

Compositional Control of Phase Separated Cellular Bodies

Approved by Supervisory Committee

Michael K. Rosen, Ph.D.

Hongtao Yu, Ph.D.

Philip J. Thomas, Ph.D.

Zhijian “James” Chen, Ph.D.

Dedicated to Alyssa, Ferozali, Yasmin, and Farhan.

Compositional Control of Phase Separated Cellular Bodies

by

Salman F. Banani, B.S.

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS SOUTHWESTERN MEDICAL CENTER

Dallas, Texas

May 2018

Copyright

by

Salman F. Banani, 2018

All rights reserved

Acknowledgments

I first thank my advisor, Dr. Michael Rosen, Ph.D., for giving me the opportunity to pursue my dissertation in his laboratory and under his guidance, and for exposing me to a new and exciting field in cell biology and biophysics. I thank him for teaching me, directly or by example, what it takes to be a scientist—in particular, for teaching me the importance of clear, rigorous, and quantitative thinking; of being grounded in theory; of using simple model systems to elucidate general principles in biology; of not being intimidated by tough problems, but rather systematically deconstructing them to find the solution; of being a meticulous analyst; of being a clear, concise writer; of thinking before I speak; of treating people with fairness and respect; and most of all, of finding joy and excitement in good science. I continually strive to improve in these traits.

I thank all members of the Rosen lab for being wonderful colleagues, professionally and socially, and for creating a stimulating and supportive work environment. The Rosen lab is a rigorous scientific environment, and a place where I have been able to learn and grow as a scientist. In particular, I thank Dr. Shae Padrick, Ph.D. for teaching me many facets of experimental biochemistry and biophysics, for sharing with me a part of his large fund of scientific knowledge, and for indulging my often naive questions and tangential discussions; Dr. Jon Ditlev, Ph.D., for important and insightful discussions about image analysis; and Dr. Sudeep Banjade, Ph.D., for teaching me protein purification and basic biochemistry during my rotation.

I thank Dr. Rama Ranganathan, M.D., Ph.D., for his unconditional mentorship and for his scientific guidance and encouragement. I thank him for sharing his wisdom and insights into thinking deeply about biology; asking good scientific questions; picking

fundamental scientific problems to work on; designing one's research program; and looking for simplicity in biology. I also thank him for allowing me to attend his lab's group meetings and events on numerous occasions. This enabled me to immerse myself in a complementary scientific environment and encouraged me to think more broadly, deeply, and creatively.

I thank all my colleagues at UTSW for helping me to considerably expand the scope of my training. Their unique experiences have helped me to think more broadly and creatively about my work at present and my interests going forward. In particular, I thank Arjun Raman for countless insightful discussions about biological systems, epistasis and complexity, as well as about the unique challenges that face MSTP students interested in basic research; Dr. Mark Kittisopikul, Ph.D., for sharing with me numerous insights about statistical analysis, image analysis, and quantitative thinking; Dr. Yuan Lin, Ph.D., for many insightful discussions about the logic, philosophy, and different styles of doing science; Ryan Golden for sharing numerous experimental insights about molecular and cell biology. These discussions have been invaluable for me during graduate school education.

I thank all the faculty of the Biophysics graduate program and department for always being approachable whenever I was in need of advice or their particular expertise. In particular, I thank my thesis committee—Drs. Hongtao Yu, Ph.D.(chair), Philip Thomas, Ph.D., and James Chen, Ph.D.—for their support and guidance throughout the years. I also thank Dr. Luke Rice, Ph.D., for providing thoughtful and extremely constructive criticism during many phases of my training—ranging from my qualifying exam to my thesis work—and for answering numerous questions about the process of science.

I thank all my collaborators and co-authors, who contributed time, energy, resources, and data to help me to complete my thesis work—Allyson Rice, Dr. Yuan Lin, Ph.D., Ben Peeples, Dr. Saumya Jain, Ph.D., and Dr. Roy Parker, Ph.D.. In particular I thank Drs. Saumya Jain and Roy Parker for their contribution of P body experiments in yeast; Drs. Andrea Rabellino, Ph.D., and Pier Paolo Scaglioni, M.D., (for providing the *PML*^{-/-}

cells, for help with generation of stable cell lines, and for assistance with apoptosis assays; Jennifer Eitson, Katrina Mar, and Dr. John Schoggins, Ph.D. for their assistance with viral assays.

I thank the numerous core facilities and the personnel therein that provided crucial assistance to complete this work—Louie Kerr (MBL Central Microscopy Facility) and Dr. Kate Luby-Phelps, Ph.D. and Abhijit Bugde (UTSW Live Cell Imaging Facility) for assistance with imaging; Drs. Chad Brautigam, Ph.D., and Thomas Scheuermann, Ph.D., (UTSW Molecular Biophysics Resource) for assistance with ITC; Dr. Chaoying Liang, Ph.D., (UTSW Genomics & Microarray Core) for assistance with RNASeq; Dr. Larry Dangott, Ph.D. (Texas A&M Protein Chemistry Lab) for assistance with 2D Gel Electrophoresis.

I thank all my undergraduate scientific mentors. In particular, I thank Drs. Warner Greene, M.D., Ph.D., and Renee Reijo Pera, Ph.D., for giving me the opportunity to work in their labs as an undergraduate. I also thank Drs. Mario Santiago, Ph.D., Amander Clark, Ph.D., and Cory Nicholas, Ph.D., for their experimental training and mentorship. Their support and encouragement enabled me to pursue graduate training.

Finally, I thank my wonderful family for their unconditional love, support, and patience—my wife, Alyssa Premji; my parents, Ferozali and Yasmin Banani and Amin and Nilofer Badrudin; and my siblings Farhan Banani and Natasha Premji.

Compositional Control of Phase Separated Cellular Bodies

Salman F. Banani, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2018

Supervisor: Michael K. Rosen, Ph.D.

Cellular bodies such as P bodies and PML nuclear bodies (PML NBs) appear to be phase separated liquids organized by multivalent interactions among proteins and RNA molecules. Although many components of various cellular bodies are known, general principles that define body composition are lacking. We have proposed a model for the formation of cellular bodies that is based on the polymerization-driven phase separation of key scaffold components of cellular bodies. We modeled cellular bodies using several engineered multivalent proteins and RNA. *In vitro* and in cells these scaffold molecules form phase separated liquid droplets that are strongly enriched with the scaffold molecules. Analytical theories of polymerization suggest the resulting second phase contains large polymers of the multivalent scaffolds. Low valency client molecules partition differently into these structures depending on the stoichiometric ratio of the scaffolds, with a sharp switch in recruitment across the phase diagram diagonal. Composition can switch rapidly through changes in scaffold concentration or valency. Natural PML NBs and P bodies show analogous partitioning behavior, suggesting how their compositions could be controlled by levels of PML SUMOylation or cellular mRNA concentration, respectively. Indeed, the engineered polySUMO/polySIM engineered scaffolds recruit many of the natural PML NB clients in a manner that depends on the SUMO:SIM stoichiometric ratio. Together, these data suggest a conceptual framework for considering the composition and control thereof of cellular bodies assembled through heterotypic multivalent interactions.

Table of Contents

Acknowledgments	v
Abstract	viii
Publications	xii
List of Tables	xiii
List of Figures	xiv
Chapter 1. Introduction	1
1.1 Cellular Bodies	1
1.2 Mechanisms of Assembly	2
1.2.1 Folded Domains	5
1.2.2 Intrinsically Disordered Regions	7
1.2.3 Cooperativity between IDRs and Folded Domains	9
1.2.4 Transition to Fibrillar States, Maturation, and Disease	10
1.3 Regulation of Assembly	13
1.3.1 Molecular Regulation of Assembly	13
1.3.2 Effects of Energy-Dependent Processes	14
1.4 Regulation of Function	15
1.4.1 Mechanisms of Compositional Control	15
1.4.2 Regulation of Biochemical Processes	17
Chapter 2. Gelation in Phase Separated Droplets	19
2.1 Rationale	19
2.2 Cyclization in the sol-gel transition	19
2.3 Discussion	21
Chapter 3. Compositional Control of Phase Separated Cellular Bodies	24
3.1 Rationale	24
3.2 Scaffold Stoichiometries Dictate Client Recruitment	25
3.3 Valency of Client Affects Client Recruitment	31

3.4	Mass Action Explains Switch-like Partitioning of Low Valency Clients	31
3.4.1	Overview of the Mass Action Model	36
3.4.2	Behavior of the Model	39
3.4.3	Fitting to Data	42
3.4.4	Quantitative Limitations of the Model	42
3.4.5	Qualitative Limitations of the Model	44
3.5	Compositional States Interchange on Cellular Timescales	45
3.6	Engineered Cellular Puncta Selectively Concentrate Low Valency Clients . . .	48
3.7	Scaffold Stoichiometries Control Client Recruitment into Natural Cellular Bodies	52
3.8	Discussion	57
3.8.1	Hierarchical Organization of Cellular Bodies	57
3.8.2	Complexities of Natural Cellular Bodies	59
3.8.3	Biological Mechanisms of Regulating Body Composition	60
3.8.4	Changes in Body Composition may Dictate Changes in Function . . .	61
3.8.5	Implications for Cellular Body Function	62
3.8.6	Conclusion	62
Chapter 4. Reconstitution of Functional Cellular Bodies		64
4.1	Rationale	64
4.2	Recruitment of PML NB Components	65
4.3	Cellular Functions of PML	66
4.3.1	Antiviral Response	68
4.3.2	Apoptotic Sensitization	70
4.3.3	Regulation of Cellular SUMOylation	71
4.3.4	Regulation of Transcription	72
4.4	Discussion	73
4.4.1	Reconstitution of PML NB Composition	73
4.4.2	Reconstitution of PML NB Function	74
4.4.3	Conclusion	76
Chapter 5. Concluding Remarks		77
5.1	Practical Considerations in Studying Phase Separated Structures	77
5.1.1	Biochemical considerations	78
5.1.2	Genetic considerations	79
5.2	Remaining Questions	80
5.2.1	The Function of Cellular Bodies	80
5.2.2	Structural Organization of Cellular Bodies	81

5.2.3 Cellular Body Composition	81
5.2.4 The Relationship of Cellular Bodies to Disease	81
5.2.5 Generality of Phase Separation in Biology	82
5.3 Conclusion	82
Appendices	84
Appendix A. Experimental Procedures	85
A.1 Genes, RNA, and Plasmids	85
A.2 Modules and Proteins	85
A.3 Protein Expression and Purification	87
A.4 Phase Diagram Mapping	89
A.5 Partitioning Assays	89
A.6 Image acquisition and analysis	90
A.7 Isothermal Titration Calorimetry	90
A.8 Mass Action Model for Client Partitioning	91
A.9 Fluorescence Recovery After Photobleaching	91
A.10 Immunofluorescence Staining of Endogenous PML	92
A.11 Western Blotting of Transfected PML	92
A.12 Viral Assays	92
A.13 Apoptosis Assays	92
A.14 2D Gel Electrophoresis	93
A.15 RNASeq	93
Appendix B. Image Analysis and Quantification of Partitioning	94
B.1 Preprocessing	94
B.2 Identification of Droplet and Bulk Regions	94
B.3 Calculation of PC or IR	95
Bibliography	97

Publications

1. **Salman F Banani***, Hyun O Lee*, Anthony A Hyman, and Michael K Rosen. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol.*, 18(5):285-298, May 2017.
2. **Salman F Banani**, Allyson Rice, William Peeples, Yuan Lin, Saumya Jain, Roy Parker, and Michael K Rosen. Compositional control of phase-separated cellular bodies. *Cell*, 166(3):651-663, July 2016.
3. Pilog Li, Sudeep Banjade, Hui-Chun Cheng, Soyeon Kim, Baoyu Chen, Liang Guo, Marc Llaguno, Javoris V Hollingsworth, David S King, **Salman F Banani**, Paul S Russo, Qiu-Xing Jiang, B Tracy Nixon, and Michael K Rosen. Phase transitions in the assembly of multivalent signalling proteins. *Nature*, 483(7389):336-340, March 2012.

*, equal contributions

List of Tables

3.1	Apparent SUMO-SIM Affinities as a Function of Valency	30
3.2	Values of Free Parameters (Client-Scaffold Affinities in Bulk and Droplet) from Fitting of Partitioning Data to Mass Action Model .	42
3.3	Values of Rate Constants from Single-Exponential Fits of FRAP Experiments	45
A.1	Sequences of Modules and Proteins Used in this Study	86

List of Figures

2.1	Microrheology of polySUMO/polySIM droplets	23
3.1	Phase Diagram Position Dictates Client Recruitment	26
3.2	Partitioning into polySUMO/polySIM Droplets Requires SUMO-SIM Binding Interactions	27
3.3	Client Valency Affects Partitioning	29
3.4	Apparent Affinity of SUMO to SIM Increases as a Function of Valency	30
3.5	A Mass Action Model Predicts Client Partitioning Behavior	32
3.6	Subtle Features in Concentration Changes of Scaffolds affect Qualitative Behaviors of Predicted Client Partitioning	33
3.7	Mass Action Model Qualitatively Recapitulates Partitioning Behavior of Low Valency Clients but has Quantitative Limitations	35
3.8	Droplets Interchange Composition On Cellular Timescales Without Compromising Structural Integrity	46
3.9	Rearrangement of Scaffold Polymers Limit the Rate of Compositional Interchange in Droplets	47
3.10	Cellular PolySUMO-PolySIM Puncta Selectively Recruit Low Valency Clients	49
3.11	Cellular Poly-SUMO-PolySIM Puncta Selectively Recruit Low Valency Clients in 3T3 Cells	50
3.12	<i>In cis</i> Scaffold Behavior is Analogous to the <i>in trans</i> Scaffold System (polySUMO and polySIM)	51
3.13	Partitioning of Clients into PML NBs Requires SUMO-SIM Binding and Depends on Levels of PML SUMOylation	53
3.14	Client Recruitment Into Natural Cellular Bodies Is Affected By Scaffold Stoichiometries	55
3.15	A Model for Compositional Control of Cellular Bodies	58
4.1	Synthetic PolySUMO/polySIM Nuclear Bodies Recruit Natural PML NB Components	67
4.2	Reported Cellular Functions of PML NBs Fail to Recapitulate in <i>PML</i> ^{-/-} MEFs	69
4.3	Global Gene Expression Profiling in <i>PML</i> ^{-/-} MEFs Fails to Detect Significant Changes in Expression Due to PML	72

Chapter 1

Introduction

This chapter contains material adapted from previously published manuscripts [7, 6].

1.1 Cellular Bodies

Eukaryotic cells compartmentalize biological processes to achieve spatial and temporal control over biochemical reactions. Compartmentalization has long been studied in the context of membrane-bound organelles, where mechanisms of biogenesis and transport of molecules into and out of the organelle are well understood. Cells also contain numerous membrane-less organelles, collectively referred to as cellular bodies [99]. These structures, which include P granules, P bodies, nucleoli and Promyelocytic Leukemia nuclear bodies (PML NBs), are micron-sized assemblies of proteins and often RNA found in the cytoplasm and nucleoplasm of eukaryotic cells. They appear to be functionally important, as inferred from their conservation among evolutionarily distant species [61] and their tendency to concentrate functionally related groups of molecules [99, 101].

