, we employed a well-developed technique in the laboratory for measuring the CMC - encapsulating pyrene in the core of the particle. Depending on the polarity of the solution, pyrene will exhibit a different fluorescent signal of peaks 1 and 3 (I1 and I3). Upon a concentration above the CMC, pyrene will be sequestered by the core of the particle, modulating this
ratio of I1 and I3. For RLP40-ELP80 and RLP80-ELP80 we approximate a CMC that is between 100 and 500 nM according to the point of decreasing I1/I3 ratio (Figure 75) which suggests that these micelles are slightly more stable than previously measured block co-polypeptides[126]. Interestingly, this I1/I3 ratio has been tabulated for various solvents suggesting that the polarity of the interior of our particles are closer to that of acetone (1.4) than water (1.8). Thus, we hypothesized that these micelles may be capable of sequestering hydrophobic moieties similar to block co-polymer micelles. To this end, we designed a series of RLP40-ELP80 proteins that were intended to have similar assembly dimensions but different core chemistries. For the core sequence, we chose the repetitive units (GRGDSPYS), (GRGDSPYQ) and (GRGDQPYQ) which all have similar UCST binodal lines but different non-charged polar residues which may form different strength H-bonds with the drug. For the small molecule drug, we chose paclitaxel as it is notoriously insoluble in aqueous solvents and suffers side effects associated with its delivery vehicle.
We successfully cloned the sequences (GRGDSPYS)40-ELP80, (GRGDSPYQ)40-ELP80, (GRGDQPYQ)40-ELP80 and expressed the protein sequences recombinantly similar to other RLP-ELP80 block co-polypeptides. Cryo-TEM images suggested that we had successfully designed a system that exhibit similar, spherical assemblies with core diameters ranging from 33-36 nm (Figure 76).
Figure 76: Cryo-TEM images of RLP40-ELP80 block co-polypeptides where the core sequence contains varying amount of Ser and Glu. Data collected at 15°C, 1 mg ml⁻¹ in 140 mM PBS.

Using these particles, we incubated the polypeptide chains in the presence of 10X molar excess of paclitaxel (PTX) in H₂O or 30% acetone + H₂O mixture overnight at 4°C to maintain a soluble, assembled micelle in the presence of large excess of insoluble PTX (Figure 77). We then centrifuged out the still insoluble PTX in the H₂O sample and dialyzed out the acetone into milliQ H₂O, also removing the still insoluble PTX after dialysis by centrifugation. We then added 100% acetonitrile to a final concentration of 30% v/v to completely dissolves the micelle and resuspend the PTX into a homogenous mixture of PTX plus protein. This sample is run on a C12 analytic HPLC column to separate the protein and PTX peaks. Using the relative size of the peaks and
the known extinction coefficients of PTX and the protein at 230 nm and 275 nm respectively, we can determine a molar ratio of PTX to protein that remained soluble after we subtract from a no-vehicle control. In essence, this experience tells us how much the solubility of PTX increases in the presence of various micelle systems, with the supposition being that this increase in suspension concentration is due sequestration into the core of the particle.

![Figure 77: Schematic of paclitaxel loading of RLP40-ELP80 micelles and analytic procedure.](image)

Our results support the conclusion that the various RLP-ELP micelles are capable of increasing the observed solubility of PTX in normally unsuitable solvents. We observe a dramatic increase in loading when using acetone as a cosolvent, perhaps because it can diffuse readily into the core of the particle which has similar polarity to acetone. We also observed dramatic differences between the subtle particles
chemistries, suggesting that primary amino acid sequence can control this partitioning coefficient by 2 to 3-fold (Figure 78). In our best performing cases, 100% Gln substitution for Ser, we observed PTX/protein molar ratios that correspond to similar conjugation efficiencies achieved with PTX to protein (~8 PTX per protein chain in 100% Gln substitution case). Thus, we suggest that this physical loading procedure may provide an alternative delivery scheme opposed to direct chemical conjugation to Lys or Cys residues.

Figure 78: The relative molar ratio of paclitaxel (PTX) to RLP-ELP as determined by analytical high-performance liquid chromatography. After an area for each molecule was derived using the absorption peak for each molecule (230 nm for PTX, 275 for RLP-ELP80), this peak was normalized to the extinction coefficient of the molecule and then compared to one another.
4.3.2 UCST-LCST presentation of ligands

Using this RLP-ELP block co-polypeptide platform, we sought to develop a micelle capable of multivalent display. We selected a targeting domain, the 10th type III domain from human fibronectin (Fn3) that targets the human αvβ3 integrin, a receptor that is upregulated in the endothelium of many tumors[225-227] and is also overexpressed on several tumor cells such as glioblastoma[228], renal cell carcinoma[229], ovarian carcinoma[230] and breast cancer metastases[231]. We chose a Fn3 variant that binds the αvβ3 integrin with low affinity [232, 233] (K₀ > 1x10⁻⁷ M), and can be expressed in E. coli as a fusion to repetitive polypeptides such as ELPs[234]. The low affinity of the parent Fn3 domain is important as, multivalent presentation could amplify its avidity, which may not be possible with ligands that possess intrinsically high affinity, so that we could test for the effect of self-assembly and multivalency on binding avidity and cellular uptake[234, 235].

After assembly of the genes in the expression vector, each vector was transformed into the BL21(DE3) strain of E. coli and overexpressed by a previously published protocol. The block co-polypeptides were isolated from the soluble fraction of the cell lysate and purified by inverse transition cycling, a non-chromatographic method, to >95% purity as determined by SDS-PAGE (Figure 79). Yields of all polypeptides were >20 mg L⁻¹ of shaker flask culture without any optimization of the expression protocol, typical to other Fn3 expression and purification schemes that yield 5-20 mg L⁻¹ [237].
Figure 79: SDS-PAGE of RLPXX-ELP80-Fn3. Ladder units are in kilodaltons. Wells are labeled with the appropriate protein in the gel. All constructs have a band around 2X the molecular weight of the main band likely indicating the formation of dimers, in the presence of the gel loading buffer. Excluding this band, all materials are ≥95% pure.

Each block co-polypeptide was analyzed by dynamic light scattering (DLS) at several temperatures between 4 ºC and 37 ºC to determine the thermal stability of the micelles, and to determine their radius of hydration (R_h). The R_h of RLP20-ELP80 and RLP20-ELP80-Fn3 were ~7 nm, indicating that these constructs did not assemble within this temperature range and exist as soluble disordered polypeptides, as their R_h is similar to that of denatured proteins with a similar molecular weight (R_h ~ 8 nm) and other elastin like polypeptides of similar size[238-240]. In contrast RLP40-ELP80 and RLP40-ELP80-Fn3 self-assembled into micelles with a R_h of 30 and 32 nm, respectively, between 20-37°C (Figure 80).
Figure 80: Thermal stability of RLP-ELP-Fn3 micelles. Spherical (RLP\textsubscript{40}-ELP\textsubscript{80}-Fn3) and worm-like micelle (RLP\textsubscript{80}-ELP\textsubscript{80}-Fn3) stability between room temperature (20°C) and physiological temperature (37°C). Data collected at 10 µM in 140 mM PBS. Filtered with 0.45 µm filter.

Likewise, RLP\textsubscript{80}-ELP\textsubscript{80} (112 nm) and RLP\textsubscript{80}-ELP\textsubscript{80}-Fn3 (47 nm) formed stable micelles over the same temperature range (Figure 81).
Figure 81: Thermal stability of block co-polypeptide micelles. Spherical (RLP40-ELP80) and worm-like micelle (RLP80-ELP80) stability between room temperature (20°C) and physiological temperature (37°C). Data collected at 10 µM in 140 mM PBS. Filtered with 0.45 µm filter.

Interestingly, the Rh of RLP80-ELP80 is dramatically affected by the presentation of the Fn3 domain on the hydrophilic C-terminal end of the block co-polypeptide (Table 11).

Table 11: Light scattering data of Fn3 block co-polypeptides

<table>
<thead>
<tr>
<th>Block co-polypeptide</th>
<th>( R_g ) (nm)</th>
<th>( R_h ) (nm)</th>
<th>( N_{agg} )</th>
<th>( \rho = R_g/R_h )</th>
<th>Shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLP_{20}-ELP_{80}-Fn3</td>
<td>-</td>
<td>6.95</td>
<td>-</td>
<td>-</td>
<td>Unimer</td>
</tr>
<tr>
<td>RLP_{40}-ELP_{80}-Fn3</td>
<td>29.3</td>
<td>29.3</td>
<td>201</td>
<td>1.0</td>
<td>Spherical</td>
</tr>
<tr>
<td>RLP_{80}-ELP_{80}-Fn3</td>
<td>39.2</td>
<td>48.7</td>
<td>630</td>
<td>0.8</td>
<td>Sphere &amp; Wormlike</td>
</tr>
</tbody>
</table>
This result makes sense as RLP80-ELP80 exists on the edge of the phase boundary that separates spherical and worm-like micelles[238]. Therefore, it is plausible that the incorporation of a small folded protein could result in a change of shape. It also appears that the Fn3 domain is not stable at temperatures above 37°C, as there is a precipitous increase in the $R_h$ of RLP40-ELP80-Fn3 between 36-40°C. Based on this result, samples were maintained on ice prior to flow cytometry and confocal microscopy.

Based on our previous results, we speculated that micelles with a $R_h$ in the 30-40 nm range are likely to be spherical, while micelles with a $R_h >$100 nm are likely to be cylindrical or worm-like in structure. To deduce the morphology of these particles and to calculate their aggregation number ($N_{agg}$), we next carried out static light scattering (SLS) measurements. Increasing the size of the core-forming block from 40 to 80 repeats of (QYPSDGRG) increases the radius of gyration ($R_g$) from 29 nm to 39 nm, the $R_h$ from 29 nm to 49 nm and the $N_{agg}$ from 201 to 630 chains per micelle (Table 11 & Figure 82).
Figure 82: Static and dynamic light scattering raw data for RLP-ELP block copolypeptides. A-D. Plot of $R_h$ vs. angle, extrapolated to $0^\circ$ for reported $R_h$ of A. RLP20-ELP-Fn3, B. RLP40-ELP-Fn3-10, C. RLP80-ELP-Fn3. D,E. Partial Zimm plot obtained by static light scattering of D. RLP40-ELP-Fn3, E. RLP80-ELP-Fn3.

These results suggest that the larger particles have a higher aspect ratio than the smaller particles. Unfortunately, the SLS results were not conclusive, as the form factor
(\(\rho = \frac{R_g}{R_h}\)) did not change dramatically across particles with putatively different morphologies. Typically, \(\rho\) depends on the morphology of the particles with typical values for spheres around 0.7 and increases as the scattering molecule becomes more elongated (i.e. disc shaped, cylindrical structures)[241]. Therefore, we next directly visualized the particles with cryo-TEM to confirm their morphology.

Previous studies of ELP-based micelles have demonstrated that the desolvated core of the micelle can be visualized by cryo-TEM, as it has significant differential contrast than the surrounding water, but the corona is far too solvated to be visualized[238]. Increasing the core-forming RLP block size from 40-80 units without a Fn3 domain (Figure 83A,B respectively) shifted the morphology from spherical to worm-like micelles, as reported previously. The core of these micelles increased in diameter (Table 12) from 24 nm to 59 nm and the spacing between the particles changed from 27 nm to 42 nm, indicating the elongation of the corona ELP chain. RLPXX-ELP80-Fn3s behave similarly. Increasing the core size increased the core diameter of the micelles from 27 nm to 51 nm and the core spacing from 27 nm to 42 nm (Figure 83C,D respectively).
Figure 83: Cryo-TEM micrographs of RLPXX−ELP80 and RLPXX−ELP80-Fn3 A. spherical micelles formed by RLP40−ELP80, B. worm-like micelles formed by RLP80−ELP80, C. spherical micelles formed by RLP40−ELP80-Fn3, D. spherical and worm-like and spherical micelles formed by RLP80−ELP80-Fn3. All scale bars represent 200 nm. All data collected at 15°C in 140 mM PBS at 10 µM.