Ultrastructural analysis of cellular bodies suggests they are porous structures with densities comparable to those of the nucleo- or cytoplasm [61]. Analysis in live cells has revealed that macroscopically, the bodies persist for hours to days. Yet, they are highly dynamic at the molecular level, turning over their contents within seconds to minutes [40, 144]. Recent work has demonstrated that bodies exhibit liquid-like properties [17, 18, 138, 86, 110, 33, 39]. These and other behaviors suggest that cellular bodies are condensed phases that form through liquid-liquid phase separation of the nucleo- or cytoplasm [72, 71].

Cellular bodies are often enriched in multivalent molecules [91]—proteins that harbor multiple modular domains or stretches of low-complexity amino acid sequence with repeated interaction motifs [37, 116, 59, 81] or charged elements [107, 43]; RNA species that contain multiple protein-binding elements; or combinations thereof. Interactions between multivalent macromolecules can drive polymerization and concomitant phase separation [91, 107, 54, 8, 100], resulting in the formation of a condensed, droplet phase suspended in the bulk solution phase. It has been suggested that this fundamental macromolecular behavior may be an important principle governing the organization of cellular bodies [91, 54, 100, 107]. Indeed, expressing engineered multivalent proteins or ectopically tethering high copy numbers of body components (high local valency) in cells is sufficient to form dynamic, membrane-less puncta that resemble bona fide cellular bodies [91, 79, 34, 98, 126].

1.2 Mechanisms of Assembly

Cellular bodies are enriched in multivalent molecules [91]. Fundamental concepts in polymer chemistry indicate that multivalent molecules can assemble into large oligomers or polymers, and that this assembly process will inherently decrease the solubility of the resulting assemblies, promoting their phase separation. The coupled assembly and phase separation of multivalent macromolecules has emerged as an important organizing principle for cellular bodies. To understand how this molecular behavior can mechanistically explain the formation of cellular bodies (and ultimately their regulation, composition and function), we consider here both the polymerization and the solubility of multivalent molecules and the factors that can modulate these properties.

We first consider the solubility of molecules, which drives phase separation. All macromolecules exhibit varying degrees of weak, non-specific interactions with each other and with solvent (water, in biology). These interactions tend to be very low in affinity, short-lived and distributed throughout the surface of the molecule. They characteristically

lack stereospecificity and are too weak to be measured by conventional techniques that quantify binding affinity. Essentially, solubility is the balance between the weak interactions between macromolecules versus those between the macromolecule and water. When the interactions between macromolecules are weaker than those between macromolecules and water, the macromolecules remain miscible in solution. However, when the macromolecule-macromolecule interactions are sufficiently stronger than macromolecule-water interactions, the macromolecule reaches its solubility limit and phase separation occurs [50, 69]. That is, the system separates into two phases: a large-volume, low concentration dilute phase, and a small-volume, high concentration condensed phase. Although the resulting condensed phase can be either solid or liquid, the properties of biological molecules often result in phases with liquid-like properties (with exceptions, see below). The switch from miscibility to immiscibility occurs very sharply with changes in macromolecule concentration or solution conditions (temperature, salt, pH, etc.) that tip the balance of interactions.

We next consider the oligo- or polymeric assembly of multivalent molecules. Biological macromolecules also form complexes through relatively long-lived interactions that occur with high affinity and high stereospecificity. Contrary to the interactions governing solubility, the affinity of these strong interactions can readily be quantified experimentally. Such interactions have long been the basis of understanding macromolecular recognition and its role in biochemical and biological function. A given molecule and its ligand may contain any number, or valency, of elements that engage in these kinds of strong interactions, which in turn dictates the properties of the resulting complexes. The behaviors of multivalent interaction systems can be described by classical theories in polymer chemistry [131, 51]. Molecules with sufficiently high valency (≥ 3) are capable of forming a distribution of complexes with varying stoichiometries. For a fixed valency, the average size of this distribution increases with the degree of binding that is, the degree to which all the sites in the system are saturated. For a fixed concentration of interacting modules, the size distribution also increases

as the modules are connected to higher valencies. Above a critical degree of binding, these systems can generate large, macroscopic polymers. The transition from small oligomers to macroscopic structures is termed the sol-gel transition or gelation, and is the basis for many modern polymeric materials Wang:2015iy. For example, agarose gels used to analyze nucleic acids in molecular biology represent a familiar example.

A key concept from polymer chemistry mechanistically connects the interactions governing solubility with those governing polymerization. In poor solvents where polymer-polymer interactions are stronger than polymer-solvent interactions the solubility of a complex decreases as its size increases. This phenomenon arises because the entropic cost of confining molecules into the condensed phase is less when the molecules are bound together than when they are individual. This phenomenon may also be viewed as the increased avidity of the weak, solubility-driving interactions as their effective valency is increased by polymerization. Thus, phase separation is promoted by polymerization. Experimentally, this has been demonstrated for gamma-crystallin [3] and lysozyme [140], which both phase separate at lower concentrations (normalized by mass) when crosslinked into oligomers than when they are monomeric. For non-covalently associating systems, this relationship provides a critical link between the weak interactions that govern solubility and the stronger interactions that mediate the formation of non-covalent complexes, polymers, or gels. Thus, for multivalent molecules, the processes of polymerization and phase separation are energetically coupled. On one hand, polymerization, by increasing the average size of complexes, can enhance the weak, non-specific interactions thereby decreasing solubility and promoting phase separation. On the other hand, because phase separation concentrates molecules into a condensed phase, it increases the degree of binding in that phase, thereby promoting formation of larger complexes.

It is important to emphasize that although these processes may be coupled, polymerization and phase separation are fundamentally distinct and can occur independently.

For example acrylamide polymerizes and undergoes gelation while remaining as a single, homogeneous phase; lysozyme phase separates at high concentrations without undergoing polymerization. Nevertheless, in most multivalent biological systems examined, the two processes go hand in hand.

We note that, to our knowledge, previous conceptions of the assembly of multivalent macromolecules have focused largely on the networks created by strong interactions, without consideration of the weak interactions that govern solubility, and how they would be affected by the assembly process. We argue that the coupling between the strong and weak interactions, and thus the ability of multivalency to promote phase separation, is essential to understanding the behavior of multivalent biological molecules [124]. These basic ideas can be applied to understand the phase separation behavior of numerous biological systems. It is becoming apparent that many biological systems, especially those comprising cellular bodies, exhibit phase separation that occurs cooperatively and concomitantly with multivalency-driven polymerization. We review such systems below and describe the mechanisms by which they undergo these processes.

1.2.1 Folded Domains

Biological systems are often driven by interactions between folded, modular protein domains and well-defined ligands. When both species have high valency, such systems have the propensity to polymerize and phase separate. This behavior has been observed for a variety of engineered macromolecules composed of tandem repeated elements connected by flexible linkers. Examples include polySH3 proteins binding to polyProline-Rich-Motif (polyPRM) ligands, polyRNA-binding-domain proteins binding to repeated RNA oligonucleotides and polySUMO proteins binding to polySUMO-Interaction-Motif (polySIM) ligands [91, 7]. In the latter system, phase separation has been observed both when the cognate multivalent molecules interact *in trans* and when they are connected *in cis* [7]. Physical studies

of these systems demonstrate that higher valency promotes assembly into larger structures and enables phase separation at lower module concentrations, consistent with the idea that assembly promotes phase separation. This valency dependence of phase separation also suggests powerful mechanisms of controlling phase separation *in vivo*, as discussed below. Further, the dynamics of molecules within phase separated droplets in the polySH3-polyPRM system, as measured in fluorescence recovery after photobleaching experiments, correlates with the monomeric SH3-PRM binding affinity, consistent with the presence of dynamic polymers within the droplets whose rearrangements are limited by dissociation of individual SH3 domains from their PRM ligands.

Similar behaviors have also been demonstrated with several natural multivalent protein systems. The actin-regulatory signaling pathway consisting of the multivalent proteins Nephrin (which contains three phosphotyrosine motifs), its ligand Nck (which recognizes phosphotyrosines with its single SH2 domain and also contains three SH3 domains), and the Nck ligand N-WASP (which contains several PRMs, ligands for SH3 domains), produces both three-dimensional phase separated liquid droplets when all components are in solution [91], as well as analogous two-dimensional phase separated puncta when Nephrin is attached to membranes [8]. The latter have been observed both *in vitro* [8] and in cells [?]. An analogous system controlling actin in T cells also forms membrane puncta *in vitro* and in cells in response to stimulation of the T cell receptor [133]. There, the tyrosine-phosphorylated adaptor protein LAT, binds its SH2- and SH3-containing ligands Grb2 and Gads, which in turn bind their PRM-containing ligands SOS and Slp76. In both the Nephrin and LAT pathways, as in the engineered systems above, the concentrations necessary for phase separation depend on the valency of tyrosine phosphorylation, as well as the valency of SH3 domains (in Nck and Grb2, respectively). Further, in the Nephrin system, phase separated puncta or droplets can be disrupted with high affinity mono-valent pTyr competitors, further supporting the model of assembly-driven phase separation. Finally, in a biologically and molecularly

unrelated system, the P body components Dcp2 and Edc3, which bind through interactions of the helical leucine motifs in the former with the Lsm domains in the latter, also phase separate in a valency-dependent manner, supporting the generality of these behaviors for multivalent proteins [54].

Not all biological multivalent systems undergo phase separation when they assemble into large structures, however. As an example, a series of engineered proteins consisting of tandem repeated WW domains readily undergo polymerization and gelation when mixed with multivalent PRM-containing partners. However, under all conditions reported, the system remains a single, macroscopically homogeneous phase [53, 104]. Although not explored in this work, the combined surface features of the WW-PRM complexes and properties of the interdomain linkers evidently generated sufficiently strong interactions with solvent that the system did not phase separate at any oligomer/polymer size. These observations further illustrate the idea that assembly/gelation and phase separation of multivalent systems are distinct phenomena, even if often coupled. It remains to be seen how frequently this simpler behavior is observed in multivalent biological systems.

1.2.2 Intrinsically Disordered Regions

Many eukaryotic proteins contain intrinsically disordered regions (IDRs) with functional binding elements [5]. The binding elements that contribute to phase separation may be disordered. With sufficient valency, these elements, much like folded domains, can promote assembly and phase separation. Indeed, IDRs from many cellular body components are sufficient for [42, 102, 24, 107], or to contribute to [149, 100], phase separation *in vitro*.

IDRs promote phase separation on either their own or in the presence of binding partners. For the former class, aromatic residues appear to play an important role in phase separation. Many of the proteins that localize to RNA granules—for example, Fus, Taf15, Tdp43, Tia1, et al.—contain regions that are highly enriched in a small number of amino acid types,

including aromatic residues. These IDRs contain multiple [GS][FY][GS] sequences and/or poly-Q/N tracts [116, 81, 84] and are important for targeting the corresponding molecules to RNA granules *in vivo* [81] and for phase separation *in vitro* [24, 102, 95]. Aromatic residues in these IDRs appear to be important for phase separation.

The IDRs from Fus, Taf15, hnRNPA2, EWS, and CIRBP also form solid-like hydrogels *in vitro* [81, 59, 87]. Based on a combination of X-ray diffraction and electron microscopy data, these hydrogels contain long filaments that appear to be generated from cross-beta interactions of the polypeptide backbone, typical of those observed in numerous amyloid fibers [59, 81]. By chemical footprinting methods, the hnRNPA2 IDR exhibits a residue accessibility signature consistent with cross-beta interactions in both the hydrogel state as well as in the liquid droplet state [148], suggesting that the cross-beta interactions that drive hydrogel and fiber formation may contribute, at least in part, to the phase separation of these molecules that is, the same structures that generate amyloid fibers when occurring thousands at a time, may, when occurring only a few at a time, provide the weak multivalent adhesions that drive phase separation. Though the mechanistic basis of the role of aromatic residues in this context remains unclear, these residues likely stabilize the backbone cross-beta interactions observed in these proteins.

Aromatic residues may also promote phase separation via interactions between residues (as opposed to backbones). The IDR in the nuage protein Ddx4, for example, contains numerous FG repeats that likely engage in cation-pi interactions with the RG repeats within the same IDR. Moreover, IDR phase separation may also be promoted by repetitive patterns of charged residues, which could provide interactions that are structurally less defined than cross-beta elements. For example, as mentioned, Ddx4 is enriched in RG repeats that engage in cation-pi interactions. The P granule protein Laf-1 also contains within its IDR multiple RG blocks that may engage in cation-pi interactions (as in Ddx4) with the numerous GY repeats in Laf-1. Moreover, Laf-1 also contains numerous blocks of negatively charged

residues that intervene and may interact with the RG blocks to satisfy charge neutrality.

In contrast, the model IDR comprising of disordered C-terminal domain of Nephrin (NICD) contains alternating blocks of negatively charged and hydrophobic residues, with few positively charged residues [108]. Thus, while the IDRs in Laf-1 and Ddx4 readily undergo phase separation *in vitro* on their own [43, 107], NICD requires the presence of cationic molecules to phase separate [108]. This process of *in trans* charge neutralization is referred to as *complex coascervation*, and it may be a commonly utilized mechanism that promotes phase separation of cellular body components. For example, the disordered acidic tracts in NPM1, known to interact with proteins containing R-motifs, are required for phase separation [100] and likely promote phase separation through the process of complex coascervation.

IDRs can thus undergo a variety of types of homotypic and heterotypic interactions. The individual interacting motifs may be less well-defined than those in folded domains. However, together they may nevertheless be thought of as contributing to the multivalency of the system. Though further investigation is needed to pinpoint the structural basis of these interactions, they likely utilize a combination of charge-charge (including complex coascervation), cation-pi, pi-stacking, dipole-dipole interactions [19].

1.2.3 Cooperativity between IDRs and Folded Domains

Cellular body components often contain a combination of folded domains and IDRs. For example, Fus, Laf-1, Whi3, and Nucleophosmin each contain IDRs as well as folded domains involved in RNA binding. Multivalent folded domains without IDRs readily phase separate *in vitro* [91], as do IDRs without folded domains [42, 102, 24, 107]. When both of these types of interactions are present within the same system, they can act cooperatively to favor phase separation. For example, fusing a prion-like IDR to multivalent RRM domains increases the propensity of the molecule to phase separate with a multivalent RNA ligand for RRM [95]. Moreover, while Whi3, Laf-1, and hnRNPA1 can each phase separate on their

own, their propensity to do so is increases in the presence of an RNA ligand [149, 43, 102].

Cooperativity may occur between any type of multivalent interactions within the same system. For example, both the nucleolar protein NPM1 can phase separate either with cationic peptides that bind to its N-terminal acidic tracts; with nucleic acids that bind to its C-terminal nucleic acid binding domain; or both. Thus, in general, multiple types interacting elements, whether folded or disordered, can cooperate to contribute to the multivalency of the system [95].

1.2.4 Transition to Fibrillar States, Maturation, and Disease

A large body of data has demonstrated that in the absence of inhibitory electrostatic interactions, most unfolded polypeptides can assemble over time into thermodynamically stable amyloid-like fibers [47]. Thus, when concentrated into phase separated droplets, which often requires charge neutralization (§1.2.2), prion-like IDRs should have a strong propensity to form such fibers. This feature may distinguish phase separated IDRs from phase separated modular domains, which should not form amyloids if they remain well-folded. Thus for IDRs, superimposed on the drive to phase separation (a thermodynamic process) is the propensity to nucleate and form amyloid-like fibers (a kinetic process). Such behaviors have been observed for several IDR systems, suggesting important implications for biological regulation.

NMR experiments characterizing phase separated droplets of the Fus IDR suggest that the protein remains disordered in the phase separated state, as it is in the dispersed monomeric state [24]. Moreover, both Fus and hnRNPA1 droplets have rapid turnover rates by FRAP experiments and macroscopically exhibit liquid-like behaviors [24, 102, 110]. However, droplets composed of many types of IDRs (e.g., from Fus, Pub1, Lsm4, eIF4GII, or Tia1) or full length IDR-containing , proteins (hnRNPA1, Fus, Whi3) while initially dynamic, cease to be liquid-like over the course of several hours, concomitant with macroscopic

formation of filamentous structures a process that has been referred to as maturation [95, 149].

These behaviors suggest that within condensed phases the IDRs, at least initially, engage in transient, dynamic interactions and do not appear to contain large-scale, stable, ordered structures. Yet, when concentrated into droplets, sometimes as much as several-hundred fold [95, 108], the propensity of IDRs to engage in additional cross-beta interactions or form amyloid fibers should increase, through increased rate of nucleation and/or growth from pre-existing nuclei. This occurs due to both the increased concentration of the IDR in the condensed phase, as well as its ability to sample more extended conformations in the favorable environment of the condensed phase, which promotes intermolecular contacts [58]. Thus, the formation of an amorphous, vitrified network of cross-beta interactions and/or amyloid fiber formation may be the structural basis for droplet maturation, generating hydrogels analogous to those formed by these proteins upon concentration [95, 149, 102, 110, 58]. The material properties of the macroscopic structure at any point in the maturation process likely result from a combination of the fiber length distribution, the strength of the individual cross-beta interactions, and the degree of fiber crosslinking through lateral interactions. However, experimental perturbations can uncouple the phase separation equilibrium and the kinetics of fiber formation. For example, a mutant of hnRNPA1 IDR that lacks a steric zipper motif, known to be crucial for the fibrilization of the molecule, readily undergoes phase separation, but these droplets do not mature as readily [95, 102].

Cells must have mechanisms to limit the natural tendency of IDRs to mature in order to tune the dynamics and liquid-like properties of cellular bodies into functionally appropriate regimes. One possible mechanism to provide such resistance involves the use of energy-dependent processes or machines to control the degree of fiber formation within bodies, limiting fibers when bodies need to be dynamic and enabling their formation and/or growth when bodies need to be static. Indeed, chaperones/disaggregases/ATPases are present in many RNA granules. In the case of stress granules, they play a role in maintaining the

dynamic properties of the structures [75, 86]. As discussed below, it thus appears that cells expend energy (by consuming ATP) to control the liquidity of cellular bodies [86, 21, 75, 18].

Granule maturation may be relevant to the pathogenesis of neurodegenerative disease. Neurodegenerative diseases—including amyotrophic lateral sclerosis (ALS), frontotemporal lobar degeneration (FTLD) and multisystem proteinopathy (MSP), among others—constitute a family of as yet untreatable diseases, in which patients present with a range of neurological and/or musculoskeletal abnormalities [113, 49]. Pathological protein/RNA aggregates, or “plaques”, in the cytoplasm or nucleus appear to be a hallmark of many of these diseases [78]. These collection of proteins found in these plaques includes many RNA binding proteins with self-assembling IDRs, many of which are found in RNA granules such as P bodies and stress granules [116, 59, 81, 84]. However, in disease states, these pathological assemblies lack rapid molecular dynamics, reminiscent of matured droplets described above, and fail to be properly cleared from cells by proteostatic mechanisms [142, 115]. Further, mutations that give rise to familial forms of these diseases frequently occur in RNA-binding proteins that localizing to RNA granules or in proteins that influence RNA granule dynamics [93, 115]. Mutants of hnRNPA1, Fus, and TDP-43 found in patients with familial forms of ALS undergo fibrilization much more readily than their wild-type counterparts [95, 102, 110]. Moreover, mutations that impair the ATPase VCP, involved in maintaining the dynamics of stress granules in cells, are causative for MSP [115]. These studies suggests that dysfunctions stress granule dynamics may play a role in the pathogenesis of these diseases. Since the implicated molecules are involved in RNA homeostasis, neurodegenerative diseases may ultimately be caused by dysfunctions in RNA metabolism due to dysregulation of RNA granules [115].

The solid-like state of granules, however, may not be limited to disease. Electron and super-resolution light microscopies as well as recent biochemical studies have all revealed that cellular bodies contain substructures that seem to be solid-like in their makeup [15, 18,

138, 75, 90]. These may represent fibrous structures embedded within a larger liquid phase. Though the function of these sub-compartments remains to be determined, it is conceivable how cells could regulate the relative amounts of granular material versus liquid material to functional effect.

1.3 Regulation of Assembly

1.3.1 Molecular Regulation of Assembly

The critical concentration for phase separation depends on molecular properties including valency and the binding affinity between interacting elements [91]. These dependencies immediately suggest mechanisms to control the formation and disassembly of cellular bodies. Such mechanisms span timescales from the rapid molecular changes that can occur physiologically to the slow changes to genes that occur during the process of evolution.

Physiological changes (seconds to minutes) can alter the valencies and affinities of cellular body components. Post-translational modifications can rapidly control valencies by addition or removal of moieties that participate in binding reactions between scaffolds. This has been demonstrated in both the Nephrin and LAT signaling systems, where increasing the number of phosphorylated tyrosine (pTyr) residues on both proteins promotes their phase separation with downstream SH2-containing ligands [133, 91, 8]. In both systems this effect has been shown through measuring the critical concentration as a function of pTyr valency, as well as in real time through monitoring of the appearance and disappearance of phase separated structures driven by kinases and phosphatases, respectively. Analogously, phase separation of the nuage protein Ddx4 is hindered by the methylation of arginine residues, which likely decreases the number of effective cation-pi interactions [107]. Moreover, in some cell types, the number and structure of PML NBs are influenced by the degree of SUMOylation of the PML protein [12, 46, 38, 63].

Cellular changes on slower timescales (hours to days) may also regulate body assem-

bly. Alternative splicing may be a facile way to control valency (although it has not yet been observed as a means of regulation for cellular bodies). Moreover, increasing the expression of a high-affinity, low valency competitor could disrupt the polymerization and concomitantly body assembly (also not observed naturally). Finally, regulating the expression levels of key assembly components about their critical concentrations may correspondingly regulate body assembly. For example, experimentally increasing the concentration of untranslated mRNAs, by blocking their degradation increases the size and numbers of P bodies in yeast [109].

On the much slower evolutionary timescales, changes in the genetic sequences could contribute to changes in valency, affinity, and perhaps even solubility of molecules to regulate body assembly. Thus, understanding the molecular basis of phase separation provides insights into how this process may be regulated by mechanisms accessible to the cell and to the evolutionary process.

1.3.2 Effects of Energy-Dependent Processes

The mechanisms of assembly and phase separation that we have described above are based entirely on equilibrium thermodynamics. However, living cells are maintained far from equilibrium through the consumption of energy. The study of energy consuming, non-equilibrium materials—“active matter”—is an area of current interest in physics [11, 20]. It is likely that energy utilization modulates the equilibrium behaviors of cellular bodies, impacting their assembly, material properties and ultimately functions. In some cases, this may involve ATP-dependent molecular machines acting directly on body components. For example, the transcription of rRNA, a key component of nucleoli, influences the nucleation and spatial distribution of condensing nucleoli in *C. elegans* embryos [11]. Additionally, depleting ATP leads to loss of liquid-like macroscopic dynamics in nucleoli concomitant with decreased mobility of their component molecules [75, 18]. Perhaps relatedly, chaperones or

disaggregases (e.g. VCP or Cdc48) are required for the maintenance of the dynamic, liquid-like state of RNA granules [21, 86]. Together, these properties suggest cellular body behavior is dictated by an underlying phase separation equilibrium that may constantly be modulated by energy-consuming processes in cells.

1.4 Regulation of Function

1.4.1 Mechanisms of Compositional Control

The compositions of most cellular bodies are dynamic. While some components are constitutively concentrated in a given body, many components are recruited only during particular stages of the cell cycle or in response to particular stimuli [57, 38, 67, 22, 96]. Such compositional regulation is likely critical to cellular body function. We recently described a conceptual framework that provides insight into how cellular body composition may be controlled [7]. This framework is based on a qualitative separation of body components into two distinct classes. Of the 10s-100s of components that localize to cellular bodies [23, 52], where examined, only a small subset is typically required for body assembly. For example, the PML protein is essential for the formation of PML NBs [73], and the NEAT1 non-coding RNA is essential for paraspeckles [35]. We define these key components as *scaffolds*. While other bodies appear to have more than a single essential component, the concept that such components constitute the minority of body components appears to be general. The remaining majority of components appear to be dispensable for body assembly. We define these components as *clients*. Clients often localize to bodies through regulated interactions with scaffolds, and reside there only under specific conditions [57, 38, 67, 22, 96]. These observations suggest that the composition and consequent functions of cellular bodies can be regulated.

The general properties of compositional control can be understood using simple multivalent model systems composed of multivalent scaffolds and their cognate low valency

clients [91, 7]. Studies of these systems have elucidated a generalizable framework for compositional control of cellular bodies. This model, based on mass action, suggests that the relative stoichiometries of scaffold components will dictate whether and which type of clients are recruited to the body [7]. Moreover, the framework suggests plausible mechanisms by which cellular body compositions could be regulated in cells via changes in scaffold expression levels or by post-translational modifications of the scaffolds or clients. This framework is described in more detail in Chapter 3.

Recruitment of low valency clients has been observed in a few different systems in addition to the model multivalent systems mentioned above. For example, Tubulin can be recruited to droplets formed by the spindle matrix protein BuGZ *in vitro* by interacting with the microtubule binding domain on BuGZ [77]. And various IDR-containing clients can be recruited to droplets whose scaffolds also contain IDRs [95, 148] *in vitro* or to stress granules in cells [81]. Moreover, in two natural systems PML NBs in mammalian nuclei and P bodies in yeast cytoplasm perturbing scaffold stoichiometries resulted in changes in client recruitment consistent with our mass action model. Removing several of the PML SUMOylation sites switched the preference of PML bodies from a SIM-recruiting to a SUMO-recruiting state. And increasing the total concentration of untranslated mRNAs in yeast was sufficient to increase the recruitment of Xrn1 into P bodies. Thus, despite the complexity of natural cellular bodies, this study suggests that their compositions may be governed by simple mass action.

Clients may also be recruited by binding to scaffold elements that are not involved in scaffold assembly, or by exhibiting appropriate physicochemical properties (e.g. surface charge or size). For example, negatively charged molecules were excluded from phase separated membrane clusters of the T Cell signaling protein LAT with its downstream effectors, while positively charged molecules were enriched [133], a process that may be important for the regulation of T Cell signaling. These changes in enrichment, however, were of a smaller

magnitude than the changes observed in client recruitment mediated by specific binding interactions, so the latter mechanism may play a dominant role in determining the composition of a given body.

Such compositional regulation could afford cellular bodies the ability to regulate biological pathways by a variety of mechanisms, as discussed below.

1.4.2 Regulation of Biochemical Processes

The characterized physical properties of cellular bodies suggest they are compatible with biochemical processes occurring within them. They are porous structures with comparable densities as their surrounding environments [61]. They exhibit liquidity and rapid dynamics (timescales of seconds to minutes), both internally and with respect to exchange with the surrounding environment [110, 18, 17, 33, 86, 139, 144, 40, 111]. These properties can enable substrates and products of reactions to diffuse into and out of the bodies, allowing the bodies to function as catalytic sites in cells.

Although chemistry within condensed phases has yet only been examined in model phase separating systems [132, 25, 133] and not in any bona fide cellular bodies, the model systems have provided significant insights. Cellular bodies compartmentalize their components by concentrating them many-fold over the surrounding bulk. This property could be functionally utilized by cells by a variety of mechanisms. First, bodies could accelerate reactions by concentrating components. Simple calculations suggest that when both an enzyme and its substrate are concentrated (but not when either is concentrated alone) within the condensed phase, the overall reaction rate of the system should increase. Indeed, the total activity of a self-splicing RNA increases when it is concentrated into one phase of a model aqueous two-phase system [132]. And actin polymerization rates can be substantially accelerated by concentrating Arp2/3 and N-WASP into droplets or phases on model membranes (though factors beyond concentration may also contribute to this effect) [91, 8, 133]. Second,

Colocalization of molecules that act in the same biological pathway—enzyme-coenzyme systems, for example—and concomitant exclusion of components of alternative reactions would enhance flux through the pathway, especially if the molecules function cooperatively. Colocalization may also enhance specificity within metabolic or signaling pathways if desired enzyme-substrate or allosteric regulator-target pairs are co-recruited to the bodies. Consistent with this, clustering the branch point enzyme CarB with a downstream enzyme PyrB and not ArgI could redirect the metabolic fate of the carbamoyl phosphate to pyrimidine from arginine in *E. coli* [29]. Finally, chemistries may be modulated within cellular bodies via their physical properties. These studies suggest that beyond the simple effects of concentration, the physical nature of biochemical reactions may itself be different within bodies as compared to the surrounding bulk. Several cellular body properties may affect biochemical reactions within them. Diffusion of particles, for example, within polymeric networks exhibits a strong dependence on the size of the particle relative to the size of the polymer meshwork [26]. The diffusion of molecules of different sizes will be influenced differently by polymer matrix; molecules smaller than pore size of matrix will be largely unaffected, while assemblies larger than pore size will be greatly slowed and move at rates dictated by polymer fluctuations. These size-dependent constraints within the bodies could potentially influence the dynamics of reactions differently than in the surrounding bulk. Moreover, since clients bind to relatively slowly diffusing scaffolds, catalysis by enzymes bound to relatively immobile surfaces may give rise to reaction properties not observed in the cytoplasm.

The biochemical environment within the bodies may therefore be fundamentally different than that in the surrounding cytoplasm or nucleoplasm, and it may afford the cells access to unique ways of regulating cellular reactions.

Chapter 2

Gelation in Phase Separated Droplets

This chapter contains material adapted from a previously published manuscript [91].

2.1 Rationale

Multivalent molecules have the propensity to form large polymeric assemblies by undergoing sol-gel transitions. Our model for cellular body is based on phase separation that occurs concomitantly with polymerization. But have the polymerizing multivalent molecules undergone a sol-gel transition, or have they simply formed large, oligomeric species that are not large enough to be considered gels [51]? Addressing this questions may provide insights into the macroscopic material properties of cellular bodies, and how they may be controlled in cells (see §1.2.4) [104, 137].

Here, I address this question by using the theory of sol-gel transitions [51] to estimate the likelihood the polymers within the droplets formed by model SH3/PRM systems [91] are gels.