Table 12: Cryo-TEM core diameter and spacing of RLPXX-ELP80-Fn3s

<table>
<thead>
<tr>
<th>Block co-polypeptide</th>
<th>Core Diameter (nm) [n=30]</th>
<th>Intra-core Spacing (nm) [n=30]</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLP_{40}-ELP_{80}</td>
<td>24.3 ± 2.8</td>
<td>23.6 ± 4.5</td>
</tr>
<tr>
<td>RLP_{40}-ELP_{80}-Fn3</td>
<td>27.9 ± 4.3</td>
<td>27.4 ± 4.7</td>
</tr>
<tr>
<td>RLP_{80}-ELP_{80}</td>
<td>39.4 ± 5.4</td>
<td>35.0 ± 4.5</td>
</tr>
<tr>
<td>RLP_{80}-ELP_{80}-Fn3</td>
<td>44.7 ± 8.0</td>
<td>34.4 ± 5.8</td>
</tr>
</tbody>
</table>
Image analysis suggests a shift in assembly as RLP40-ELP80-Fn3 has >90% of particles with an aspect ratio <2 while RLP80-ELP80-Fn3 has 45% of micelles with an aspect ratio <2 (Figure 84).

Figure 84: Histogram of Observed Aspect Ratios of RLPXX-ELPYY-Fn3s. Using the data from Figure 83, particularly the longest straight line of an individual particle and a corresponding perpendicular measurement were made to describe the aspect ratio of particles observed. Only particles that were clearly individual particles were used for the analysis (equal contrast around the particle, suggesting a camera-normal orientation). n = 50.

These results both indicate a shift in morphology from spherical to a mixture of spherical and worm-like micelles. These results also corroborate the DLS and SLS experiments measurements indicating that increasing the core block length elongates the micelle morphology, increasing the density of chains in the corona and maintaining the overall shape of the parent block co-polypeptide.
We next used surface plasmon resonance (SPR) to characterize the avidity of the RLPXX-ELP80-Fn3 fusions to the ectodomain of human αvβ3 integrin. SPR sensorgrams were generated for binding of the Fn3-functionalized RLP-ELP80 block copolypeptides at concentrations ranging between 2.5 and 10 µM. Kinetic association and (k\text{on}) and dissociation constants (k\text{off}) are summarized in Table S3. As the core block size increases, the k\text{on} increases in magnitude and the k\text{off} decreases in magnitude, both consistent with the increase in size of the binding unit (unimer or larger diameter micelle)[242]. As seen in previous studies[192, 243], Fn3-decorated spherical micelles showed a 10-fold increased avidity for the αvβ3 integrin compared to the RLP20-ELP80-Fn3 construct that does not self-assemble and hence only presents a single copy of the Fn3-domain. Interestingly, elongating the particle from a spherical to worm-like geometry can increases the avidity for the integrin by ~1000-fold compared to the monomer ligand, driving avidity into picomolar concentrations (Figure 85). This result is remarkable when one considers the unoptimized nature of the Fn3, which has K\text{D} in the micromolar range for the αvβ3 integrin. The effective K\text{D} of the RLPXX-ELP80-Fn3 worm-like micelles is in fact is many orders of magnitude lower than a clinically relevant therapeutic antibody —LM609— which has a K\text{D} of ~20 nM[244]. For context, these binding constants are at the upper threshold of antibodies that are used for targeted cancer therapy targeting, highlighting their clinical relevance[245].
Figure 85: Shape dependent avidity of RLPXX-ELP80-Fn3. Multivalency increases the observed $K_D$ as does increasing the aspect ratio of the micelle. Representative SPR sensor grams shown on top show a marked decrease in $k_{off}$ between unimer, and spherical and worm-like micelles. In contrast the $k_{on}$ is similar for all constructs interest. SPR sensorgram data collected in PBS at 10 µM.

<table>
<thead>
<tr>
<th>Block co-polypeptide</th>
<th>Shape</th>
<th>$k_{on}$ (L/(mol·s))</th>
<th>$k_{off}$ (s⁻¹)</th>
<th>$K_D$ (nmol/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLP$<em>{20}$-ELP$</em>{60}$-Fn3</td>
<td>Unimer</td>
<td>4790</td>
<td>4.57E-03</td>
<td>954 nM</td>
</tr>
<tr>
<td>RLP$<em>{40}$-ELP$</em>{60}$-Fn3</td>
<td>Sphere</td>
<td>4245</td>
<td>3.35E-04</td>
<td>78.9 nM</td>
</tr>
<tr>
<td>RLP$<em>{60}$-ELP$</em>{60}$-Fn3</td>
<td>Worm</td>
<td>4065</td>
<td>3.20E-06</td>
<td>0.786 nM</td>
</tr>
</tbody>
</table>

To assess the intracellular uptake of these particles, we used a cell line stably transfected with the ανβ3 integrin. The native cell line, K562, has endogenously low levels of expression of this receptor and therefore serves as the receptor-negative control, and the un-decorated RLPXX-ELP80 micelles serve as ligand-negative controls for each type—size and shape—of micelle. Cells were incubated for 2 hours with a 10 µM solution of various block co-polypeptides at 37°C, a concentration that is well above the CMC and $K_D$ of all micelles. Confocal microscopy was first used to study the internalization of the block co-polypeptides by the ανβ3 integrin transfected cell line.
Ligand-negative spherical micelles showed low levels of uptake, while that of ligand-negative worm-like micelles was slightly higher (Figure 86A), consistent with previous observations that shape plays a role in controlling non-specific uptake of nanoparticles[246, 247].

However, far more dramatic differences were seen for Fn3-decorated micelles. Compared to the parent spherical micelle which showed low levels of uptake (Figure 86A) that was barely above that of the autofluorescence of the WT —untransfected— cell line, presentation of the Fn3 domain on the RLPXX-ELP80 block co-polypeptide that forms spherical micelles significantly increased its uptake (Figure 87) as quantified by the number of particles inside the cell membrane (Figure 86B) and the mean fluorescence of the cell —(Figure 3B p < 0.001 unpaired student’s t-test).
Figure 87: Representative images of cellular uptake of RLPXX-ELP block copolypeptides labeled with Alexa488 fluorophore (green) overlaid with DIC images after 2.5 hr of incubation in serum free minimal media at 10 µM. Scale bar = 20 µm.

The worm-like micelles that are decorated with the Fn3-ligand similarly showed a much greater level of cell uptake compared to the parent worm-like micelles (Figure 87). In contrast, without overexpression of the αvβ3 integrin on K562 cells, there were low levels of internalization and uptake of the spherical and Fn3-decorated micelles, showing that most of internalization of ligand-decorated micelles is driven by ligand-receptor engagement (Figure 88).

Figure 88: Cellular uptake of RLPXX-ELPYY-Fn3 Polypeptides in αvβ3 Negative K562 Cell Line. A-D Representative images of cellular uptake of block polypeptides labeled with Alexa488 fluorophore (green) overlaid with DIC images.
(grey) after 2.5 hr of incubation in serum free minimal media at 10 µM. A. LM609 Antibody B. RLP20-ELP80-Fn3 C. RLP40-ELP80-Fn3 D. RLP80-ELP80-Fn3. Scale bar = 20 µm.

The LM609 antibody showed completely different cell uptake than RLP-ELP80-Fn3 micelles. Although it has a high level of fluorescence (Figure 87), much of the fluorescence was localized at the cell membrane and far lower levels of intracellular fluorescence, especially compared to the Fn3-decorated micelles, indicating that this antibody-integrin binding event does not trigger internalization.

Flow cytometry was next used to quantify the cell uptake. Unstained K562 cells had a background cell fluorescence of 2912 ± 3236 (Geometric Mean ± StDev) (Figure 89) which increased to 8686 ± 8787 when incubated with RLP40-ELP80 spherical micelles and to 24904 ± 13884 for the RLP80-ELP80 indicating that there is a low level but shape-dependent non-specific uptake of the micelles (p < 0.001, unpaired student’s t-test) (Figure 3B).
The positive control, the LM609 antibody had a statistically significantly higher uptake of $36708 \pm 255175$, consistent with its known specificity for the $\alpha v \beta 3$ integrin (Figure 90). A closer look at the flow cytometry data indicates that there is a high-level and low-level receptor expressing cell population, as seen by the two distinct peaks in Figure 89. Interestingly, spherical micelles formed by RLP40-ELP80-Fn3 and worm-like micelles formed by RLP80-ELP80-Fn3 have higher geometric fluorescent intensity means of $15539 \pm 286229$ and $71382 \pm 251919$ that are 2-fold and 3-fold greater than the undecorated controls (Figure 89).
Figure 90: Quantification of cellular uptake by flow cytometry. *** = p < 0.001. Box indicates 25th and 75th percentile and bars indicate 10th and 90th percentile.

Clearly receptor-mediated endocytosis is shape dependent, as seen by the significantly higher cell uptake exhibited by the worm-like micelles compared to spherical micelles, that is also consistent with their higher avidity for the integrin. The Fn3-decorated spherical and worm-like micelles also only exhibited a single flow cytometry peak, unlike LM609 that has a bimodal distribution of cell uptake. This result implies that high valency micelles are not sensitive to the heterogeneity of receptor expression, presumably as long as the receptor expression is above a certain threshold to enable multiple ligands to engage the receptors on the cell surface. High valency micelles may therefore provide more robust strategy to target cells with inhomogeneous levels of receptor expression than antibodies.
Figure 91: Cellular uptake of RLPXX-ELPYY-Fn3 polypeptides with variable aspect ratio in αvβ3 transfected cell line. A-D Representative images of cellular uptake of block polypeptides labeled with Alexa488 fluorophore (green) overlaid with DIC images (grey) after 1.5 h of incubation in serum free minimal media at 10 µM. A. RLP80-ELP80-Fn3 B. RLP80-ELP160-Fn3 C. RLP40-ELP80-Fn3 D. RLP40-ELP40-Fn3. There is a much lower level of uptake of all constructs with spherical morphologies (B,C) versus particles with elongated morphologies (A,D). Scale bar = 20 µm.

We believe that morphology is more important than size, as worm-like micelles with the same hydrophilic weight fraction as RLP80-ELP80-Fn3, but that are smaller in size, exhibit higher levels of cell uptake than a spherical micelle of comparable size- RLP40-ELP80-Fn3 (Figure 92). Likewise, spherical particles of a similar size than the worm-like micelle of RLP80-ELP80-Fn3 exhibit very low levels of uptake (Figure 91). These data indicate that the elongated shape and flexibility of the worm-like micelles increased the number of accessible Fn3 ligands available to bind the receptor.
Figure 92: Cryo-TEM Characterization of “shape control” RLPXX-ELPYY-Fn3s. A. spherical micelles formed by RLP80–ELP160–Fn3, B. spherical and worm-like micelles formed by RLP40–ELP80–Fn3. All data collected at 15°C in 140 mM PBS at 10 μM.