2.2 Cyclization in the sol-gel transition

Consider a generic two-component system capable of undergoing condensation polymerization with monomers of configuration \mathbf{A}_f and \mathbf{B}_g , where f and g designate the valencies of \mathbf{A} and \mathbf{B} modules, respectively, and where \mathbf{A} modules bind to \mathbf{B} modules in a 1:1 fashion. Systems of this configuration have the propensity to polymerize and yield a distribution of polymer complexes of different sizes, as well as to undergo a sol-gel transition above a critical

degree of saturation [51]. The critical condition for sol-gel transition in this system has been derived by Stockmayer [131] as

$$p_{\mathbf{A}}p_{\mathbf{B}} \geq \frac{1}{(f-1)(g-1)}, \quad (2.1)$$

where the degrees of saturation $p_{\mathbf{A}}$ and $p_{\mathbf{B}}$, estimated by module affinity and concentrations, denote the probability that a given \mathbf{A} or \mathbf{B} module, respectively, is in the bound state. This derivation of the condition for gel formation assumes ideal, Caley-tree binding configurations throughout the system, where each additional bond in any given complex is accompanied by an addition of a monomeric unit to the complex [51, 131]. Thus, intra-complex binding, or cyclization, is prohibited in this analysis. The possibility of cyclization cannot be ignored in biological systems, where the effects of avidity are prevalently observed.

To account for the possibility of cyclic binding in our systems, I define the degrees of cyclization as

$$\sigma_{\mathbf{A}} \equiv \frac{\text{cyclic } \mathbf{A} \text{ modules}}{\text{cyclic } \mathbf{A} \text{ modules} + \text{acyclic } \mathbf{A} \text{ modules}}; \text{ and} \quad (2.2)$$

$$\sigma_{\mathbf{B}} \equiv \frac{\text{cyclic } \mathbf{B} \text{ modules}}{\text{cyclic } \mathbf{B} \text{ modules} + \text{acyclic } \mathbf{B} \text{ modules}}, \quad (2.3)$$

where $\sigma_{\mathbf{A}}$ and $\sigma_{\mathbf{B}}$ are the probability that a bound \mathbf{A} or \mathbf{B} module, respectively, participates in intra-complex binding. Allowing for the possibility of cyclization in our system, I redefine the critical condition in Equation 2.1 as

$$(1 - \sigma_{\mathbf{A}})(1 - \sigma_{\mathbf{B}})p_{\mathbf{A}}p_{\mathbf{B}} \geq \frac{1}{(f-1)(g-1)}, \quad (2.4)$$

where I note that only acyclic (ideal) bonds will contribute in the approach to the sol-gel transition point as $p_{\mathbf{A}}$ and $p_{\mathbf{B}}$ increase. Consider the simplifying case of equal module concentrations. For this case, $p_{\mathbf{A}} = p_{\mathbf{B}} \equiv p$; $\sigma_{\mathbf{A}} = \sigma_{\mathbf{B}} \equiv \sigma$; and Equation 2.4 reduces to

$$[(1 - \sigma)p]^2 \geq \frac{1}{(f-1)(g-1)}. \quad (2.5)$$

Equation 2.5 indicates that a system above the gel point with degree of saturation p can tolerate some level non-ideal behavior in the form of cyclic bonds without being pushed below the sol-gel transition critical point. To determine this maximally tolerable degree of cyclization σ_{tol} , I equate the two sides of Equation 2.5 and solve for σ to obtain

$$\sigma_{\text{tol}} = 1 - \frac{1}{p} \sqrt{\frac{1}{(f-1)(g-1)}}. \quad (2.6)$$

Equation 2.6 indicates that systems at higher degrees of saturation or with higher valencies can tolerate higher degrees of cyclization before being pushed below the sol-gel transition critical point. The SH3₅ + PRM₅ and SH3₅ + PRM₈^{N-WASP} systems at degrees of saturation 81 % (based on module concentration of 8 mM in the droplet phase and $K_d = 356 \mu\text{M}$ for SH3₁ + PRM₁ [91]) and 90 % (based on module concentration of 13.1 mM in the droplet phase and $K_d = 147 \mu\text{M}$ for SH3₅ + PRM₁^{N-WASP} [91]) could tolerate up to 69 % and 79 %, respectively, of all bonds in the system being cyclic and still remain above the sol-gel transition critical point. In the absence of cyclization ($\sigma = 0$), Equation 2.5 predicts the critical degree of saturation p_{critical} of 0.25 and 0.17 for the SH3₅ + PRM₅ and the SH3₅ + PRM₈^{N-WASP} systems, respectively.

2.3 Discussion

Our analysis provides an estimate of the conditions under which the droplets formed by the SH3/PRM model systems contain polymers that have undergone the sol-gel transition. These calculations suggest that a very large fraction of the binding sites would have to be engaged in cyclic interactions for the resulting polymer distribution to not be a gel.

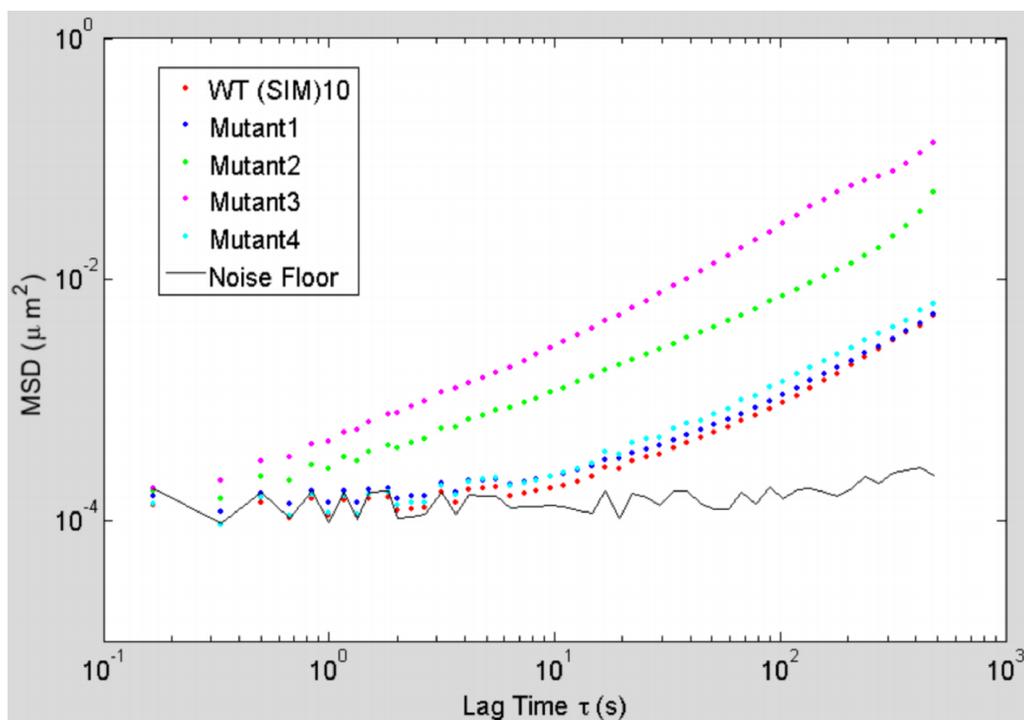
However, this analysis has gross limitations. It assumes that the degree of binding ($p_{\mathbf{A}}$ and $p_{\mathbf{B}}$) can be calculated from the known modular binding affinity of \mathbf{A} for \mathbf{B} and their module concentrations, treating each module as a monovalent molecule. More recently, our work with the SUMO/SIM model system (not done at the time this analysis was published),

suggests that the effective modular affinities in multivalent systems are in fact much stronger than in their monovalent counterparts (Figure 3.4 and Table 3.1). Figure 3.4 suggests that monovalent affinities will likely underestimate the effective modular affinities in multivalent systems, and that the degree of saturation of all modules is likely higher than our analysis. The SUMO/SIM experiments suggest that the degree of binding ought to be higher than what I estimate in our calculations above. This suggests that an even higher fraction total bonds, σ_{tot} (Equation 2.6), would have to be cyclic in order for the droplets to not have undergone a sol-gel transition.

Microrheology may be a way to ascertain whether gels have formed within droplets. In these studies, the Brownian motion of sub-micron particles (e.g. fluorescent beads) is tracked and analyzed for its mean squared displacement (MSD) over a range of timescales (τ). MSD has a power law dependence on τ such that $\text{MSD} \propto \tau^\alpha$, where α is the critical exponent (or the slope of the log-log relationship between MSD and τ) [145]. The behavior of MSD over timescales is related to the rheological properties of the material [145]. Polymers of multivalent WW domains and their multivalent ligands exhibit elastic, gel-like behavior ($\alpha \sim 0$) [104]. Importantly, this behavior did not occur with lower valency components, which display microrheological behaviors consistent with viscous liquids and thus inconsistent with gelation ($\alpha \sim 1$).

In collaboration with the laboratory of Dr. Cliff Brangwynne, Ph.D., (Princeton University, Princeton, NJ) we measured the rheological properties of polySUMO/polySIM droplets (Figure 2.1). Sara Chuang and I observed that while the viscosity of the droplets was ~ 60 -fold that of water (analysis not shown), the droplets did not exhibit elasticity, a behavior inconsistent with gelation. Together, these studies suggest that while gelation is within the range of possible behaviors for multivalent systems, further rheological characterization of additional multivalent biological systems is required to determine the generality of this behavior within cellular bodies.

Figure 2.1: Microrheology of polySUMO/polySIM droplets



Average mean squared displacement (MSD) versus timescale (τ) for 500 nm fluorescent polystyrene beads within droplets made from with $60 \mu\text{M}$ (SUMO)₁₀ + $80 \mu\text{M}$ (SIM)₁₀ (red dots). MSD/ τ plots for droplets made with (SUMO)₁₀ and SUMO-binding site mutants of (SIM)₁₀ are also shown: V2A (Mutant 1, blue dots); V2Y, V4Y (Mutant 2, green dots); I5Y (Mutant 3, magenta dots); E10A (Mutant 4, cyan dots). Black curve depicts the uncertainty in MSD as a function of τ , estimated from analogous analysis of immobilized beads.

Chapter 3

Compositional Control of Phase Separated Cellular Bodies

This chapter contains material adapted from a previously published manuscript [7].

3.1 Rationale

Cellular bodies typically contain 10s-100s of types of molecules [23, 52]. Where characterized in detail, only a small number of these components appears to be essential for the structural integrity of the body [60, 73, 35]. We refer here to such molecules as *scaffolds*. In contrast, the remaining majority of components are dispensable for body assembly, and often reside in the bodies only under certain conditions [57, 38, 67, 22, 96]. These molecules, which we refer to here as *clients*, often contain elements that specifically bind to elements in the scaffolds, often via low valency interacting elements of the same class as those in the scaffolds (e.g. [30, 94, 55]). For example, P bodies assemble in part via scaffolding interactions between RNA binding proteins and RNA but also recruit several RNA binding proteins that are not important for P body assembly [23]. Within cellular bodies, clients diffuse much more rapidly than scaffolds [40, 144], suggesting that client-scaffold interactions are more transient than the interactions among scaffolds.

Compositional regulation is a general property of many cellular bodies and may be crucial to their function. Cellular body compositions change during the phases of the cell cycle or in response to stresses [57, 38, 67, 22, 96]. Despite their importance, the fundamental principles governing cellular body composition have been experimentally difficult to

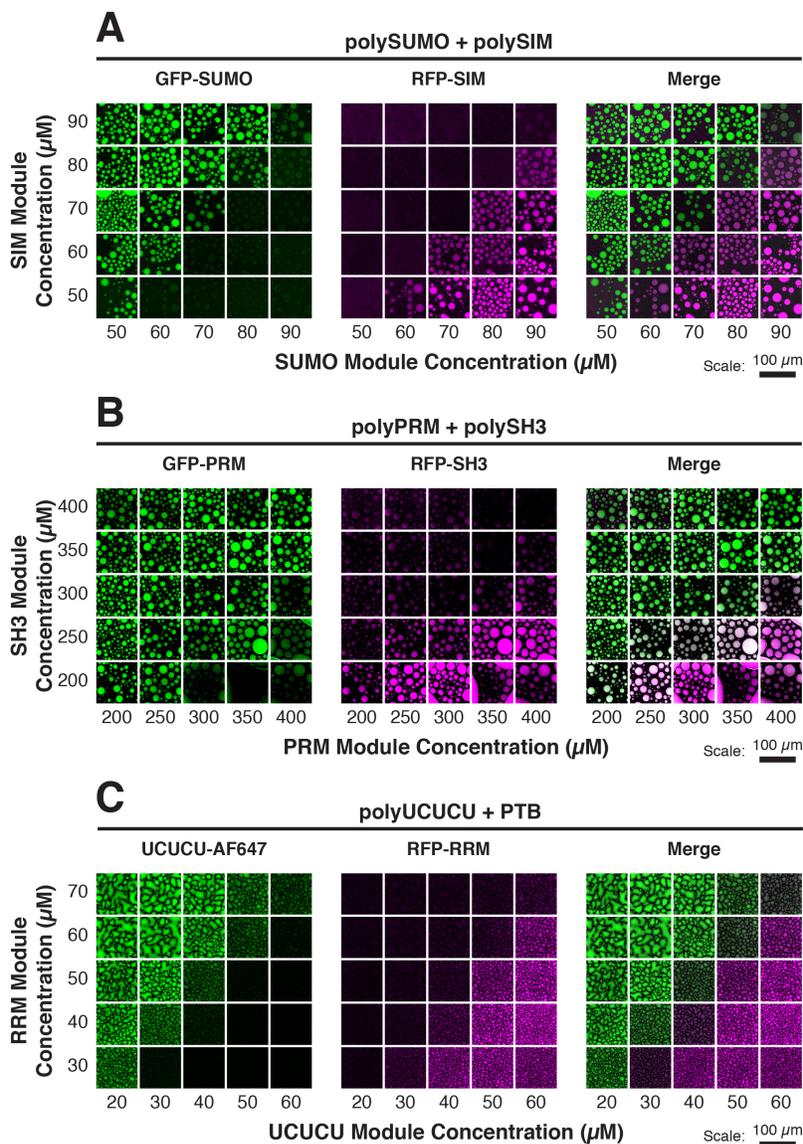
elucidate, owing to the complex nature of both scaffolds and clients and the diversity of species that reside within bodies. However, simplified model systems composed of few types of molecules, each with well-defined interaction elements, can help isolate key molecular parameters and thus have the potential to reveal generalizable concepts.

Here we describe the biochemical and cellular behavior of three different sets of engineered molecules as simplified but representative multivalent scaffolds and low valency clients, which form model cellular bodies. Clients were differentially recruited into the bodies based on the relative stoichiometries of the scaffolds. Changes in client recruitment occurred sharply and on cellular timescales as the scaffold stoichiometries or valencies changed. Client partitioning also depended on client valency. These findings lead to a simple mass action model that predicts many features of the observed client partitioning behavior and suggests how cellular body compositions could be regulated in cells. Behaviors analogous to those of the model systems were observed in PML NBs in mammalian nuclei and P bodies in yeast cytoplasm. Thus, although natural cellular bodies are complex, their compositions may be governed by simple underlying rules and could be altered based on parameters that are easily tunable through cellular and evolutionary mechanisms.

3.2 Scaffold Stoichiometries Dictate Client Recruitment

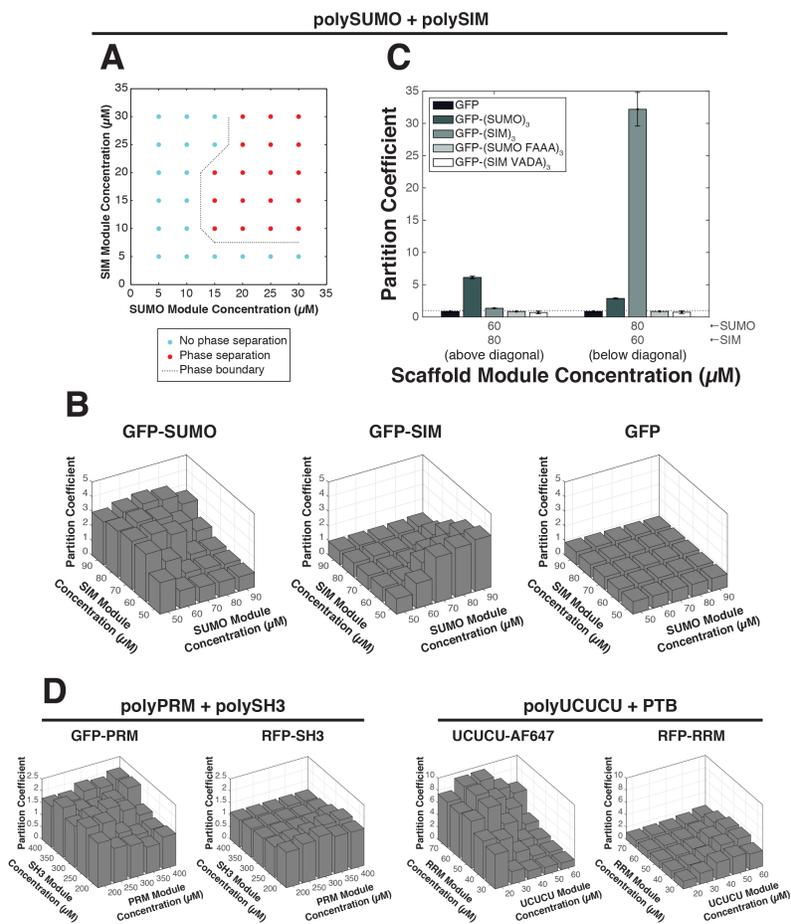
We began by studying three independent pairs of interacting multivalent scaffolds *in vitro*. These systems consisted of (i) a protein with ten repeats of human SUMO3 (polySUMO) and a protein with ten repeats of the SUMO Interaction Motif (SIM) from PIASx (polySIM); (ii) a protein with four repeats of the second SH3 domain from Nck (polySH3), and a protein containing four repeats of a Proline-Rich Motif (PRM) from Abl1 (polyPRM) [91]; and (iii) the PTB protein [contains four RNA recognition motifs (RRMs)], and an RNA with five repeats of the RRM recognition element 5'-UCUCU-3'. (polyUCUCU) [91]. Each of these pairs phase separated when mixed together, but not when individual components were

Figure 3.1: Phase Diagram Position Dictates Client Recruitment



See also Figure 3.2. Solutions of multivalent scaffolds plus the indicated clients were imaged for client fluorescence. AF, Alexa fluorophore. (A) GFP-SUMO (green) and RFP-SIM (magenta) (100 nM each) were mixed with the indicated module concentrations of polySUMO and polySIM. (B) GFP-PRM (green) and RFP-SH3 (magenta) (200 nM each) were mixed with the indicated module concentrations of polyPRM and polySH3. (C) UCUCU-AF647 (green) and RFP-RRM (magenta) (200 nM each) were mixed with the indicated module concentrations of polyUCUCU and PTB.

Figure 3.2: Partitioning into polySUMO/polySIM Droplets Requires SUMO-SIM Binding Interactions



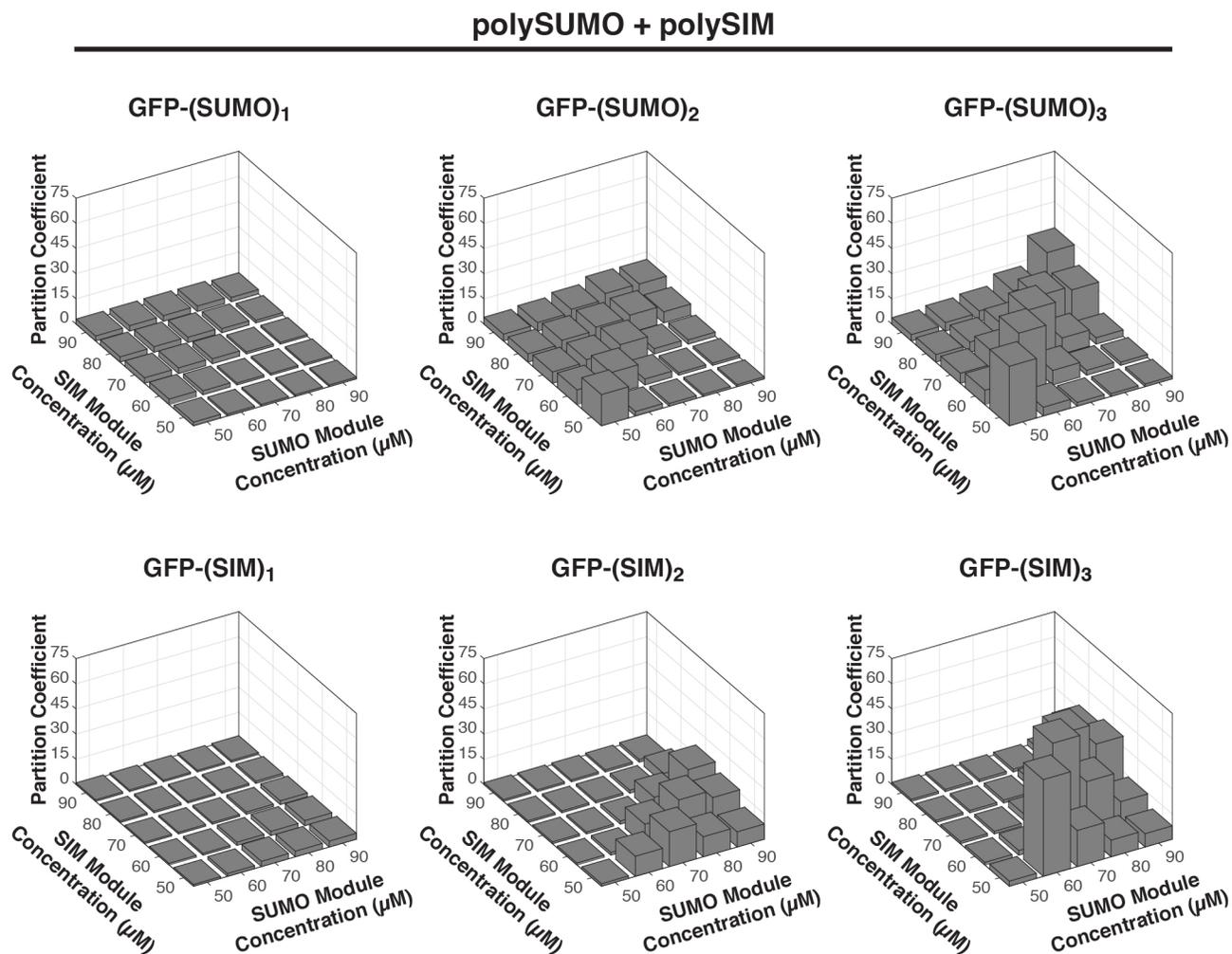
Related to Figure 3.1. (A) Phase diagram for polySUMO and polySIM, indicating where the solution consists of a single phase (red dots) or of two phases (blue dots). Axes show scaffold module concentrations. Dotted line, apparent phase boundary. (B) *PC*s for GFP-SUMO, GFP-SIM, or GFP (100 nM) into droplets formed by the indicated module concentrations of polySUMO and polySIM. (C) *PC*s for the binding site mutant clients GFP-(SUMO FAAA)₃ and GFP-(SIM VADA)₃ (100 nM) into droplets formed by the indicated module concentrations of polySUMO and polySIM. Partitioning data for GFP-(SUMO)₃, GFP-(SIM)₃, and GFP [from Figure 3.3 and panel (B)] at the indicated positions on the phase diagram are reproduced here alongside mutant clients for comparison. Averages of duplicate experiments are shown. Error bars represent SEM. Dotted line, *PC* = 1. (D) (Left) *PC*s for GFP-PRM and RFP-SH3 (200 nM each) into droplets formed by the indicated module concentrations of polyPRM and polySH3. Average of duplicate experiments is shown. (Right) *PC*s for UCUCU-AF647 and RFP-RRM (200 nM each) into droplets formed by the indicated module concentrations of polyUCUCU and PTB.

alone in solution ([91], Figure 3.2A, and data not shown).

To model client recruitment into the bodies, I engineered a series of fluorescently-labeled, mono-valent clients (containing a single element that binds the scaffold), and characterized their partitioning into droplets generated by their cognate scaffolds. I mixed (i) GFP-SUMO and RFP-SIM (or GFP-SIM) with polySUMO/polySIM (Figure 3.1A); (ii) GFP-PRM and RFP-SH3 with polySH3/polyPRM (Figure 3.1B); and (iii) GFP-RRM and UCUCU-AlexaFluor647 with PTB/polyUCUCU (Figure 3.1C). Partition coefficients (PC s) for the clients, defined as the ratio of concentrations in the droplet to the bulk phases, ranged from ~ 1 to 10 across the phase diagram (Figure 3.2B,D). Client recruitment in all three systems was qualitatively similar. Clients partitioned asymmetrically about the diagonal of the phase diagram (the line of equal scaffold stoichiometry) or near to it; each client was enriched only on the side where its cognate scaffold was in stoichiometric excess in the solution. For example, when polySIM was in excess (above the diagonal) GFP-SUMO was enriched in the droplets ($PC \sim 3$), but when polySUMO was in excess (below the diagonal) GFP-SUMO concentrated nearly equally in both phases ($PC \sim 1$) (Figure 3.2B). GFP-SIM showed an opposite pattern of enrichment ($PC \sim 3$ when polySUMO was in excess; $PC \sim 1$ when polySIM was in excess). Recruitment preference transitioned sharply, in switch-like fashion, as the diagonal was crossed. For the polySUMO/polySIM system, neither GFP alone nor clients mutated at their binding sites were enriched in droplets on either side of the diagonal (Figure 3.2B-C). Thus, binding to the scaffold proteins is necessary and sufficient for enrichment into the droplets ($PC > 1$).

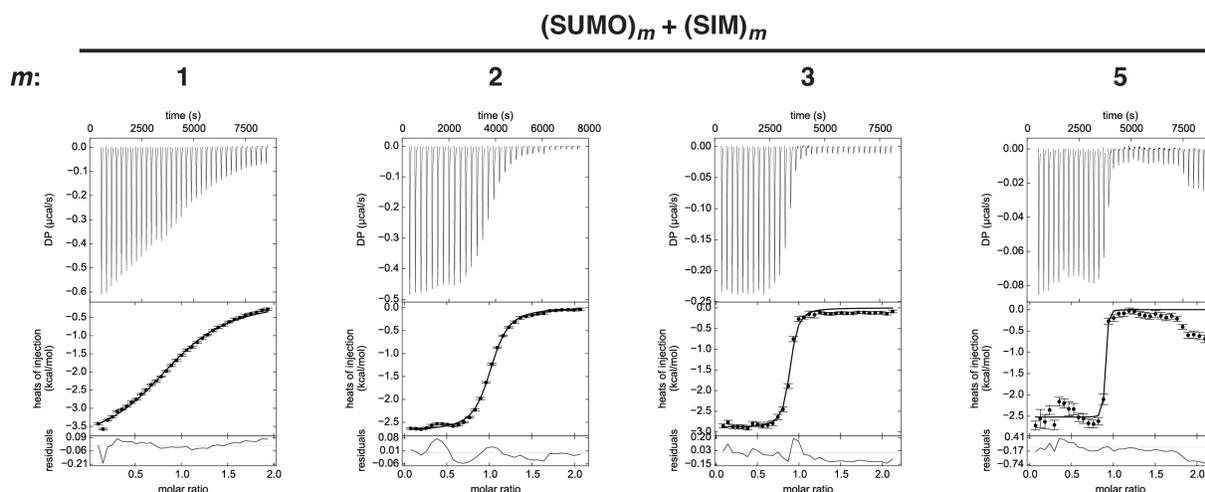
Together, these data show that regardless of the molecular system, low valency clients partition asymmetrically into droplets formed by heterotypic scaffold interactions, with a sharp switch in client recruitment preference across the diagonal.

Figure 3.3: Client Valency Affects Partitioning



See also Figure 3.4 and Table 3.1. *PC*s (means of duplicate samples) of the indicated clients (100 nM) into droplets formed by the indicated module concentrations of polySUMO and polySIM.

Figure 3.4: Apparent Affinity of SUMO to SIM Increases as a Function of Valency



Related to Figure 3.3. Thermograms and integrated heats from a series of ITC experiments with $(\text{SUMO})_m + (\text{SIM})_m$, where $m = 1, 2, 3$, or 5 , as indicated. Integrated heats were fit to a 1:1 heteromeric association model. Residuals are shown in the lowest panels. K_d values obtained from fitting are listed in Table 3.1.

Table 3.1: Apparent SUMO-SIM Affinities as a Function of Valency

Experiment	[Cell] (μM module)	[Syringe] (μM module)	K_d (μM module)
$(\text{SUMO})_1 + (\text{SIM})_1^c$	71.0	640.0	10.0
$(\text{SUMO})_2 + (\text{SIM})_2^c$	63.0	687.4	0.7
$(\text{SUMO})_3^c + (\text{SIM})_3$	28.7	298.6	0.07
$(\text{SUMO})_5^c + (\text{SIM})_5$	12.0	120.0	0.001

Related to Figure 3.3. ITC was used to assess binding affinities with the indicated titration concentrations. ^cProtein selected as titrand (cell) for the experiment.

3.3 Valency of Client Affects Client Recruitment

Since the clients of a given cellular body can differ in their valencies, I examined how client valency affected partitioning. I fused to GFP 2 or 3 tandem repeats of SUMO or SIM and measured the PC for these clients across the polySUMO/polySIM phase diagram (Figure 3.3). Like their monovalent counterparts, the di- and trivalent clients partitioned into the droplets predominantly on one side of the phase diagram, transitioning sharply in their PC s across the diagonal. However, both the di- and trivalent clients had larger magnitudes of maximum partitioning than their monovalent counterparts, a feature that increased with valency: max PC was 19 and 37 for GFP-(SUMO)₂ and GFP-(SUMO)₃, respectively; and 21 and 61 for GFP-(SIM)₂ and GFP-(SIM)₃, respectively. In all cases, maximum partitioning occurred just past the diagonal, substantially enhancing the sharpness of the switch between client preferences. The increased partitioning was likely due to higher apparent affinity of the di- and tri-valent clients for the scaffold. Indeed, isothermal titration calorimetry (ITC) experiments verified that apparent affinity of the clients to cognate sites increases with increasing valency (Figure 3.4 and Table 3.1).