We next decided to visualize the kinetics of internalization by imaging the cells at 20, 45, 90, 120 and 240 min post-incubation (Figure 93). Particle and area analysis of the Alexa488 dye was performed on all cells in the visual field for at least 3 separate images resulting in ~50 individual measurements for each sample. The analysis area was gated to exclude the cell membrane to eliminate non-internalized areas of fluorescence.
Figure 93: Cellular uptake of block co-polypeptides over time. Representative confocal images of antibody (LM609), RLP20-ELP80-Fn3, RLP40-ELP80-Fn3 & RLP80-ELP80-Fn3 uptake as a function of time. Scale bar = 20 µm.

Compared to the LM609 antibody, that remains largely associated with the cell membrane, with only a few isolated fluorescent puncta within the cell at later points, spherical (RLP40-ELP80-Fn3) and worm-like micelles (RLP80-ELP80-Fn3) are internalized faster, resulting in more particles within the cell at all time points (Figure 94A).
Using a 3-way ANOVA for time, shape and decoration state (Fn3 +/-) we observed a main effect of shape, decoration state and time for both the number of particles and the area which they cover in the cell. Pairwise interactions indicate that there are significant effects of micelle shape on the number of particles per cell between spherical and worm like micelles ($p < 0.01$) and spherical micelles and the positive — antibody— control ($p < 0.05$). There were significant differences in the area these particles occupied in the cell between spherical micelles and the antibody control over time ($p < 0.05$) and worm-like micelles and the antibody control ($p < 0.05$) over time (Figure 94B). There was not a significant difference in particle area between spherical and worm-like micelles over time. The overall percentage of the cell population that contained fluorescent signal was evaluated but did not produce a significant effect over time. The only significant effects of this ANOVA analysis were pairwise effects of micelle shape over time ($p < 0.001$) and decoration state over time ($p < 0.01$) (Table S4, Table
S5, Table S6). In summary these data show that: 1) there is a statistically significant increase in cellular uptake with respect to particle morphology 2) a statistically significant increase in cellular uptake by micelles that present an integrin-binding Fn3 domain.

4.3.3 UCST-UCST block co-polypeptide assembly

Considering the large library of RLPs generated in Chapter 2 and the vast difference in transition temperatures observed, we set out to be the first group to create the first block co-polypeptide made of two UCST blocks. From our work with ELP block co-polypeptides and RLP-ELP block co-polypeptides we know that there must be a large difference in transition temperatures of the two blocks as each block will influence the other, driving the transitions of each individual block closer together. To begin, we fused the two RLPs with a large difference in $T_t$, [S]-40,80 and GRGDQPHN ([QHN]-40). We also varied the block length of the core forming block, as previous experiments demonstrated the importance of the overall hydrophilic weight fraction on the assembled morphology. We also varied the core block by changing GRGDQPHN to GRGDNPHQ ([NHQ]-40). This is a hydrophobic change but maintains the same overall composition of the polypeptide.

The first notable observation is that the choice of these two RLP sequences resulted in variable self-assembly. RLPSS-40 – RLPQHN-40 assembled into an identifiable nanoscale morphology with a $R_g$ of 35.1 nm, $R_h$ of 28.4 nm and a $N_{agg}$ of 43 (Table 13). The form factor ($R_g/R_h$) suggests that [S]-40 – [QHN]-40 assembles into worm-like micelles. Increasing the molecular weight of the core block increases the $R_g$,
Another interesting observation is that the simple change between \([S]-XX-[QHN]-40\) and \([S]-XX-[NHQ]-40\) resulted in disassembly. This result was unexpected as the difference in hydrophobicity between \([QHN]\) and \([NHQ]\) is rather small. However, this result can be understood in the sense that this small change narrows the \(T_t\) gap such that the two blocks have similar enough \(T_t\). The similar enough \(T_t\) between the two blocks results in co-polypeptide behaving as a unimer unit, with just one UCST temperature.

To gain more insight into the morphology of the UCST-UCST constructs, samples were prepared for cryo-TEM. We knew from previous experiments that the vitrification process increases the concentration of the solution. In these images, we observe that the sample polypeptide has undergone liquid like phase separation due to the concentration increase. This phase separation makes it impossible to discern the nanostructure of the two constructs. However, what is interesting is that within the liquid
like droplet, there appears to be an interconnected structure (Figure 95). There also appears to be three distinct areas of contrast – the water/buffer that surrounds the droplets, the contrast of the droplets themselves and then the further contrast of the internal microstructure. This indicates that different RLPs have different contrast upon their phase transition, indicating that RLP-ELPs and RLP-RLPs with different RLPs at the core, may have measurable differences in core contrast. Both [S]-40 – [QHN]-40 (Figure 95A,C) and [S]-80 – [QHN]-40 (Figure 95B,D) form this microstructure.

Figure 95: A. [S]-40 – [QHN]-40 cryo-TEM image. Scale bar 500 nm. B. [S]-80 – [QHN]-40 cryo-TEM image. Scale bar 500 nm. C. [S]-40 – [QHN]-40 cryo-TEM image. Scale bar 200 nm. D. [S]-80 – [QHN]-40 cryo-TEM image. Scale bar 200 nm. All constructs were vitrified at 2mg ml⁻¹, 100% humidity, 37°C in 140mM PBS.
As previously, the temperature dependent turbidity was determined with UV-Vis spectrophotometry. Utilizing this method, we were unable to visualize a clear soluble to assembled transition upon cooling. Additionally, both [S]-40 – [QHN]-40 and [S]-80 – [QHN]-40 had clear UCST aggregation temperatures. For both we determined the concentration dependence of this UCST aggregation (Figure 96A). Both constructs have much higher UCST values than would be predicted by just the corona block alone, indicating that the fusion to the more hydrophobic RLP brings the observed UCST behavior somewhere between the two unimer blocks alone. Increasing the size of the RLP increased the UCST, indicating that the corona is sensitive to the size of the attached core polypeptide. In comparison to RLP-ELP block co-polypeptides, RLP-RLP block co-polypeptides retain much of their concentration dependence which explains our cryo-TEM results. It is also interesting that [S]-80 – [QHN]-40, which would be predicted to be the more worm-like of the two, has higher concentration dependence. This is also different than the previously observed trend with RLP-ELP block co-polypeptides.
Figure 96: A. UCST phase behavior of [S]-40-[QHN]-40 and [S]-80-[QHN]-40 block co-polypeptides as determined by UV-Vis spectrophotometry. B. pH effect on UCST behavior of RLP-RLP block co-polypeptides as observed via temperature dependent DLS. Data is taken at 2mg ml^{-1} in 140mM PBS where the UCST of both block co-polypeptides is very similar and hence a similar pH triggered UCST deflection is observed.

RLP-RLP block co-polypeptides retain the unique pH responsiveness of the corona unimer. Temperature dependent DLS measurements in different buffered pH conditions demonstrated a maximum in the observed UCST aggregation temperature, again around the isoelectric point of His (Figure 96B). Both constructs UCST increases in almost a linear fashion from pH 8.4 to pH 6.4 and then decreases from pH 6.4 to 3.4. In the context of what was observed earlier, this result makes sense. As the pH decreases towards the isoelectric point of His, there is an increase in the $T_t$ because the corona is behaving more hydrophobic. Although this process is not understood, it remains consistent with previous observations. After reaching the isoelectric point, the $T_t$ decreases because the massive amount of positive charge in the corona chain which
would dramatically increase chain repulsion. It is important to note, that in each of these aggregation DLS curves, that a stable micelle regime was observed. [S]-40 – [QHN]-40 did not appear to have a large change in \( R_h \) as the pH decreased to the isoelectric point indicating that there was likely not a change in the morphology of the structure. [S]-80 – [QHN]-40 did exhibit a size change as the pH decreased towards the isoelectric point and then actually dropped below the original size measured at pH 7.4 at pH 3.4. This would indicate a morphology change from worm at neutral pH to a more elongated worm at pH 5.4 and then possibly to a sphere at pH 3.4.

Our UV-Vis measurements did not answer important questions about the temperature dependent disassembly. Therefore, to get a more accurate picture, we monitored the \( R_h \) as we slowly cooled the solution. We observed two different overall behaviors. [S]-40 – [QHN]-40 upon cooling underwent a clear unimer to micelle transition at ~55°C and a clear micelle to aggregate transition @ 28°C (Figure 97B). A unimer to micelle transition that is temperature dependent is known as the critical micellization temperature (CMT). As we know from before, this UCST aggregation temperature is dependent on the solution concentration. [S]-80 – [QHN]-40 remained a micelle at the highest temperature we could measure with the instrument. Upon cooling the size of the aggregate increased slightly until a clear aggregate transition at 18°C. These two results indicate that controlling the core block sequence provides control over the aggregate UCST and the CMT.
Figure 97: A. Critical micelle concentrations (CMCs) of [S]-40-[QHN]-40 and [S]-80-[QHN]-40 determined to be 3 µM and 0.4 µM respectively by a shift in I1/I3 of pyrene fluorescence. Sigmoidal fit to triplicate data is shown. CMC is determined by the inflection point of the sigmoidal fit. B. Full thermal characterization of [S]-40-[QHN]-40 and [S]-80-[QHN]-40 indicates that increasing the core block shifts the disassembly UCST phase behavior.

In addition to these block co-polypeptides, we sought to understand the difference in Tt required for the assembly of UCST-UCST micelles. Using the library of repetitive IDPs from Chapter 2, we fused progressively more hydrophilic blocks to our core sequence [S]-40, by replacing Tyr residues with Val. We know that this is the most efficient substitution, an aliphatic reside for an aromatic residue. Additionally, we know that [S]-40 is above the minimum core block size required for assembly. Thus, first we created three block co-polypeptides with ~50% hydrophilic weight fraction where the intended corona chain contains the sequences [Y:V]-40, [Y:3V]-40 and [V]-40 and the core is comprised of [S]-40.

UV-vis spectrophotometry and dynamic light scattering (DLS) measurements were quite instructive in determining the minimum difference required for self-assembly.
In these experiments, it became clear to us that coronas comprised of [Y:V] and [Y:3V] only resulted in particulate systems that do not assemble but aggregate in a way approximating liquid-liquid coacervation. Only when Tyr was completely replaced by Val did we observe self-assembled structures. The DLS and UV-vis spectrophotometry show that upon cooling there is an intermediary phase of assembly where the unimer sequences (~10 nm) transition into ~500 nm particles before settling into stable 30 nm micelles (Figure 98). Here, unlike our [S]-40 and [QHN]-40,80 constructs, we only observe a single mode of assembly with DLS suggesting that the core collapses, but the corona chains remain soluble across the entire temperature range. These results also suggest that a minimum different of the core block is 12.5% with an approximate difference in Ti of 80-100°C.

Figure 98: UV-Vis spectrophotometry and dynamic light scattering of [S]-40-[V]-40 and [S]-40-[Y:3V]-40 block co-polypeptides.

These experiments give us a concept of the core and corona necessary for assembly. From our experiments with UCST and LCST diblocks we know that the relative block sizes influence the assembly of the micelles. Therefore, to investigate the
assembly size, we changed the corona size and evaluated their assembly by cryo-TEM. The cryo-TEM images show that our first construct, [S]-40-[V]-40 assembled into a mixture of small micelles and large phase separated domains. These large phase-separated domains grow as the corona size decreases (Figure 99). Likewise, increasing the corona size decreases the size of the phase separated domains with a larger percentage of the field of view forming small spherical particles of ~30 nm Rₚ.