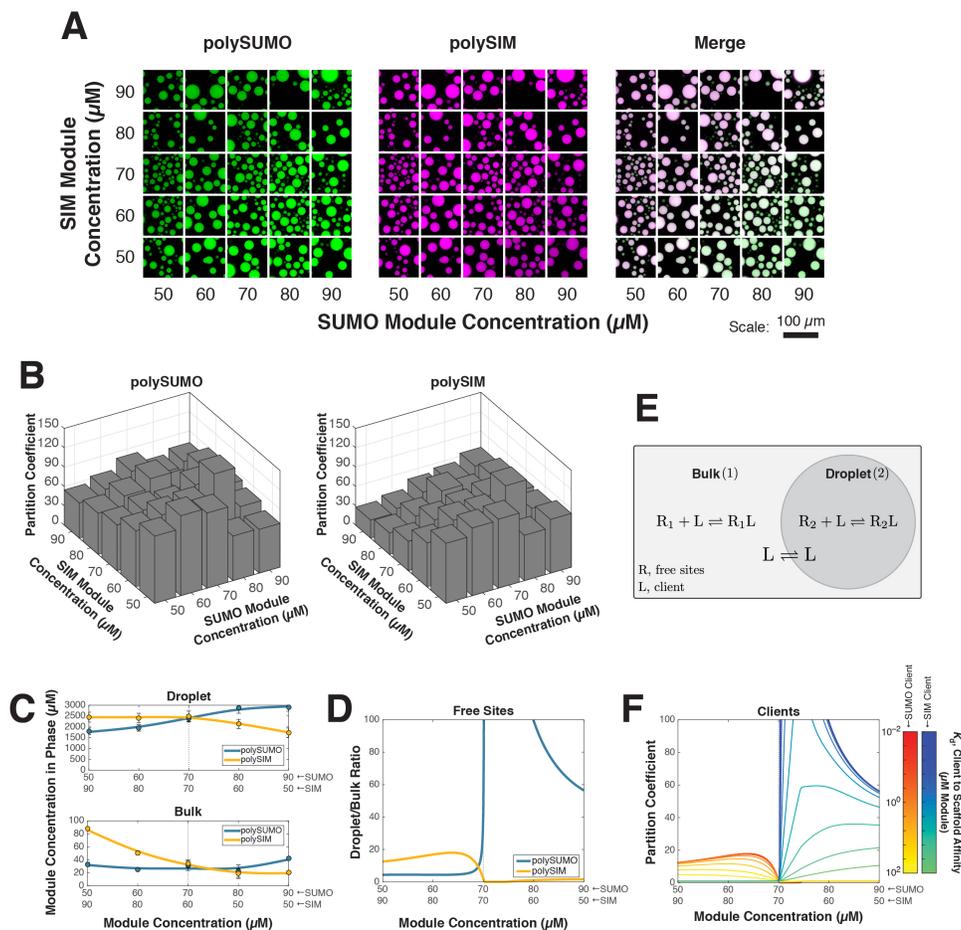
These data demonstrate that in addition to position on the phase diagram, client valency can strongly influence client partitioning and thus droplet composition.

3.4 Mass Action Explains Switch-like Partitioning of Low Valency Clients

We sought to understand the origin of the switch-like nature of client partitioning. Our data suggest that partitioning depends strictly on SUMO-SIM interactions between clients and scaffolds (Figure 3.2B-C). We reasoned that client partitioning should be governed by the relative concentrations of available scaffold binding sites in droplets versus the bulk.

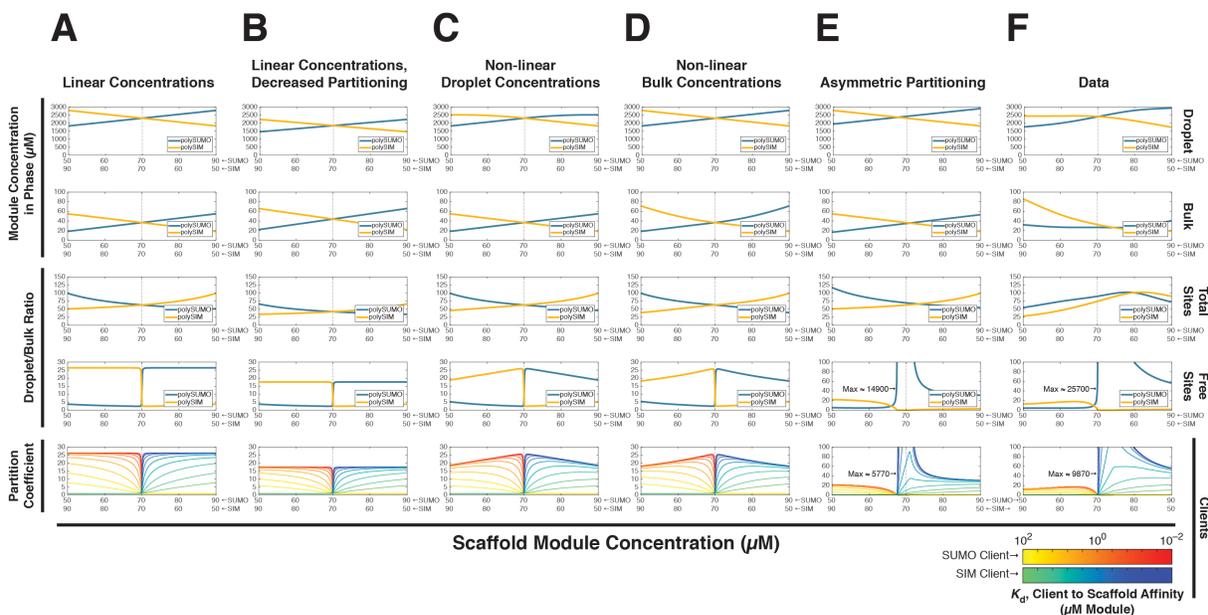
The apparent dissociation constant for polySUMO/polySIM [$K_d \leq 1$ nM, based on ITC measurements with (SUMO)₅/(SIM)₅] is much less than the scaffold concentrations in

Figure 3.5: A Mass Action Model Predicts Client Partitioning Behavior



See also Figures 3.6 and 3.7 and Table 3.2. (A) Imaging of polySUMO and polySIM (1 % labeled with AF488 (green) and AF647 (magenta), respectively) fluorescence. (B) PC s (means of duplicate samples) of polySUMO (left) and polySIM (right) calculated from imaging (see Methods). (C) Scaffold module concentrations (blue and yellow dots) in the droplet (top) and bulk (bottom) phases at the anti-diagonal data points from panel (B). To model client partitioning, values of concentrations were smoothed and interpolated with a cubic spline to yield continuous curves from discrete data. The continuous, interpolated values were used for subsequent calculations. Error bars represent SEM. Dotted line, phase diagram diagonal. (D) Blue curve shows the ratio of free SUMO sites in the droplet phase to free SUMO sites in the bulk phase. Yellow curve shows the analogous ratio for free SIM sites. (E) Mass action model for the partitioning of a low valency client, L, that binds to free scaffold sites R1 and R2 in the droplet and bulk, respectively (see Methods). Predicted PC of clients as a function of affinity for scaffolds. Free site concentrations computed in (E) were used to parameterize the model (C) and predict partitioning of client as a function of their apparent affinity (ranging from 10^{-2} - $10^2 \mu\text{M}$ module) for the scaffolds (see §3.4.1).

Figure 3.6: Subtle Features in Concentration Changes of Scaffolds affect Qualitative Behaviors of Predicted Client Partitioning



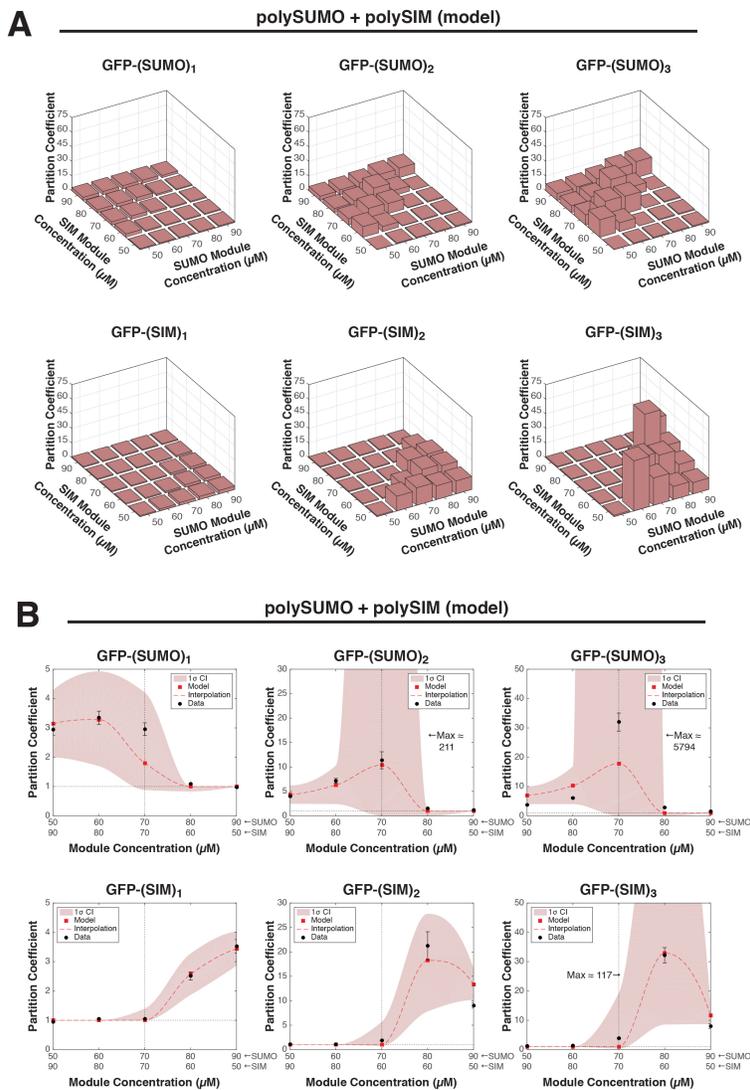
Related to Figure 3.5. Iterative alterations to partitioning data on the anti-diagonal from Figure 3.5B to examine the resulting behavior of predicted client partitioning. All computations were done and are displayed as in Figure 3.5. (A) Scaffold concentrations in each phase were linearized by averaging the data about the diagonal to make them symmetric followed by fitting to a line. (B) Linearized droplet concentrations were reduced by 20 % to decrease partitioning by a constant factor. (C) A quadratic term was added to the linearized droplet concentrations in to curb increase beyond the diagonal. (D) A quadratic term was added to the linearized bulk concentrations in to accelerate increase beyond the diagonal. (E) Concentrations of the two scaffolds were made asymmetric about the diagonal. polySIM concentrations are identical to those in (A), while polySUMO droplet concentrations were increased by 5 % of the mean droplet concentration. Concomitantly, polySUMO bulk concentrations were decreased by 5 % of the mean bulk concentration. (F) Scaffold concentrations obtained from data in Figure 3.5C. Analyses from Figure 3.5C-D,F are reproduced here for comparison.

either phase (Figure 3.5C), suggesting that most scaffold sites are occupied. Moreover, the apparent client-scaffold dissociation constants ($K_d = 70 - 10,000$ nM, estimated from ITC measurements with $(\text{SUMO})_m + (\text{SIM})_m$, for $m = 1, 2$ or 3) (Figure 3.6 and Table 3.1) are much weaker than the apparent polySUMO/polySIM affinity. Thus, clients should be poor competitors of scaffold-scaffold interactions. This analysis suggests that, in either phase, only the scaffold that is in stoichiometric excess will have free sites accessible to its cognate client. Conversely, the scaffold that is stoichiometrically deficient will effectively be saturated by scaffold-scaffold interactions, and invisible to its cognate client in either phase.

The scaffold composition of the droplet and bulk phases varied in a smooth, graded fashion across the phase diagram, with PC s ranging from ~ 30 -125 (Figures 3.5A-B). For most of the phase diagram, the PC of the two scaffolds was similar (within a ~ 2 -fold range), so that the droplets were essentially concentrated counterparts of the bulk solution (Figure 3.5C). Thus, at each point in the diagram, the scaffold in excess has a higher concentration of free sites in the droplet than the bulk (Figure 3.5D), and consequently concentrates its cognate client into the droplets. The scaffold that is stoichiometrically deficient has few free sites in either phase, and its cognate client remains uniformly distributed. Since the stoichiometric relationship between the two scaffolds switches abruptly in both phases near the diagonal, the capacity of the droplets to recruit one client over the other also switches abruptly.

I modeled client partitioning by mass action (Figure 3.5E). I allowed clients to equilibrate between two simulated phases while binding to free sites at concentrations computed from our experiments (Figure 3.5C, Table 3.1 and §3.4.1). This simple mass action model suffices to recapitulate the key qualitative features of observed client partitioning (Figure 3.5F): (i) selective partitioning of clients, restricted to only one side of the diagonal; (ii) a sharp change in partitioning as the diagonal is crossed; and (iii) the dependence of partitioning on the apparent client-scaffold affinity. As described in §3.4.2, the model also predicts less

Figure 3.7: Mass Action Model Qualitatively Recapitulates Partitioning Behavior of Low Valency Clients but has Quantitative Limitations



Related to Figure 3.5. (A) Client partitioning data with the polySUMO/polySIM system (Figure ?? were fit to the free sites mass action model (Figure 3.5; see Methods), and the best-fit client partitioning predictions are shown. Compare to Figure 3.3. (B) The antidiagonal of partitioning data from Figure 3.5 (black dots) and model predictions from panel (A) (red squares) are shown alongside. To aid visualization, discrete points in the model were interpolated with a piecewise cubic spline (red dashed curve). Error bars represent SEM of the data, while the pale red region represents the 1σ confidence interval of the model. Note large confidence intervals in model near the diagonal, suggesting that the free sites mass action model is not robust in this region of the phase diagram. This precludes precise quantitative prediction in this region of the phase diagram. Vertical black dotted line, the diagonal; horizontal black dotted line, $PC = 1$.

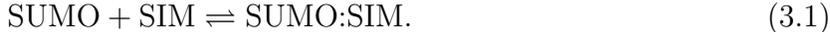
intuitive features of the data, including non-monotonic partitioning as well as dramatically high partitioning of one client and attenuated partitioning of the other near the diagonal (Figures 3.5F; 3.3, trivalent clients; 3.6 and 3.7).

Collectively, this analysis suggests that switch-like changes in client partitioning fundamentally arise from the sharp inversion of scaffold excess across the diagonal of the phase diagram.

The following sections describe the mass action model in more detail.

3.4.1 Overview of the Mass Action Model

I model partitioning of the clients as an equilibrium mass action phenomenon. I assume, as has been done previously for polymers of multivalent monomers [51, 131], that each module in the polymeric scaffold or client acts independently of every other module. With this assumption, the interactions between the multivalent scaffold proteins reduce to the binding of their modules:



I further assume that this condition is maintained after phase separation of the solution into droplet and bulk phases, with each phase undergoing the scaffold binding reaction independently. With these assumptions, I first solve for $[\text{SUMO}]$ and $[\text{SIM}]$, the unbound module concentrations (that is, the free sites concentrations) of each scaffold, in the system of binding and mass conservation equations,

$$K_{d,\text{pol}} = \frac{[\text{SUMO}][\text{SIM}]}{[\text{SUMO:SIM}]} \tag{3.2}$$

$$[\text{SUMO}]_{\text{tot}} = [\text{SUMO}] + [\text{SUMO:SIM}] \tag{3.3}$$

$$[\text{SIM}]_{\text{tot}} = [\text{SIM}] + [\text{SUMO:SIM}]. \tag{3.4}$$

Two analogous systems of Equations 3.2-3.4 are solved separately for the two phases. Here, $K_{d,pol}$ is the apparent dissociation constant (in units of module concentration) for the reaction between SUMO and SIM sites in the scaffold, and $[SUMO]_{tot}$ and $[SIM]_{tot}$ are the measured or simulated total module concentrations of polySUMO or polySIM in the phase in question. I invariably set $K_{d,pol}$ throughout this study to the value of the apparent dissociation constant for the binding reaction of $(SUMO)_5$ to $(SIM)_5$ (Figure 3.4 and Table 3.1). Because experimental module concentrations of the scaffolds are much, much larger than $K_{d,pol}$, the uncertainty in the exact value of $K_{d,pol}$ is of little consequence in subsequent calculations (not shown).

Next, to calculate partitioning of clients between the two phases, I assume that the binding of the client to the scaffold does not appreciably affect the binding equilibrium between the two scaffolds. Therefore, the client only interacts with the free SUMO and SIM sites that remain in each phase after the binding of polySUMO to polySIM has equilibrated (Equations 3.1-3.4). With this assumption, the partitioning of the client between the two phases is analogous to the binding of a ligand L to two receptors R_1 and R_2 , each in one of two compartments separated by a boundary that is permeable to L but not to R_1 or R_2 . The binding of L to the receptors R_1 and R_2 is governed by



The two compartments (1 and 2) are related by their free ligand concentrations, equal in both phases, which I define as

$$[L] \equiv [L]_1 = [L]_2. \quad (3.7)$$

I compute the concentrations of free and bound L in each compartment by solving

the system of mass conservation and coupled binding equations,

$$K_{d,1} = \frac{[R_1][L]}{[R_1L]} \quad (3.8)$$

$$K_{d,2} = \frac{[R_2][L]}{[R_2L]} \quad (3.9)$$

$$[R_1]_{\text{tot}} = [R_1] + [R_1L] \quad (3.10)$$

$$[R_2]_{\text{tot}} = [R_2] + [R_2L] \quad (3.11)$$

$$[L]_{\text{tot,system}} = \phi([L] + [R_1L]) + (1 - \phi)([L] + [R_2L]), \quad (3.12)$$

where the two K_d values are the apparent affinities of the client to the free sites in the scaffold in the two respective compartments. Initial versions of the model using only a single K_d value for client-scaffold affinities in both phases showed significantly larger deviations from the experimental data. I justify use of different K_d values by the fact that the scaffold-scaffold assemblies are likely quite different in size and molecular organization in the two phases. ϕ is defined as the volume fraction of compartment 1, and, assuming incompressibility of the system, relates to the volume fraction of compartment 2 by

$$\phi \equiv \phi_1 = 1 - \phi_2. \quad (3.13)$$

Solving this system of equations yields the concentrations of all species, bound and free, in each compartment, from which the partition coefficient PC , the ratio of the total concentrations of L in the two compartments, can be computed as

$$PC = \frac{[L] + [R_2L]}{[L] + [R_1L]}. \quad (3.14)$$

Because obtaining an exact analytical solution to this system of equations is intractable (not shown), I solved this system of equations using numerical methods in MATLAB and used the solution to obtain predicted values of PC given $[R_1]_{\text{tot}}$; $[R_2]_{\text{tot}}$; $K_{d,1}$; $K_{d,2}$; $[L]_{\text{tot,system}}$; and ϕ . I assigned compartment 1 as the bulk phase and 2 as the droplet phase,

and assigned $[R_1]_{\text{tot}}$ and $[R_2]_{\text{tot}}$ to correspond to free sites concentrations in the respective phase (Equations 4-6). Free sites of polySUMO or polySIM were used to calculate partitioning of SIM or SUMO clients, respectively. I assigned $[L]_{\text{tot,system}}$ to correspond to the total module concentration of client added to solution. Values for ϕ were obtained using mass and volume conservation (assuming incompressibility of the solution) from measured or simulated values of total polySUMO and polySIM module concentrations in the droplet and bulk phases, as well as the total module concentrations of the scaffolds added to the solution, with either of the two relations,

$$[\text{SUMO}]_{\text{tot,solution}} = \phi[\text{SUMO}]_{\text{tot,bulk}} + (1 - \phi)[\text{SUMO}]_{\text{tot,droplet}} \quad (3.15)$$

$$[\text{SIM}]_{\text{tot,solution}} = \phi[\text{SIM}]_{\text{tot,bulk}} + (1 - \phi)[\text{SIM}]_{\text{tot,droplet}} \quad (3.16)$$

The two values of ϕ computed by solving each of these two equations separately for ϕ were averaged and the resulting value was used in the procedure to compute PC (Equation 3.14).

3.4.2 Behavior of the Model

Using the model I sought to understand which features of the scaffold partitioning give rise to the qualitative behaviors in client partitioning observed experimentally (Figure 3.3)—(i) the selective partitioning of one client on each side of the phase diagram diagonal; (ii) the sharp switch in client partitioning across the diagonal; (iii) the increase in client partitioning with increasing client affinity; (iv) the fall off in client partitioning away from the diagonal (non-monotonic partitioning); and (v) the explosive partitioning of high affinity clients near the diagonal. To examine these features, I iteratively simulated subtle perturbations to scaffold concentrations across the phase diagram and mapped their effects on predicted client partitioning outcomes.

I started with the simplest case of constant concentrations and equal partitioning of both scaffold proteins across the phase diagram. Equal scaffold partition coefficients did not

give rise to changes in client partitioning across the phase diagram, contrary to observed behavior, because free site ratios (the ratio of free scaffold sites in the droplet and bulk phases—the driving force for client partitioning into the droplets) did not change across the phase diagram in this case (not shown). Next, I evaluated the effects of changes in scaffold concentrations across the phase diagram that were linear and symmetrically related across the diagonal (Figure 3.6A, rows 1-2). Linear, symmetric changes produced free site ratios that were > 1 for only one scaffold on each side of the diagonal. The free site ratios were relatively flat on either side of the diagonal, but sharply changed across the diagonal from a value on one side close to that of the scaffold partition coefficient to ~ 1 on the other side (Figure 3.6A, row 4). In turn, high affinity client partitioning tracked closely with the free sites ratios, with an analogous sharp switch across the diagonal (Figure 3.6A, row 5, high affinity client). Client partitioning on the favored side progressively decreased with decreasing client-scaffold affinity, consistent with our experimental observations. Thus, linear, symmetric scaffold partitioning was sufficient to produce features (i)-(iii) above.

Decreasing droplet concentrations of the scaffold across the phase diagram by a scalar factor (decreased partitioning) (Figure 3.6B, rows 1-3) caused a decrease in the corresponding free sites ratio, and likewise the client partitioning (Figure 3.6B, rows 4-5). But client partitioning remained monotonic, and did not fall off away from the diagonal. Thus, linear, symmetric changes in scaffold concentrations are sufficient to produce the sharp switch in client partitioning across the diagonal, but not the decrease in client partitioning away from the diagonal, as observed experimentally for the divalent and trivalent clients (Figure 3.3). However, when the concentrations of scaffold in the droplet (Figure 3.6C) or bulk (Figure 3.6D) changed non-linearly, such that the scaffold partition coefficients fell off toward the corners of the phase diagram, as observed experimentally in Figure 3.5B, client partitioning became non-monotonic, rising steeply at the diagonal and then falling off toward the corners. This occurred because non-linear changes in scaffold concentrations curbed the rate

of increase of the cognate scaffold compared to the non-cognate scaffold, thereby decreasing the free sites ratio after the initial increase. Thus feature (iv) above can occur, at least in part, because of the non-linear changes in scaffold concentrations across the phase diagram.

When the two scaffolds partitioned asymmetrically across the phase diagram (even with linear changes in concentrations) (Figure 3.6E), two features emerged that were also observed in our experiments. First, the clients now partitioned asymmetrically; that is, one maximally partitioned much more strongly than the other. Second, the more strongly partitioning client now could achieve very large partition coefficients near the diagonal (Figure 3.6E, row 5). This effect did not result from a large increase in client concentration in the droplet, but rather from a strong depletion from the bulk (not shown). These behaviors arose because the stoichiometric equality points of the two scaffolds in the droplet and bulk phases are displaced from one another across the diagonal. When the scaffolds are (nearly) equal in the bulk, but unequal in the droplet, the driving force for one client to enter the droplet will be very strong. This behavior of multivalent scaffold-client systems could enable bodies to function by sequestering certain clients *in vivo*. That is, even though the total volume of a particular type of body is only a small fraction of total cellular volume, in some scaffold concentration regimes, client partitioning could be extremely high, enabling most of the client molecules in the cell to localize to the body. In the absence of such high partitioning, the small volume fraction of a body would argue against sequestration as a likely function of the body.

Analogous analyses of our experimental data revealed that the observed client partitioning was likely affected by all of these effects simultaneously (Figure 3.6F). I conclude, therefore, that observed partitioning behaviors derive fundamentally from the precise way that polySUMO and polySIM partition between the droplet and bulk phases, which in turn dictates the free sites ratios that govern client partitioning.

Table 3.2: **Values of Free Parameters (Client-Scaffold Affinities in Bulk and Droplet) from Fitting of Partitioning Data to Mass Action Model**

GFP-tagged Client	$K_{d,1}$ (μM module)	$K_{d,2}$ (μM module)	RMSD
(SUMO) ₁	42 ± 17	104 ± 19	0.49
(SUMO) ₂	0.2 ± 1.7	0.6 ± 4.6	4.62
(SUMO) ₃	0.005 ± 1	0.01 ± 2	11.83
(SIM) ₁	$2\text{E}5 \pm 1\text{E}8$	476 ± 36	0.19
(SIM) ₂	13 ± 3	33 ± 3	2.21
(SIM) ₃	1.6 ± 0.3	7.0 ± 0.6	6.77

Related to Figure 3.5. $K_{d,1}$, bulk; $K_{d,2}$, droplet.

3.4.3 Fitting to Data

Partitioning data for each SUMO or SIM client (Figure 3.3) were fit to the 2 compartment mass action model using non-linear least squares methods in MATLAB. I used the polySUMO and polySIM free sites concentrations, computed from total module concentrations of polySUMO and polySIM in each of the two phases (Equations 3.2-3.4), to obtain values for $[R_1]_{\text{tot}}$ and $[R_2]_{\text{tot}}$. I used the experimentally determined module concentrations of the clients in solution to obtain values for $[L]_{\text{tot,system}}$. Using $[R_1]_{\text{tot}}$, $[R_2]_{\text{tot}}$, and $[L]_{\text{tot,system}}$ as fixed parameters, I fit the data allowing $K_{d,1}$ and $K_{d,2}$ to float as free parameters. With two free parameters, the model qualitatively recapitulated many features of the observed partitioning data (compare Figures 3.7A and 3.3), including a sharp transition in client partitioning across the diagonal as well as non-monotonic partitioning beyond the diagonal.

3.4.4 Quantitative Limitations of the Model

Despite the qualitative correspondence with the data, the model did not predict the partitioning quantitatively. Calculations of partitioning using our calorimetry-based estimates of the client-scaffold affinities (Figure 3.4 and Table 3.1) to parameterize the model overestimate client partitioning relative to experiment (compare Figures 3.5F and 3.3). This suggests that the client-scaffold affinities might be weaker than the values measured in our

ITC experiments. Because our ITC experiments were conducted with low valency molecules binding each other, these measurements might not faithfully represent the potentially hindered access to free sites within a crosslinked network or oligomer. As described, I fit our data to the mass action model by allowing the client-scaffold affinities in both the droplet and bulk phases to float as free parameters (Table 3.2). The resulting predictions better matched the magnitude of partitioning observed experimentally (Figure 3.7A). Indeed, the client-scaffold affinities obtained from the fits were weaker than those measured by ITC (compare Tables 3.1 and 3.2), suggesting that free sites in both the droplet and bulk phases are less accessible than free sites on free molecules. Consistent with the notion that scaffolds in the droplets are crosslinked to a higher degree than molecules in the bulk, the fitted values of client-scaffold affinities were generally weaker in the droplets than in the bulk ($K_{d,2} > K_{d,1}$) (Table 3.2). However, the model remained a poor quantitative predictor of partitioning, more so for the higher valency clients (see RMSDs in Table 3.2).

Because the fixed parameters in the model are subject to uncertainties in their measurements, I also wondered how the experimental error in these measurements propagates to the prediction of client partitioning, and whether the discrepancies between model and data were statistically significant. To estimate the confidence interval in our model, I simulated 10^4 predictions for partitioning, each with slightly varying concentrations of polySUMO and polySIM in the two phases. These values were randomly drawn from normal distributions with means set to measured values and standard deviations set to the corresponding standard errors for those values. I define the 1σ confidence interval of the model as the range encompassing 68.3 % of all 10^4 randomly generated curves about their median (15.9 to 84.1 percentile). I compared the data to the model along with its confidence intervals, shown for the antidiagonal in Figure 3.7B. This analysis revealed that despite reasonably precise measurements of scaffold concentrations in the droplet and bulk phases (average SEM of all measurements was 11.2 %), the model had very large confidence intervals near the diagonal.

This arises from the fact that near the diagonal the free sites in both the droplet and bulk phases represent, to first order, a small difference between two much larger numbers. This behavior precludes precise quantitative predictions near the diagonal, especially for higher affinity clients (Figure 3.7B and Table 3.2). Though I may be able to reasonably predict qualitative features of client partitioning by systematically analyzing the behavior of scaffold concentrations across the phase diagram (Figure 3.6), a fundamental limitation of the model is the high degree of sensitivity near the diagonal to small changes in input parameters, limiting its predictive power in that region of the phase diagram. Because experimental measurements of concentrations (used as fixed parameters in the model) are always subject to uncertainties, precise predictions of client partitioning near the diagonal may remain infeasible.

3.4.5 Qualitative Limitations of the Model

The partitioning of polySUMO and polySIM is quite different from that of the clients, changing smoothly from one corner of the phase diagram to the other (Figure 3.5A-B). I reason that the partitioning behavior of the clients should approach that of the scaffolds as the valency of the clients increases; in the limit of equal valency, clients and scaffolds are chemically indistinguishable. The transition in behavior would likely occur because the initial assumption in our model, that clients are unable to compete with scaffolds for binding sites, breaks down as client valency (and therefore the apparent client-scaffold affinity) increases. It may be more appropriate, therefore, to consider the clients in this regime as additional polymerizing agents (scaffolds) that drive phase separation rather than species recruited secondarily. This exposes a limitation in predictive ability of our model for the partitioning for such high valency molecules. The partitioning of such molecules should be governed primarily by their propensity for polymerization and phase separation.