![Cryo-TEM images of [S]-40-[V]-20, [S]-40-[V]-40, [S]-40-[V]-60. Data collected at 15°C in 140 mM PBS. Scale bar = 500 nm.](image)

Finally, combining these two insights we made a third systematic library where we change ~10% of the corona chain with an aliphatic amino acid but one with varying hydrophobicity as predicted by other hydrophobicity scales. We know from Chapter 2 that Ala, Ise, Val all exhibit similar effects on the UCST phase separation behavior of polypeptides, the only difference between each of these amino acids being extra hydrocarbons on the side group. Thus, we made the proteins [S]-40-[I]-40 and [S]-40-[A]-40 in addition to our existing protein [S]-40-[V]-40.
These cryo-TEM images suggest that decreasing the hydrophobicity of the chain eliminates the presence of phase separate domains and shifts the assembly phase into a mixture of worm like micelles and spherical micelles, including the presence of a few vesicles. Increasing the hydrophobicity of the corona increases the size and hydrophobicity of the phase separated domains so that they associate with the hydrocarbon grid.

4.4 Conclusions

Our employment of a core UCST block was overall very successful in creating predictable self-assembling block co-polypeptides. The avenues of mutation that we pursued were governed by simple principles of diblock assembly from polymer physics and generally resulted in the effects predicted by the theory. Unlike previous protein assembly systems, whose behavior is more difficult to a priori predict, the self-assembly of RLP-ELP block co-polypeptides can be understood by the hydrophilic weight fraction,
polypeptide-polypeptide and polypeptide-solvent interactions. This finding is crucial because provides a route for the de novo design of desired nanoscale morphologies from first principles, into a wide variety of shapes and sizes.

Using these block co-polypeptides as scaffolds for assembly, our results clearly show that RLPXX-ELP80 block co-polypeptides are a robust platform for multivalent display of Fn3 domains via self-assembly. The morphology of the parent micelles—spherical versus worm-like—can be tuned by modulating the block ratios and the molecular weight of the core. Decreasing the hydrophilic weight fraction from ~0.7 to ~0.46 to ~0.30 changes the morphology from unimers to spherical micelles to worm-like micelles, respectively. Importantly, the gene-level fusion of Fn3 domain that targets the αvβ3 integrin at the hydrophilic, C-terminal end of the block RLP-ELP co-polypeptide does not abrogate self-assembly and enables the high-density presentation of a Fn3 domain on the corona of the micelles. Fn3 presentation does, however, have an impact on morphology, as the parent micelle, RLP80-ELP80, that exists on the phase boundary between spherical and worm-like micelles, converts to worm-like micelles upon presentation of the Fn3 domain on the corona of the block co-polypeptide.

Cell uptake studies of Fn3-presenting RLP-ELP block co-polypeptides with an αvβ3 overexpressing cell line yielded four notable results: first, compared to the parent—ligand negative micelles—that in the best case—a worm-like micelle—has 3-fold greater cell uptake at 2 hours, demonstrating that multivalency can greatly enhance the targeting potency of a ligand simply by virtue of the avidity effect. Second, compared to a RLPn-ELP-Fn3 fusion that does not self-assemble into a micelle and hence only
presents a single copy of the Fn3 domain that target the αvβ3 integrin, we found that multivalent spherical and worm-like micelles have an higher avidity and greater cellular uptake, showing the importance of multivalency in amplifying the avidity of a ligand by presentation of multiple copies on a nanoscale scaffold. Third, we observed a dramatic difference in cell uptake as a function micelle morphology, where Fn3-decorated worm-like micelles showed a 5-fold increase in cell uptake compared to spherical micelles. Fourth, we believe that morphology is more important than size, as worm-like micelles with the same hydrophilic weight fraction as RLP80-ELP80-Fn3, but that are smaller in size, exhibit higher levels of cell uptake than a spherical micelle of comparable size (RLP40-ELP80-Fn3). Likewise, spherical particles of a similar size than the worm-like micelle of RLP80-ELP80-Fn3 exhibit very low levels of uptake. These data indicate that the elongated shape and flexibility of the worm-like micelles increased the number of accessible Fn3 ligands available to bind the receptor. Fifth, the avidity and cell uptake of the best performing worm-like micelles is greater than a therapeutically relevant antibody that targets the same receptor.

This class of self-assembling RLP-ELP block co-polypeptides provide an exceptionally robust and versatile system for the molecular design and recombinant synthesis of micelles for delivery of drugs and imaging agents for the following reasons, compared to other ELP-based nanostructures[182, 234, 248-253]: First, RLP-ELP block co-polypeptides, unlike ELP block co-polypeptides[124], follow canonical rules of polymer self-assembly via genetic encoded sequences, which make it easier to program their morphology de novo for specific applications. Second, these micelles have
significantly greater thermodynamic stability than ELP micelles, as they have CMCs in
the ≤ 0.1 µM range, compared to the 5-10 µM CMC of ELP micelles. Third, these
micelles enable presentation of a Fn3 domain on their corona, which is an attractive
choice as a targeting ligand, as the Fn3 scaffold is an enormously mutable targeting
scaffold and allows variants to be discovered by library screening approaches against
diverse targets. Fourth, we note that these targeted micelles can be loaded with drug
simply by conjugation of small molecule drugs into the core-forming, hydrophobic
domain, in a manner similar to our previous ELP micelles[182, 248, 249]. Finally, their
manufacturing — and hence clinical translation — can leverage the bacterial
fermentation and downstream purification capabilities of the biopharmaceutical industry.

5. Effect of repetitive polypeptide design on \textit{in vivo} release of glucagon like peptide 1 (GLP-1) from subcutaneous depots

5.1 Introduction

In the previous chapters of this thesis we have utilized the phase behavior of the
repeat polypeptide, chemically inspired from naturally occurring IDPs, to control the bio-
availability of resources when confined by a lipid bi-layer in bacteria and to control the
assembly of micelles when sterically confined by a relatively hydrophilic ELP or RLP
molecule exerting a surfactant like effect. We were motivated to test the efficacy of
controlling bio-availability in a system where the dilute phase is attached to an infinite
sink – where the dilute phase is in equilibrium with a biological system capable of protein
clearance. The Chilkoti lab has extensive experience with subcutaneous delivery of
184
peptide molecules in vertebrate animals in this exact set up, the primary focus on delivering therapeutic molecules for diabetes and various cancers, albeit exclusively with LCST polypeptides [101, 254, 255].

The delivery of peptides remains an outstanding challenge for drug delivery [256-258]. Despite protein engineering improvements focused on improving half-life, their effective window lasts from minutes to a few hours rendering them unsuitable for therapeutic use [259, 260]. Interestingly, nature utilizes peptides in various biological applications, but regulation of their activity is often tightly controlled by a cellular population that can react to a phenotype change[261, 262]. Thus, man-made peptide drugs require a delivery solution, one that can improve the pharmacokinetics of these valuable macromolecules.

The most common approaches to improve a peptide's half-life include protein engineering and changing the formulation to prolong release and/or reducing renal clearance. Sequence engineering —such as the incorporation of D-amino acids or other chemically esoteric amino acid derivatives— can limit proteolytic degradation of the protein[263], but severely limit manufacturing choices. Encapsulation methods produce inconsistent effects on bioavailability or require harsh production conditions that limit the type of peptide drug that can be delivered via these methods[264]. Strategies to decrease renal clearance revolve around increasing the size of the molecule and reducing opsonization including attachment to synthetic or biological polymers [265-267] that extend half-life, fusion to a large protein, and conjugating chemical moieties that allow the peptide to piggyback on endogenous biomolecules with slow turnover rates like
albumin [268, 269] or antibody fragments [270, 271]. These strategies are not without limitations as they dramatically reduce potency and rely on a patient population to express these piggybacking biomarkers consistently between individuals.

Our engineered polypeptides offer an elegant solution – through their primary amino acid sequence and molecular weight they control the dense and local dilute phase of the bioactive molecule effectively controlling the bioavailability of the drug. Unless proteases possess the unique ability to diffuse into the dense phase of the polypeptide-peptide drug depot, the availability of peptide drug is mediated exclusively by the primary amino acid sequence of the polypeptide. This local dilute phase can then diffuse and convect from the subcutaneous space into circulation and exert a therapeutic effect. Using the principles of polypeptide design described in Chapter 2, we can rationally design drug release depots that can prolong the half-life of the peptide drug \textit{in vivo}.

\subsection*{5.1.1 Previous work}

As mentioned previously, our lab has extensive experience prolonging the delivery peptide drugs. For this study, the relevant peptide drug is GLP-1, which is a 31 amino acid peptide produced in the L cells of the intestines, capable of exerting blood glucose control over a large therapeutic window [272, 273]. Previous experience with GLP-1-polypeptide depots through the works of Miriam Amiram and Kelli Luginbuhl has revealed much about the design of sub-cutaneous depots for drug release. 1) Fusion of a macromolecule such as the polypeptide ELPs reduce potency of the GLP-1 molecule by about \(~30\) fold but this does not preclude \textit{in vivo} activity. 2) Zero order release can be
achieved for up to 10 days in mice and 17 days in monkeys under optimal conditions. 3) Optimal conditions are an injectable transition temperature 5-7°C below the body temperature of the animal and a molecular weight of 35 kDa or greater to avoid renal clearance. This optimal 5-7°C below body temperature corresponds to a dilute phase concentration of approximately 1-100 µM where non-optimal depots exhibited $C_{\text{sats}}$ an order of magnitude above and below this optimal range.

5.1.2 Approach

The peptide drug of choice for these experiments is GLP-1 for several reasons. 1) GLP-1 can rapidly exert a therapeutic effect in vivo. 2) GLP-1 is a prime candidate for improved pharmacokinetics with a half-life of ~5 min in vivo. 3) GLP-1 can be easily studied in established mice models of diet induced obesity where a high fat diet increases the blood glucose. 4) GLP-1 is a stable peptide drug which will eliminate confounding variables associated with genetic fusion to various polypeptide partners and myriad delivery strategies.

Previous studies in our lab suggest that transition temperature at injection concentration is the important parameter for determining efficacy. However, this misjudges the important isotherm on the phase diagram to be room temperature instead of the operating temperature of the depot which is defined by the animals resting body temperature. In subcutaneous mice models, this temperature is approximately 35°C. Thus, using language from Chapter 3, we are looking proteins with variable dilute phase concentrations, at an isotherm of 35°C (similar to a saturation concentration at 35°C)
and different molecular weights to test to observe how these two variables affect the bioavailability of GLP-1.

A potentially confounding effect of this study is the method of delivery. With polypeptides that exhibit a UCST, to achieve $C_{sats}$ that are equivalent to the range of $C_{sats}$ previously tested with LCST polypeptides, the solution cloud point of the polypeptide will often be dramatically higher than the body temperature of the animal. Thus, one of our first tests will be establishing depots using either solubilizing small molecules (urea) that can diffuse from the injection site more rapidly than the polymer, thus enabling the polypeptide to be injected in a soluble solvent that rapidly exchanges with the environment to become a poor solvent, forming a depot (Figure 101). Secondly, we will try physically implanting a desiccated depot of a prescribed shape, size and dose that will rehydrate and begin to release active polypeptide after an initial delay (Figure 101). We are searching for consistent dosing that rapidly produces depots of similar size in the sub-cutaneous space.
Next, we will design polypeptides of similar molecular weight that exhibit different \( C_{\text{sat}} \). We will achieve this feat by utilizing the same parameter as before, the aromatic:aliphatic ratio to rationally tune the binodal of the GLP-1-polypeptide fusion. This allows us to observe the effect to \( C_{\text{sat}} \) on the therapeutic efficacy of the fusion. However, \( C_{\text{sat}} \) is just one parameter that affects bioavailability. The overall size of the molecule is also critical to the diffusion/convection in the subcutaneous space [274, 275]. Thus, using similar ranges of \( C_{\text{sat}} \) we will test depots for delivery as we increase the molecular weight of the molecule overall. Previous studies by previous lab members
have demonstrated that there diminishing returns of molecular weight beyond ~35 kDa and thus we will not exceed this size.

5.2 Methods

5.2.1 Gene synthesis

Each octapeptide amino acid motif with the desired saturation concentration were genetically fused to the C terminus of GLP-1. To increase the number of total repeats of the gene, we performed iterative cloning steps of Recursive Directional Ligation by Plasmid Reconstruction adding an addition twenty repeats during each step. Transformations were performed into the desired E. Coli cell line – BL21 (DE3) for recombinant expression.

5.2.2 Protein purification

Individual liquid cultures of BL21 E. Coli strains each harboring our gene of interest were inoculated into 5 ml of Terrific Broth (TB) medium from frozen glycerol stocks and grown to confluence overnight (16-18 hours). Cultures were then inoculated at a 1:200 dilution in 1L TB media supplemented with 45 μg ml⁻¹ kanamycin. Cells were grown at 37°C in a shaking incubator (~200 r.p.m.) for 9 hr, at which time protein expression was induced by the addition of 500 μM Isopropyl-β-D-thiogalactoside (IPTG). Cells were then incubated at 37°C (shaking at ~200 r.p.m.) for an additional 18 hr. Protein was then purified from the insoluble cell suspension fraction. In brief, cell pellets were isolated by centrifuging cultures at 3500 RCF and resuspending in 20 mL of milli-Q
water. Cells were then lysed by sonicating the cell solutions for 2 minutes, with 10 seconds of pulsing followed by 40 seconds of rest on ice (Misonix; Farmingdale, NY).

Centrifuging each lysate suspension at 20,000 RCF for 20 minutes results in a soluble and insoluble fraction. The supernatant was discarded with the insoluble fraction resuspended in an approximately equal volume of 8M urea + 140 mM PBS (~6-8 mL). This suspension was heated for 10 min in a 37°C water bath and then centrifuged at 20,000 RCF for 20 minutes. The supernatant was collected from this suspension and dialyzed in a 10 kDa membrane (SnakeSkinTM, Thermo Fischer Scientific) against a 1:200 milli-Q water solution at 4°C. The dialysis water was changed twice over a 48-hour period. From inside the dialysis bag, both insoluble and soluble components were collected and centrifuged at 3500 RCF for 10 minutes and 4°C. The supernatant was removed and the remaining insoluble pellet containing the protein of interest was lyophilized for a minimum of three days to remove all water from the pellet.

Protein purity was characterized by 4–20% gradient tris-HCl (Biorad, Hercules, CA) sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and staining with either 0.5 M copper chloride or SimplyBlueTM SafeStain (Thermo Fischer Scientific). Protein yield was determined by weight after lyophilization.

5.2.3 Characterization of phase separation and in vivo activity

5.2.3.1 Temperature dependent UV-vis spectrophotometry

Turbidity profiles were obtained for each of the constructs by recording the optical density as a function of temperature (1°C min⁻¹ ramp) on a temperature-controlled
UV-vis spectrophotometer (Cary 300 Bio; Varian Instruments; Palo Alto, CA). The transition temperature ($T_t$) was defined as the inflection point of the turbidity profile. Samples were measured in PBS at 10 µM.

5.2.3.2 Animal husbandry

All experimental procedures were conducted under protocol A053-15-02 approved by the Duke Institutional Animal Care and Use Committee (IACUC). 6-week old, male C57Bl/6J mice were purchased from Jackson Labs (strain 000664) and group housed in a room with a controlled photoperiod (12 hr light/12 hr dark cycle) and allowed at least 1 week to acclimate to the facilities prior to that start of procedures. Animals had unlimited access to water and food and were observed daily for signs and symptoms of distress. The diet-induced obese (DIO) phenotype was achieved by maintaining the mice on a high-fat (60 kcal% fat) diet upon arrival to the facility.

5.2.3.3 Endotoxin removal

Constructs were endotoxin purified prior to injection by passing the solution through a sterile 0.22 µm Acrodisc filter comprised of a positively charged and hydrophilic Mustang® E membrane (Pall Corporation). Constructs were filtered in 2M urea + 140 mM PBS at 37°C and then dialyzed against milli-Q H₂O at 4°C, changing the water three separate times over the course of 72 hours. Aggregated material was removed from the dialysis bag and pelleted with centrifugation (4°C, 3500 rpm). Samples were frozen and lyophilized for a minimum of 48 hrs.
5.2.3.5 Method of establishing GLP-1 releasing sub-cutaneous depots

In one method, the polypeptide is resuspended at 175 µM in 2M urea + 140 mM PBS. A total volume of 200 µl is injected into the right hind flank after shaving and removing all hair with chemical dissolution at the site of injection. Mice were weighed to determine injection volume required for ~2100 nmole of GLP-1 per kg of animal weight. Injection volume did not exceed 200 µl. In a second method, a small incision is made on the right hind flank with surgical scissors after animals were anesthetized with isoflurane. Incision site was pre-sterilized according to Duke husbandry guidelines. Pre-weighed, dehydrated polypeptide pellets are then inserted under the skin. The pellets rapidly rehydrate and become adherent to the skin tissue and thus for sealing the incision site, only a small amount of surgical glue was used to secure the skin flap.

5.2.3.6 Blood glucose measurement and weight measurements

Mice were put into a clear restraining tube. Their tails were wiped with 50% ethanol in sterile water and then dried. A small incision was made adjacent to the tail vein using a small lancet. The first drop of blood was blotted away. Blood glucose was quantified by applying the second drop of blood to the test strip of an AlphaTRAK 2 blood glucose meter (Abbott Laboratories). Weight was measured on a scale zeroed with a container into which the mice were briefly placed.

5.2.3.7 Statistical analysis

Experimental numbers for both in vitro and in vivo studies were selected based on knowledge gleaned from previous experiments or other published data. Because of
the small sample size (n ≤ 6), normality of groups was not tested. Variance across groups was similar except in untreated versus treated in vivo groups, which is not unexpected given the lack of glucose control in the mouse models tested. Blood glucose and percent change in weight studies were analyzed using repeated measures ANOVA, followed by lower order ANOVAs and Dunnett’s Test for multiple comparisons. For comparing two groups, two-tailed Student’s t-tests were used. No blinding was performed. All analysis and data processing were performed using Igor and R software.

5.3 Results and Discussion

We envisioned two strategies with the potential to establish depots subcutaneously for the UCST polypeptides which have transition temperatures far above safe biological temperature ranges, precluding a soluble to insoluble transition upon cooling to body temperature. The two strategies are to 1) employ urea to lower the solution cloud point for injection and 2) injection of a concentrated, dehydrated GLP-1-IDP fusion that will rehydrate, releasing peptide fusion. We decided to directly visualize this effect via fluorescent tomography. We chose a model IDP, (Gly-Arg-Gly-Asp-Ser-Pro-Tyr-Gln)-40 which has a predicted transition temperature of >70°C at the injection concentration necessary (175 µM or 1.2 mg which roughly corresponds to equivalent doses of GLP-1-ELP fusions used in previous studies, 1000 nmole kg⁻¹). Using a near infrared fluorescent tag (CW800) attached via NHS-ester chemistry at free amines, we can visualize the localization of the polypeptide in the hind flank.
We know that inclusion of 1M urea in solution with the polypeptide will reduce its observed cloud by point by ~25°C. Thus, by resuspending the model polypeptide in 2M Urea + PBS at 175 µM we can inject in a soluble state under ambient conditions. We hypothesize that considering the two order of magnitude difference in molecular weight between urea and the model polypeptide, urea will rapidly diffuse out of the subcutaneous space, leaving the remaining polypeptide in a poor solvent and thus will transition in situ.

Upon injection we observe something akin to a “burst” release of the polypeptide (Figure 102). This is characterized by a large area of fluorescence with a center of mass along the axis of the injection path of the needle. Over the first 8 hours this center of mass reduces in intensity, slowly reaching an equilibrium shape over the first two days. This center of mass disappeared by day 14.

Figure 102: Fluorescence molecular tomography of (GRGDSPYQ)40 labeled with a near infrared fluorescent dye after injection in the presence of 2M urea + PBS. Injection concentration of 175 µM, corresponding to 1.2 mg of protein total.
We also injected a dehydrated coacervate that had the same total protein content as the urea experiments. Here we observe a completely different behavior. Although our initial intention was to use convective flow from the syringe to push the dehydrated depot from the needle point into the subcutaneous space, upon contact with water, the dehydrated depot becomes extremely adhesive to the hydrophobic needle (Figure 103). Thus, we implanted the material with forceps, placing the dehydrated pellet underneath the skin through a small incision. Upon implantation, we observe an extremely small area of fluorescent that slowly expands over the first 90 minutes. The center of mass of the implant does not move noticeably over the course of two weeks. However, the mass of the depot is decreasing over this time, slowly releasing material surrounding the depot primarily during the first three days of implantation. Comparing these two injection strategies, we decided to move forward with the dehydrated depot strategy due to the lack of burst release and the increased persistence of the depot.

![Figure 103: Fluorescence molecular tomography of (GRGDSPYQ)40 labeled with a near infrared fluorescent dye after injection in dehydrated state. Injection mass equal to 1.2 mg of protein total.](image_url)
Fusion of GLP-1 to the N terminus of polypeptides was generally well tolerated. We observed minimal loss in yield from recombinant expression with most constructs expressing between 25-50 mg L⁻¹. As mentioned previously, we wanted to design peptide-polypeptide fusions that have $C_{\text{sat}}$ in the general ranges of ~0.1, 10, >100 µM corresponding to slow release, optimal release and near soluble release from the depot. This roughly corresponds to the $C_{\text{sat}}$s predicted for [3Y:V]-20, [Y:V]-20 and [3V:Y]-20. [3V:Y]-20 was not expected to exhibit phase behavior under physiologic conditions and thus six His residues were fused to the C terminus of the polypeptide and purified from the soluble fraction with chromatography.

The phase behavior of these polypeptide fusions was measured as before with temperature dependent UV-vis spectrophotometry. In determining the UCST binodal line, we identify these two proteins indeed have the desired phase behavior with GLP-1-[3Y:V]-20 exhibiting a $C_{\text{sat}}$ of ~30 µM and GLP-1-[Y:V]-20 exhibiting a $C_{\text{sat}}$ of ~500 µM (Figure 104). These roughly correspond to the values predicted by the RIDP of choice alone. As predicted GLP-1-[3V:Y]-20-His6X did not exhibit any phase behavior under physiologic conditions.
Figure 104: Binodal phase boundary of GLP-1-RIDPs of lower molecular weight. Data collected in 140 mM PBS. Dotted lines are fit lines for $y = m \ln(x) + b$.

After endotoxin purification, 1.2 mg of GLP-1-[3Y:V]-20, GLP-1-[Y:V]-20 and GLP-1-[3V:Y]-20-His6X were weighed and implanted in the hind flank of C57Bl/6J mice that have been fed 60% fat diet. In addition to these 3 groups, there is an additional group that received a saline injection. Over the course of the study, we measured blood glucose via tail vein blood draws at 0, 1, 2, 4, 8, 24 hrs and then each day thereafter for a total of 8 days.

The blood glucose data can be visualized in Figure 105. Overall, our strategy of implanting dehydrated depots was successful at controlling blood glucose. It is also positive that we are observing an effect of aromatic:aliphatic ratio, even in non-optimal molecular weight polypeptides. First, it is notable that in the early time points blood
glucose drops at approximately same speed suggesting that even in the soluble control, there are limitations regarding the minimal time to observe an effect on blood glucose. Second, each experimental group exhibits elements of burst release, with the largest change observed for those constructs that form subcutaneous depots. This result suggests that upon solubilization there is a larger bolus dose that reaches the bloodstream, which is reduced upon reaching an equilibrium state between depot release and protein clearance. Third, our depot forming formulations (GLP-1-[3Y:V]-20 and GLP-1-[Y:V]-20) each control blood glucose at least one additional day compared to the soluble RIDP control.

Figure 105: Blood glucose of GLP-1-RIDP fusion proteins of ~20 kDa in size and variable Csat. Data collected from C57Bl/6J mice that have been fed 60% fat diet. Error bars are standard error of the mean (n = 5).
Figure 106: Body weight change of mice with sub-cutaneous GLP-1-RIDP depots of ~20 kDa in size and variable $C_{\text{sat}}$. Data collected from C57Bl/6J mice that have been fed 60% fat diet. Error bars are standard deviation of the mean ($n = 5$).

Measurements of the body weight of the mice provide supplementary information on the efficacy of our sub-cutaneous depots (Figure 106). Again, we observe that our depot-forming proteins exhibit the greatest level of a burst release effect, resulting in the largest change in body weight in the first 2 days. This effect appears to be somewhat depot-dose dependent where the lower $C_{\text{sat}}$ construct exhibits the largest depression in appetite. As expected, the saline injection does not affect body weight.

Body weight measurements also differentiate our two depot forming fusions from one another. The high $C_{\text{sat}}$ construct body weight measurements suggest that their efficacy has waned by day 5 whereas the low $C_{\text{sat}}$ appears to be exerting a phenotypic effect until the end of the study (day 8).
These experiments mirror similar results of first pass experiments of optimization experiments performed by Kelli Luginbuhl with GLP-1-ELP depots. Kelli had identified that there were diminishing returns of polypeptides with molecular weights exceeding 35 kDa but improvements to glucose control between 20-35 kDa. Thus, we explored creating higher molecular weight variants of GLP-1-RIDP fusions.

![Figure 107: Binodal phase boundary of GLP-1-RIDPs of higher molecular weight. Data collected in 140 mM PBS. Dotted lines are fit lines for $y = m \ln(x) + b$.]

Increasing the molecular weight of polypeptide fusions produced a series that have $C_{sats}$ of 0.5, 7 & 60 µM by progressively reducing the aromatic content with aliphatic substitutions. Another GLP-1 protein fusion was also made with 75% aliphatic content that did not exhibit UCST phase behavior under physiologic conditions. I also
synthesized a molecular weight control (GLP-1-[S]-20) that has similar $C_{\text{sat}}$ to GLP-1-[3Y:V]-40 (7 µM compared to 12 µM) but is half the molecular weight.

![Blood glucose of GLP-1-RIDP fusion proteins of ~35 kDa in size and variable $C_{\text{sat}}$. Data collected from C57Bl/6J mice that have been fed 60% fat diet. Error bars are standard error of the mean ($n = 5$).](image)

**Figure 108:** Blood glucose of GLP-1-RIDP fusion proteins of ~35 kDa in size and variable $C_{\text{sat}}$. Data collected from C57Bl/6J mice that have been fed 60% fat diet. Error bars are standard error of the mean ($n = 5$).

The blood glucose measurements of mice with 2.0 mg depots implanted in their subcutaneous space can be visualized in Figure 108. These proteins that exhibit variable $C_{\text{sat}}$ also exhibit variable release from the depot in the subcutaneous space. The most hydrophobic depot, GLP-1-S-[40], appears to release the least amount of material suggesting that the depot biophysical properties is retarding a phenotypic effect of the peptide drug. The middle hydrophobic depots, GLP-1-[3Y:V]-40 and GLP-1-[Y:V]-
40, with $C_{\text{sat}}$ between 7 and 60 µM, exert similar levels of glucose control that is nearly a full 24 hr improvement over the low molecular weight versions. The most hydrophilic “depot”, predicted to be soluble can only manage glucose control over the first 24 hr. This is still an improvement from the soluble control at a lower molecular weight. These experiments support previous conclusions of optimal depot design, identifying that optimal release kinetics can be achieved with polypeptides that exhibit a $C_{\text{sat}}$ between 7 and 60 µM.

The effect of molecular weight, independent of $C_{\text{sat}}$, on blood glucose can be visualized in Figure 109. Increasing the molecular weight but retaining a $C_{\text{sat}}$ that is within an optimal release range can prolong glucose control by an additional two days. We hypothesize that this effect is a result of delayed diffusion into the blood stream and prolonged drug half-life resulting from delayed renal clearance.
Tracking the body weight of the mice supports the conclusions inferred from the blood glucose measurements (Figure 110). Here the parabolic effect of depot hydrophobicity is clear – at the extremes of hydrophobicity and hydrophilicity there is a lesser burst release and shorter duration of weight control. Our “optimal” constructs, GLP-1-[3Y:V]-40 and GLP-1-[Y:V]-40, are still exhibiting weight control at 7 days suggesting that there must be small amounts of material releasing from the depot even at 144 hours after implantation. Our molecular weight control, GLP-1-[S]-20, exhibits lesser burst released and shorter duration of efficacy than its larger molecular weight analogue, GLP-1-[3Y:V]-40, again supporting the conclusion of delayed entry and prolonged persistence in the blood stream from higher molecular weight depots.
Figure 110: Body weight change of mice with sub-cutaneous GLP-1-RIDP depots of ~35 kDa in size and variable $C_{sat}$. Data collected from C57Bl/6J mice that have been fed 60% fat diet. Error bars are standard error of the mean ($n = 5$).

5.4 Conclusions

In summary, we leveraged our knowledge of encoding UCST binodal lines to design mimetic fusions compared to previous work with the GLP-1-ELP system. Along this journey we also discovered multiple new routes of establishing sub-cutaneous depots with unfavorable working transition temperatures. Using these two innovations we were able to create depots that performed with similar efficacy to previously optimized GLP-1-ELP depots, controlling blood glucose for up to 5 days in vivo in a DIO mice model.
6. Conclusions and Future Directions

Often during my graduate work, I felt that I was often dealing with problems that those who work with LCST polymers had moved beyond. I feel that this thesis has closed this gap, bringing LCST and UCST polypeptides to an equal playing field in terms of controlling the binodal phase boundary. However, these UCST polypeptides have opened new avenues of inquiry on several scientific fronts.

6.1 Future directions of artificial IDP design

First and foremost, for major advancements to be discovered linking primary sequence features to function in IDPs, researchers need to address the elephant in the room – throughput. Over 6 years of this thesis I have been able to design, express, characterize ~100 unimeric protein sequences that exhibit exemplary well behaved UCST phase behavior. If progress is ever to be made to understand the biophysical importance of IDPs that occur naturally, there will first need to be an innovation in the number of sequences that can be characterized simultaneously, otherwise progress will continue to occur over decades instead of months or years.

I do believe that this thesis highlights specific sequence compositions of interest. As we now know, small deviations from repetition do not appear to influence phase behavior dramatically in the context of other more important factors (chain length, aromatic:aliphatic ratio, charge balance etc.). Thus, the increase in throughput experienced by the DNA synthesis community is now available to researchers studying naturally occurring IDPs or semi-repetitive IDPs. There are still significant improvements
that need to be explored to increase phase diagram collection speed, but this may be possible to accomplish in situ.

A project that has been met with little success in our lab is the combination of disorder with intrinsic bioactivity. This is due to several reasons but the emergence of short linear motifs (SLiMs) in the IDP community has the potential to be the key that provides the breakthrough. These motifs are disordered but can still bind their target in a disordered state or undergo a disorder-order transition that enables binding. Using our information about encoding UCST phase behavior and the knowledge of these SLiM motifs, we may be able to dope in SLiM motifs onto Pro/Gly scaffold that complements the chemical propensity of the SLiM to phase separate. I hypothesize that these linear genes may radically improve the pharmacokinetics of the base SLiM motif (a similar magnitude increase to that seen with ELP-peptide fusion proteins[101, 276]) and dramatically increase their potency (via multivalency observed with various linear tandem repeat protein[277]). More generally, incorporation of non-repetitive sequences with specific target epitopes for specific macro-molecular interaction is now feasible with this large library that explains how to specifically encode UCST phase behavior using every canonical amino acid.

Protein polypeptides have long been the envy of synthetic polymer chemists due to the charge patterning but there has yet to be a real effort to understand this behavior with polypeptides. Our data would suggest that this library of RIDPs are near-ideal disordered proteins that have a molecular architecture similar to polyampholytes[278]. There are several outstanding questions regarding charge-balance and charge-
patterning in both communities that could be systematically explored with this
generalized scaffold. The ability to incorporate non-canonical amino acids also enables
site-specific incorporation of side chains that may influence phase behavior in extremely
controlled, chemically precise experiments.

These RIDPs will always be compared to ELPs if only because of their proximity
to the lab from which they were published. Although I do not believe there to be an
obvious best choice of UCST or LCST for any one application I do believe that the
chemical diversity of side chains of RIDPs lend themselves much better to post-
translational modification, either by canonical enzymes or chemical reaction. There is a
host of Tyr chemistries that may attach small molecules at extraordinary densities to
RIDPs (drug/polypeptide mole) [279, 280].

Finally, a few anecdotal observations from my experiences that may nourish
future applications of RIDPs. Historically our lab has struggled to produce proteins by
any other method other than recombinant expression in E. Coli. However, in the ladder
days of my thesis I was able to successfully express, export and purify proteins from
Expi293F cells suggesting that concerns over the tolerability of IDPs with eukaryotes is
perhaps overstated. Additionally, since we now understand composition to be the
primary concern for the UCST binodal, I believe that solid-phase peptide synthesis is a
viable alternative to produce protein sequences (without endotoxin) that is also much
more amenable to un-natural side chain chemistry than what is currently possible in
bacterial/eukaryotic expression systems.
6.2 Future directions of artificial intracellular puncta

We envision that the strict adherence to phase separation behavior predicted by polymer physics in cells could be exploited to replace existing IDPRs or simplify the IDP under study to “phase separating components” and other unknown functional components. Many previous attempts to understand the importance of phase behavior will often remove whole protein domains, a technique often employed in structural biology. However, this is less appropriate here, where subtly is critical to understanding function and interactions that are not as specific. Using RIDPs, one can start at the beginning – with a well-behaved phase separating component with exquisitely tunable $C_{sat}$ and $\phi''$.

The proof of concept experiments in Chapter 3 demonstrate that intracellular droplets can be engineered to have non-canonical functions. We have performed this with an enzyme where the partitioning of a small molecule into the phase separated droplet containing the enzyme is essential. However, using two RIDPs, one could thermodynamically co-localize two molecules, enzyme and substrate, into closer than statistical probability with one another to improve kinetics.

A concept not explored by this work was developing RIDPs that could be responsive to an environmental change. We know that engineering thermoresponsivness would be trivial but response to pH, light or a post-translational modification are all small intellectual leaps from the presented data. We were able to design molecules with $in situ$ conjugation capability suggesting that click reactions could be performed in the cell to specifically encode post-translational modifications or make
large chemical changes to the RIDP that allow for extrinsically triggerable phase separation.

I believe there are significant opportunities in the gene circuit or synthetic biology field. The ability to program “buffering” type behavior through phase separation into a gene of interest allows for more tolerant circuit designs to be implemented. Here again, the ability to possibility control miscibility with another molecule via molecular tag enables more intentional intracellular engineering.

6.3 Future directions of protein polypeptide self-assembly

In Chapter 4, we extensively characterized the assembly of block co-polypeptides comprised of UCST-LCST block architectures and UCST-UCST block architectures. We demonstrated predictable control over nanoscale assembly in terms of both size and shape of the micelles. Finally, we demonstrate one-pot synthesis of targeted nanoparticles with dramatically increased affinity. These innovations suggest that the next step for these materials is in vivo. Now that we are in possession of the universal platform for multivalent display of biologics, I believe we need to choose a few select therapeutic targets, either those with poor affinity ligands that have been developed, or doubling down on a great target to increase potency to femtomolar regimes, which is predicted to be possible if starting with a low nanomolar or high picomolar affinity targeting molecule.

Another obvious direction for this work that is currently underway is the crosslinking of the core to improve stability of the particle. This stands to further
increase potency of a targeting molecule and eliminates the high nanomolar CMC of the phase separation stabilized particles. Indeed, we are now in possession of the perfect system to cleanly test the importance of CMC on in vivo pharmacodynamics/pharmacokinetics.

My experience with physical loading of paclitaxel into the core of these particles suggests that there is significant bandwidth to engineer the core of these micelles. Now that we understand the clear rules for encoding self-assembly, we can begin to make dramatic changes to the chemical identity of the core. I believe that it may even be possible to move beyond non-specific associations that are a function of the side chain and begin to include specific epitopes that are designed to specifically sequester hydrophobic molecules. There may be an interesting synergy between phage display and primary UCST polypeptide design to create the ultimate designer core for particularly insoluble small molecule drugs. As mentioned, previously, there is also significant potential to chemical conjugate drugs at a high density at Tyr residues located on each repeating polypeptide unit. Anecdotally, I was able to label 50% of all Tyr with I-125, corresponding to ~40 conjugations per block co-polypeptide chain.

6.4 Future directions for the delivery of peptide therapeutics

Our innovations creating sub-cutaneous depots point to multiple future lines of inquiry. First, there is an obvious temptation to increase or decrease the rate of depot release using localized heating and cooling at the depot injection site. In this case, a UCST system make much sense as heating both increases the soluble concentration in
the local area that synergizes with the natural biological response of increased diffusion rates, convection rates in the subcutaneous space and blood flow in the tissue.

Our results offer another solution to a common problem. If one is using a stable peptide drug that has some tolerance for lyophilization, our depots can improve that tolerance and serve as the delivery mechanism that allows one to eliminate the so called “cold-chain”. This refers to the necessity to preserve the activity of a biologic with freezer packs and ice as it is transferred from its manufacturing location to the consumer, a major burden for the pharmaceutical industry. The depots also provide upper limit dose control, eliminating the need for a trained profession to prepare the dose for every individual patient.

I also believe that there is significant bandwidth to prolong the efficacy of depots through highly potent molecules – those with measured affinities in the picomolar and femtomolar regime. Our imaging results and evaluation of body weight suggest that these depots are still releasing material at time points >2 weeks but that the concentration is not providing a therapeutic effect. GLP-1 has a low nanomolar EC50, suggesting that an increasingly potent drug would continue to benefit from these depots beyond the results observed for GLP-1. As a part of my thesis, we have demonstrated how multivalency can dramatically increase the potency of macromolecules and thus there appears to be an obvious directive to create a GLP-1 decorated micelle with optimal coronal C_{sat} (5-100 µM).
## Appendix A

### Table S1 - Full amino acid description of proteins with a single repeating motif

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Full Amino Acid Sequence</th>
<th># Amino Acid</th>
<th>Molecular Weight (Da)</th>
<th>µM @ 37°C</th>
<th>Gel Image #</th>
</tr>
</thead>
<tbody>
<tr>
<td>[S]-20</td>
<td>SKGP-[GRGDSPYS]_{20}GY</td>
<td>166</td>
<td>17004</td>
<td>44.19</td>
<td>1</td>
</tr>
<tr>
<td>[S]-40</td>
<td>SKGP-[GRGDSPYS]_{40}GY</td>
<td>326</td>
<td>33400</td>
<td>0.7550</td>
<td>2</td>
</tr>
<tr>
<td>[S]-60</td>
<td>SKGP-[GRGDSPYS]_{60}GY</td>
<td>486</td>
<td>49797</td>
<td>0.0446</td>
<td>3</td>
</tr>
<tr>
<td>[S]-80</td>
<td>SKGP-[GRGDSPYS]_{80}GY</td>
<td>646</td>
<td>66193</td>
<td>0.0002</td>
<td>4</td>
</tr>
<tr>
<td>[Q]-20</td>
<td>SKGP-[GRGDQPYQ]_{20}GY</td>
<td>166</td>
<td>18646</td>
<td>247.9</td>
<td>8</td>
</tr>
<tr>
<td>[Q]-40</td>
<td>SKGP-[GRGDQPYQ]_{40}GY</td>
<td>326</td>
<td>36685</td>
<td>6.593</td>
<td>9</td>
</tr>
<tr>
<td>[Q]-60</td>
<td>SKGP-[GRGDQPYQ]_{60}GY</td>
<td>486</td>
<td>54723</td>
<td>1.034</td>
<td>10</td>
</tr>
<tr>
<td>[Q]-80</td>
<td>SKGP-[GRGDQPYQ]_{80}GY</td>
<td>646</td>
<td>72762</td>
<td>0.241</td>
<td>11</td>
</tr>
<tr>
<td>[T]-40</td>
<td>SKGP-[GRGDTPYT]_{40}GY</td>
<td>326</td>
<td>34522</td>
<td>#DIV/0!</td>
<td>60</td>
</tr>
<tr>
<td>[N]-20</td>
<td>SKGP-[GRGDNPYN]_{20}GY</td>
<td>166</td>
<td>18085</td>
<td>337.2</td>
<td>12</td>
</tr>
<tr>
<td>[N]-40</td>
<td>SKGP-[GRGDNPYN]_{40}GY</td>
<td>326</td>
<td>35562</td>
<td>39.18</td>
<td>13</td>
</tr>
<tr>
<td>[N]-60</td>
<td>SKGP-[GRGDNPYN]_{60}GY</td>
<td>486</td>
<td>53040</td>
<td>4.827</td>
<td>14</td>
</tr>
<tr>
<td>[H]-40</td>
<td>SKGP-[GRGDSPHS]_{40}GY</td>
<td>326</td>
<td>32359</td>
<td>8376</td>
<td>61</td>
</tr>
<tr>
<td>[H]-60</td>
<td>SKGP-[GRGDSPHS]_{60}GY</td>
<td>486</td>
<td>48235</td>
<td>286.8</td>
<td>24</td>
</tr>
<tr>
<td>[H]-80</td>
<td>SKGP-[GRGDSPHS]_{80}GY</td>
<td>646</td>
<td>64111</td>
<td>327.1</td>
<td>25</td>
</tr>
<tr>
<td>[F]-40</td>
<td>SKGP-[GRGDSPFS]_{40}GY</td>
<td>326</td>
<td>32760</td>
<td>1.923</td>
<td>62</td>
</tr>
<tr>
<td>[SQ]-20</td>
<td>SKGP-[GRGDSPYQ]_{20}GY</td>
<td>166</td>
<td>17825</td>
<td>0.0291</td>
<td>38</td>
</tr>
<tr>
<td>[SQ]-40</td>
<td>SKGP-[GRGDSPYQ]_{40}GY</td>
<td>326</td>
<td>35042</td>
<td>0.1359</td>
<td>39</td>
</tr>
<tr>
<td>[SQ]-60</td>
<td>SKGP-[GRGDSPYQ]_{60}GY</td>
<td>486</td>
<td>52260</td>
<td>3.179</td>
<td>40</td>
</tr>
<tr>
<td>[SQ]-80</td>
<td>SKGP-[GRGDSPYQ]_{80}GY</td>
<td>646</td>
<td>69478</td>
<td>15.72</td>
<td>41</td>
</tr>
<tr>
<td>[QS]-20</td>
<td>SKGP-[GRGDQPS]_{20}GY</td>
<td>166</td>
<td>17825</td>
<td>149.0</td>
<td>29</td>
</tr>
<tr>
<td>[QS]-40</td>
<td>SKGP-[GRGDQPS]_{40}GY</td>
<td>326</td>
<td>35042</td>
<td>7.941</td>
<td>30</td>
</tr>
<tr>
<td>[QS]-60</td>
<td>SKGP-[GRGDQPS]_{60}GY</td>
<td>486</td>
<td>52260</td>
<td>0.7239</td>
<td>31</td>
</tr>
<tr>
<td>[QS]-80</td>
<td>SKGP-[GRGDQPS]_{80}GY</td>
<td>646</td>
<td>69478</td>
<td>0.0861</td>
<td>32</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Full Amino Acid Sequence</td>
<td># Amino Acid</td>
<td>Molecular Weight (Da)</td>
<td>µM @ 37C</td>
<td>Gel Image #</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------</td>
<td>-------------</td>
<td>-----------------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>[3S:Q]-40</td>
<td>SKGP-[GRGDSPYSGRGDSPYSG RGDSPYSGRGDQPYQ]_{10-}GY</td>
<td>326</td>
<td>34221</td>
<td>1.6006</td>
<td>33</td>
</tr>
<tr>
<td>[S:Q]-40</td>
<td>SKGP-[GRGDSPYSGRGDQPYQ]_{20-}GY</td>
<td>326</td>
<td>35042</td>
<td>1.9073</td>
<td>34</td>
</tr>
<tr>
<td>[S:3Q]-40</td>
<td>SKGP-[GRGDQPYQGRGDQPYQGRGDQPYQGRGDSPYS]_{10-}GY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>326</td>
<td>35863</td>
<td>5.0625</td>
<td>35</td>
</tr>
<tr>
<td>[3Y:V]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPYSGRGDSPVS]_{10}^-GY</td>
<td>326</td>
<td>32760</td>
<td>24.540</td>
<td>36</td>
</tr>
<tr>
<td>[Y:V]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPVS]_{20}^-GY</td>
<td>326</td>
<td>32119</td>
<td>815.49</td>
<td>37</td>
</tr>
<tr>
<td>[Y:3V]-40</td>
<td>SKGP- [GRGDSPVSGRGDSPVSGRGDSPVS]_{10}^-GY</td>
<td>326</td>
<td>31479</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>[V]-40</td>
<td>SKGP- [GRGDSPVS]_{40}^-GY</td>
<td>326</td>
<td>30839</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>[3Y:A]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPYSGRGDSPYSGRGDSPYSG]_{10}^-GY</td>
<td>326</td>
<td>32479</td>
<td>28.465</td>
<td>45</td>
</tr>
<tr>
<td>[Y:A]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPAS]_{20}^-GY</td>
<td>326</td>
<td>31558</td>
<td>1816.7</td>
<td>46</td>
</tr>
<tr>
<td>[3Y:I]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPYSGRGDSPYSG]_{10}^-GY</td>
<td>326</td>
<td>32900</td>
<td>12.883</td>
<td>47</td>
</tr>
<tr>
<td>[Y:I]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPIS]_{20}^-GY</td>
<td>326</td>
<td>32400</td>
<td>214.85</td>
<td>48</td>
</tr>
<tr>
<td>[3Y:M]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPYSGRGDSPYSGRGDSPYSG]_{10}^-GY</td>
<td>326</td>
<td>33080</td>
<td>7.9140</td>
<td>49</td>
</tr>
<tr>
<td>[Y:M]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPMS]_{20}^-GY</td>
<td>326</td>
<td>32761</td>
<td>110.52</td>
<td>50</td>
</tr>
<tr>
<td>[3Y:H]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPYSGRGDSPYSGRGDSPYSG]_{10}^-GY</td>
<td>326</td>
<td>33140</td>
<td>7.1257</td>
<td>42</td>
</tr>
<tr>
<td>[Y:H]-40</td>
<td>SKGP- [GRGDSPYSGRGDSPHS]_{20}^-GY</td>
<td>326</td>
<td>32880</td>
<td>70.864</td>
<td>43</td>
</tr>
<tr>
<td>[Y:3H]-40</td>
<td>SKGP- [GRGDSPHSGRGDSPHSG]</td>
<td>326</td>
<td>32619</td>
<td>508.53</td>
<td>44</td>
</tr>
<tr>
<td>Peptide</td>
<td>Sequence</td>
<td>Mw</td>
<td>pIC50</td>
<td>nM</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>----</td>
<td>------</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>[R:K]-40</td>
<td>SKGP-[GRGDSPYSGKGDSPYS]20-</td>
<td>326</td>
<td>32840</td>
<td>79.887</td>
<td>52</td>
</tr>
<tr>
<td>[D:E]-40</td>
<td>SKGP-[GRGDSPYSGRGESPYS]20-</td>
<td>326</td>
<td>33681</td>
<td>1.7797</td>
<td>53</td>
</tr>
<tr>
<td>[5:7P]-40</td>
<td>SKGP[GRPDSPYSGRGDSPYSGRGDSPYSGRPDSPYSGRGDSPYSGRGDSPYSGRGDSPYSGRGDSPYSGRGDSPYSGRGDSPYS]2-GY</td>
<td>326</td>
<td>33801</td>
<td>0.7428</td>
<td>54</td>
</tr>
<tr>
<td>[G:7P]-40</td>
<td>SKGP[GRPDSPYSGRGDSPYSGRGDSPYSGRPDSPYSGRGDSPYSGRPDSPYSGRGDSPYSGRGDSPYSGRGDSPYSGRGDSPYS]10-GY</td>
<td>326</td>
<td>34202</td>
<td>0.7283</td>
<td>55</td>
</tr>
<tr>
<td>[7:5P]-40</td>
<td>SKGP-[GRPDSPYSPRGDSPYSGRPDSPYSPRGDSPYSGRPDSPYSPRGDSPYSGRPDSPYSPRGDSPYSGRPDSPYSPRGDSPYSGRPDSPYSPRGDSPYS]2-GY</td>
<td>326</td>
<td>34602</td>
<td>0.9106</td>
<td>56</td>
</tr>
<tr>
<td>[2V:18Y] [2R:8K]-4</td>
<td>SKGP-[GRGDSPYSGRGDSPVSGRGDSPYSGRGDSPVSGRGDSPYSGRGDSPVSGRGDSPYSGRGDSPVSGRGDSPYSGRGDSPVSGRGDSPYSGRGDSPYSG]2-GY</td>
<td>326</td>
<td>32264</td>
<td>76.61</td>
<td>57</td>
</tr>
<tr>
<td>5 V:15Y</td>
<td></td>
<td>1 R:5K</td>
<td>-4</td>
<td>SKGP-[GRGDSPYSGRGDSPVSG RGDPYSKGDSPYS]_{10}^{-}GY</td>
<td>326</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>[8V:12Y]</td>
<td>8R:2K]-40</td>
<td>SKGP-[GRGDSPYSKGDSPYSGRG DSPYSKGDSPYSGRGDS PYSGRKGDSPVVS]_{10}^{-}GR GDSPYSKGDSPYSGR GDSPYSKGDSPYSGY</td>
<td>326</td>
<td>32696</td>
<td>658.69</td>
</tr>
<tr>
<td>[3Y:V]-40-sfGFP</td>
<td>SKGP-[GRGDSPYSGRGDSPYSRG DSPYSGRGDPSVS]_{10}^{-}sfGFP</td>
<td>570</td>
<td>60307</td>
<td>42.468</td>
<td>28</td>
</tr>
</tbody>
</table>

Figure S1: SDS-PAGE gels of purified proteins used in this study. Lane labels correspond to Table S1 and Table S2.
### Appendix B

**Table S3: Surface Plasmon Resonance fit parameters**

<table>
<thead>
<tr>
<th></th>
<th>RLP$<em>{20}$-ELP$</em>{80}$-Fn3</th>
<th>RLP$<em>{40}$-ELP$</em>{80}$-Fn3</th>
<th>RLP$<em>{80}$-ELP$</em>{80}$-Fn3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Langmuir 1:1</td>
<td>Heterogenous ligand</td>
<td></td>
</tr>
<tr>
<td>$K_{on1}$ (L/(mol•s))</td>
<td>4.79×10$^3$</td>
<td>4.25×10$^3$</td>
<td>4.07×10$^3$</td>
</tr>
<tr>
<td>$K_{off1}$ (s$^{-1}$)</td>
<td>4.57×10$^{-3}$</td>
<td>3.35×10$^{-4}$</td>
<td>3.20×10$^{-6}$</td>
</tr>
<tr>
<td>$K_{on2}$ (L/(mol•s))</td>
<td>-</td>
<td>3.96×10$^2$</td>
<td>4.81×10$^2$</td>
</tr>
<tr>
<td>$K_{off2}$ (s$^{-1}$)</td>
<td>-</td>
<td>2.32×10$^{-2}$</td>
<td>3.86×10$^{-3}$</td>
</tr>
<tr>
<td>$K_{D1}$ (mol/L)</td>
<td>9.54×10$^{-7}$</td>
<td>7.89×10$^{-8}$</td>
<td>7.86×10$^{-10}$</td>
</tr>
<tr>
<td>$K_{D2}$ (mol/L)</td>
<td>-</td>
<td>5.86×10$^{-5}$</td>
<td>8.04×10$^{-6}$</td>
</tr>
<tr>
<td>Rmax1</td>
<td>29.43</td>
<td>191.74</td>
<td>288.90</td>
</tr>
<tr>
<td>Rmax2</td>
<td>-</td>
<td>20.81</td>
<td>25.62</td>
</tr>
<tr>
<td>Rmax1 %</td>
<td>-</td>
<td>90.21</td>
<td>91.85</td>
</tr>
<tr>
<td>Rmax2 %</td>
<td>-</td>
<td>9.79</td>
<td>8.15</td>
</tr>
<tr>
<td>Chi square</td>
<td>2.05</td>
<td>2.23</td>
<td>2.43</td>
</tr>
<tr>
<td>Shape</td>
<td>Unimeric</td>
<td>Spherical</td>
<td>Sphere &amp; Worm-like</td>
</tr>
</tbody>
</table>
### Table S4: 3-way ANOVA of cell area

<table>
<thead>
<tr>
<th></th>
<th>DoF</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fn3</td>
<td>1</td>
<td>49.13</td>
<td>49.13</td>
<td>61.008</td>
<td>8.43E-11***</td>
</tr>
<tr>
<td>Time</td>
<td>4</td>
<td>15.79</td>
<td>3.95</td>
<td>4.9</td>
<td>0.0017**</td>
</tr>
<tr>
<td>Name</td>
<td>2</td>
<td>101.12</td>
<td>50.56</td>
<td>62.779</td>
<td>1.25E-15***</td>
</tr>
<tr>
<td>Fn3:Time</td>
<td>4</td>
<td>3.39</td>
<td>0.85</td>
<td>1.054</td>
<td>0.3871</td>
</tr>
<tr>
<td>Fn3:Name</td>
<td>1</td>
<td>29.67</td>
<td>29.67</td>
<td>36.84</td>
<td>8.49E-08***</td>
</tr>
<tr>
<td>Time:Name</td>
<td>8</td>
<td>26.64</td>
<td>3.33</td>
<td>4.135</td>
<td>0.0005***</td>
</tr>
<tr>
<td>Fn3:Time:Name</td>
<td>4</td>
<td>6.94</td>
<td>1.73</td>
<td>2.153</td>
<td>0.0848.</td>
</tr>
<tr>
<td>Residuals</td>
<td>62</td>
<td>49.93</td>
<td>0.81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table S5: 3-way ANOVA of # particles/cell

<table>
<thead>
<tr>
<th></th>
<th>DoF</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fn3</td>
<td>1</td>
<td>134.09</td>
<td>134.09</td>
<td>119.75</td>
<td>4.09E-16***</td>
</tr>
<tr>
<td>Time</td>
<td>4</td>
<td>29.58</td>
<td>7.39</td>
<td>6.604</td>
<td>0.0002***</td>
</tr>
<tr>
<td>Name</td>
<td>2</td>
<td>235.34</td>
<td>117.67</td>
<td>105.09</td>
<td>&lt; 2e-16***</td>
</tr>
<tr>
<td>Fn3:Time</td>
<td>4</td>
<td>14.79</td>
<td>3.7</td>
<td>3.303</td>
<td>0.0162*</td>
</tr>
<tr>
<td>Fn3:Name</td>
<td>1</td>
<td>29.39</td>
<td>29.39</td>
<td>26.25</td>
<td>3.16E-06***</td>
</tr>
<tr>
<td>Time:Name</td>
<td>8</td>
<td>23.3</td>
<td>2.91</td>
<td>2.601</td>
<td>0.0161*</td>
</tr>
<tr>
<td>Fn3:Time:Name</td>
<td>4</td>
<td>1.95</td>
<td>0.49</td>
<td>0.436</td>
<td>0.782</td>
</tr>
<tr>
<td>Residuals</td>
<td>62</td>
<td>69.42</td>
<td>1.12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table S6: 3-way ANOVA of % of cells with fluorescent signal

<table>
<thead>
<tr>
<th></th>
<th>DoF</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fn3</td>
<td>1</td>
<td>2.3455</td>
<td>2.3455</td>
<td>348.75</td>
<td>&lt; 2e-16***</td>
</tr>
<tr>
<td>Time</td>
<td>4</td>
<td>0.0403</td>
<td>0.0101</td>
<td>1.499</td>
<td>0.2135</td>
</tr>
<tr>
<td>Name</td>
<td>2</td>
<td>0.132</td>
<td>0.066</td>
<td>9.813</td>
<td>0.0002 ***</td>
</tr>
<tr>
<td>Fn3:Time</td>
<td>4</td>
<td>0.1244</td>
<td>0.0311</td>
<td>4.624</td>
<td>0.0025 **</td>
</tr>
<tr>
<td>Fn3:Name</td>
<td>1</td>
<td>0.0192</td>
<td>0.0192</td>
<td>2.852</td>
<td>0.0963 .</td>
</tr>
<tr>
<td>Time:Name</td>
<td>8</td>
<td>0.218</td>
<td>0.0273</td>
<td>4.052</td>
<td>0.0006 ***</td>
</tr>
<tr>
<td>Fn3:Time:Name</td>
<td>4</td>
<td>0.056</td>
<td>0.014</td>
<td>2.08</td>
<td>0.0941 .</td>
</tr>
<tr>
<td>Residuals</td>
<td>62</td>
<td>0.417</td>
<td>0.0067</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


Biography

Michael graduated valedictorian from Cathedral Preparatory High School in 2008 and attended college Case Western Reserve University in Cleveland, Ohio. In 2012, he graduated magna cum laude from with a Bachelor of Science and Engineering, in Biomedical Engineering. There he was the recipient of the Jose Alcala Memorial Award for excellence in biomedical engineering research and the Robert J. Adler Award for high scholarship, technical creativity and service to his peers. After one year in industry, Michael joined the lab of Ashutosh Chilkoti. While at Duke, Michael was a Triangle Materials Research Science and Engineering Center (MRSEC) fellow and a James McElhaney fellow. His time spent in the Chilkoti laboratory has resulted in the following scientific publications (* first author).