Table 3.3: Values of Rate Constants from Single-Exponential Fits of FRAP Experiments

Fluorophore	k (min^{-1})	τ (min)	τ_1 ($\text{s}/\mu\text{m}^2$)
polySUMO-AF488	0.0263 ± 0.003	38	5.7
RFP-SIM	0.8 ± 0.1	1.3	0.2

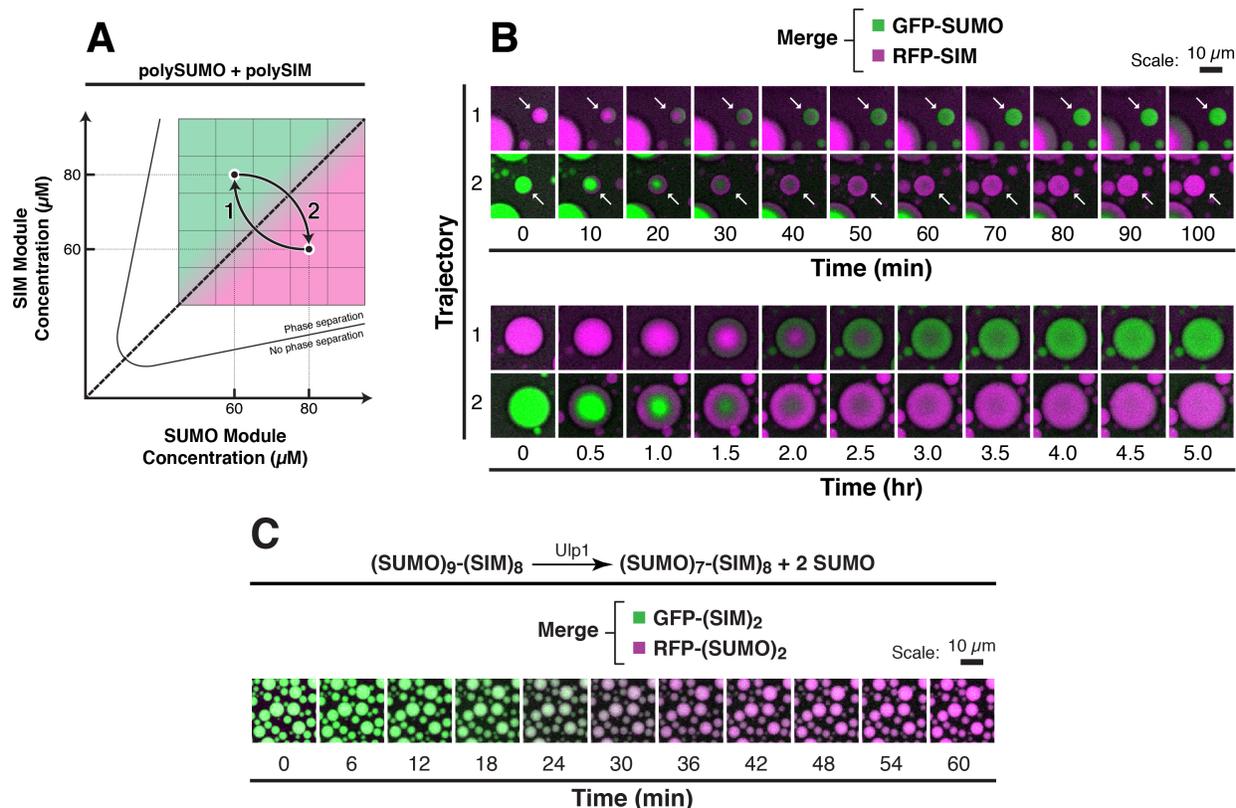
Related to Figure 3.8. k , exponential recovery constant; τ , characteristic timescale ($1/k$); τ_1 , area-normalized timescale [$(\tau/400 \mu\text{m}^2)$ for $\sim 20 \mu\text{m}$ diameter droplets; note change in units between τ and τ_1].

3.5 Compositional States Interchange on Cellular Timescales

Our data and analyses suggest how compositional states could be controlled by mass action. We wondered whether transitions between two compositional states were kinetically feasible on cellular timescales. I equilibrated polySUMO/polySIM droplets at a point on the phase diagram where only one of the clients, either GFP-SUMO or RFP-SIM, was preferentially enriched in the droplets (but both were present in solution). I then abruptly changed the concentration of the scaffold components to move the system to a point across the diagonal where the reciprocal recruitment preference was expected (Figure 3.8A). The droplets remained intact, spherical, and of relatively consistent sizes throughout the experiment. Within ~ 6 hours all droplets expelled the initially enriched client in exchange for the other client (Figure 3.8B). Recruitment of the latter started at the outer edges of the droplets and moved inward, and smaller droplets exchanged clients more rapidly than larger droplets.

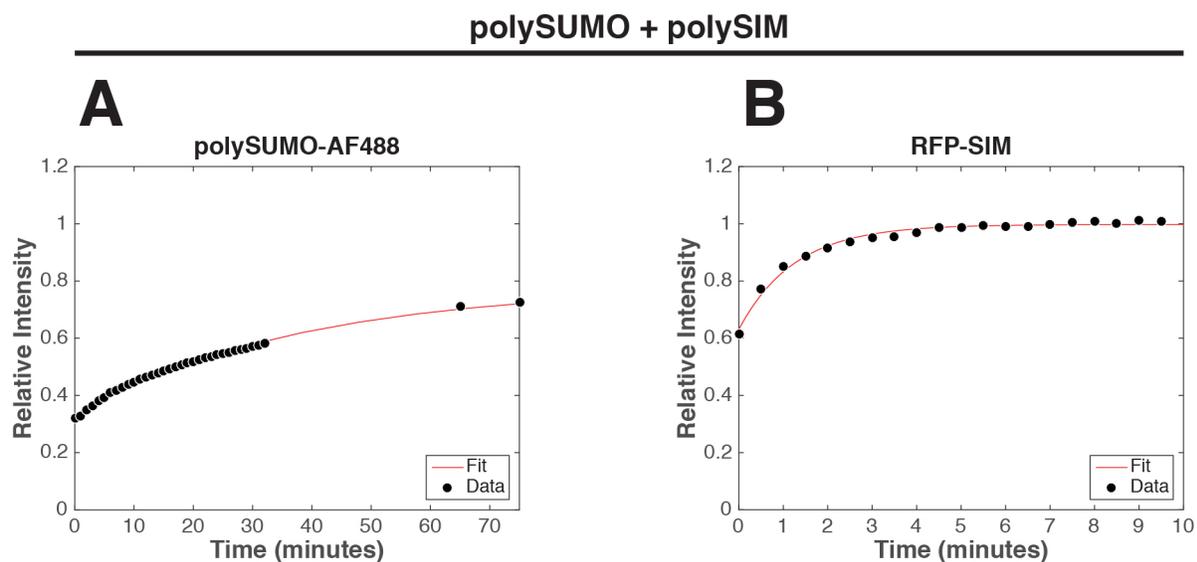
In fluorescence recovery after photobleaching (FRAP) experiments, clients diffused much more rapidly within droplets than did scaffolds (i.e. for droplets $\sim 20 \mu\text{m}$ in diameter, RFP-SIM and polySUMO had exponential recovery constants, τ , of 1.3 min and 38 min, respectively; Figure 3.9 and Table 3.3). Thus, scaffold rearrangements likely limit the rate of transitions between compositional states. Scaling recovery times to droplets of 1 μm diameter, as often observed in cells, indicates that compositions should interchange on a timescale of

Figure 3.8: Droplets Interchange Composition On Cellular Timescales Without Compromising Structural Integrity



See also Figure 3.9 and Table 3.3. (A) Schematic of experiment. After equilibration of 100 nM GFP-SUMO and 100 nM RFP-SIM with polySUMO and polySIM at module concentrations of 60 μM and 80 μM , respectively, concentrations of the polySUMO and polySIM were abruptly shifted to 80 μM and 60 μM , respectively, for trajectory 1 and vice versa for trajectory 2. (B) Time lapse imaging of droplets starting immediately after the abrupt change in concentrations of polySUMO and polySIM, showing merged, pseudocolored fluorescence signals from GFP-SUMO (green) and RFP-SIM (magenta). Note that small droplets (white arrows, top panel) interconvert more quickly than larger droplets (bottom panel). (C) 6 μM of a $(\text{SUMO})_9\text{-(SIM)}_8$ scaffold containing Ulp1 cleavage sites after only the two N-terminal SUMOs was equilibrated with 50 nM of GFP-(SIM)₂ (green) and RFP-(SUMO)₂ (magenta). Time lapse imaging was started immediately after addition of 10 nM of Ulp1. Pseudocolored images showing merged fluorescent signals from the two clients are shown.

Figure 3.9: Rearrangement of Scaffold Polymers Limit the Rate of Compositional Interchange in Droplets



Related to Figure 3.8. 100 nM of RFP-SIM was mixed with 80 μ M module of polySUMO (1 % labeled with AF488) and 60 μ M module of unlabeled polySIM. (A) Intensity recovery after photobleaching of polySUMO-AF488 fluorescence across the entire droplet (black dots; normalized to pre-bleaching intensities) is shown for a representative droplet, measuring roughly 20 μ M in diameter. Single exponential fits of the recovery profiles are also shown (red curve). Recovery rate constants are listed in Table 3.3. (B) An analogous representative fluorescence recovery curve for RFP-SIM is shown. Recovery rate constants are listed in Table 3.3.

~ 6 seconds in natural systems (Table 3.3).

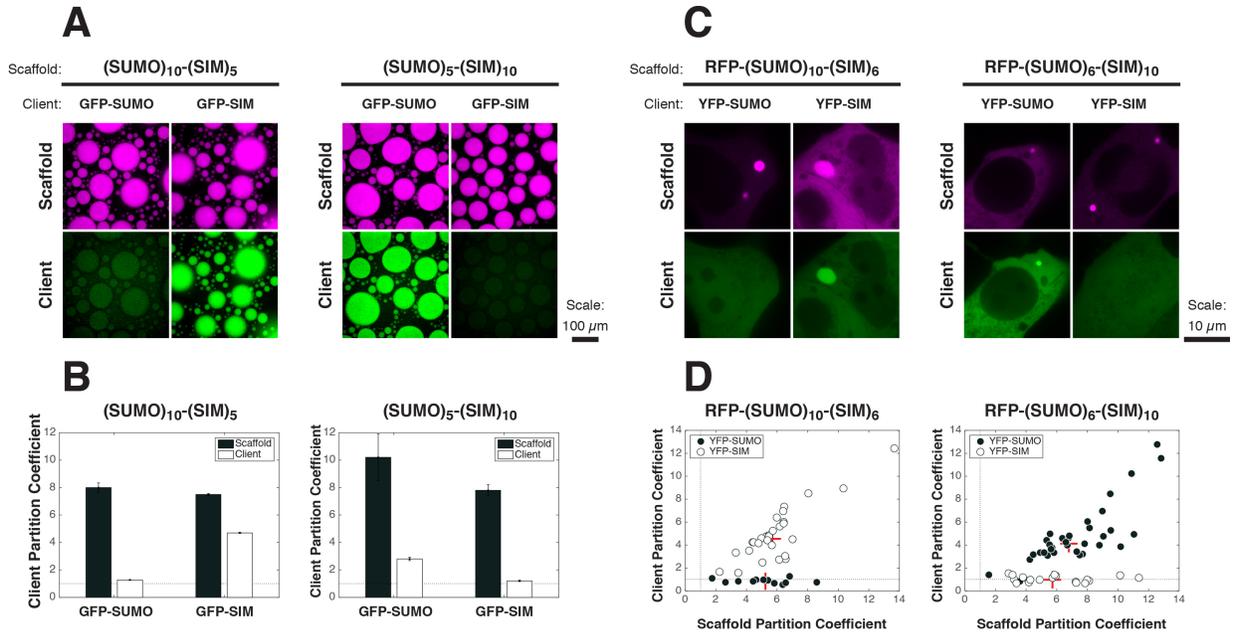
We previously demonstrated how covalent modifications of scaffolds could regulate the formation and dissolution of droplet phases [91]. We likewise wondered whether covalent modifications could also regulate droplet compositions. In cells, SUMO modifications are dynamically added by the SUMO ligase cascade and removed by SUMO proteases. Ben Peeples generated a single component, fused (SUMO)₉-(SIM)₈ scaffold that could be selectively cleaved by Ulp1, the yeast SUMO protease, to produce (SUMO)₇-(SIM)₈, mimicking natural deSUMOylation (see §A.2). Such (SUMO)_m-(SIM)_n ($m \neq n$) fusions are essentially fixed on one side of the phase diagram diagonal, and have client recruitment preferences analogous to the *in trans* systems (see Figure 3.10A). When mixed with GFP-(SIM)₂ and RFP-(SUMO)₂, (SUMO)₉-(SIM)₈ droplets recruited the former but not the latter client (Figure 3.8C). Ulp1 cleavage, which shifted the scaffold to the other side of the phase diagram, caused the droplets to expel GFP-(SIM)₂ and recruit RFP-(SUMO)₂. These data suggest that enzymatic modifications of cellular body scaffolds, such as SUMOylation and deSUMOylation, could robustly regulate body composition.

We conclude that droplets can transition, without compromising structural integrity, between substantially different compositional states on timescales accessible to cells. This can occur with only subtle changes in the concentration or covalent modifications of their polymer scaffolds.

3.6 Engineered Cellular Puncta Selectively Concentrate Low Valency Clients

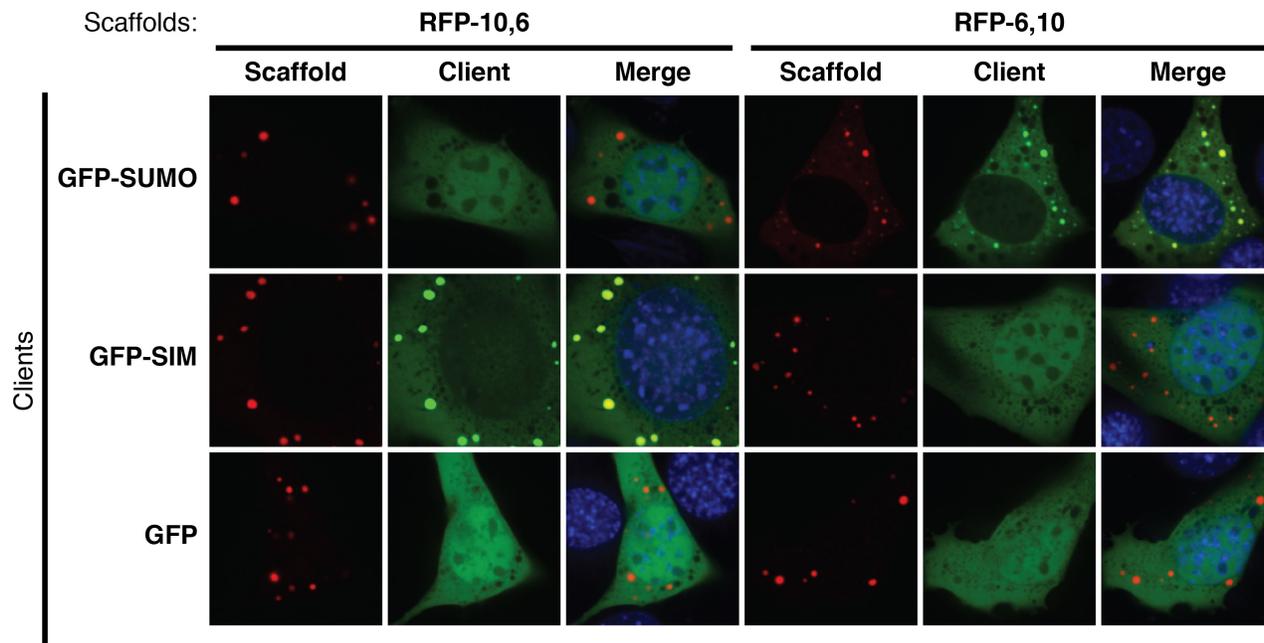
We next asked whether the partitioning behavior observed *in vitro* could also be observed in cells. For these experiments we used *in cis* [(SUMO)_m-(SIM)_n] scaffolds, which afforded tight experimental control of the relative module concentrations independent of absolute concentrations. Both (SUMO)₁₀-(SIM)₅ and (SUMO)₅-(SIM)₁₀ phase separated *in*

Figure 3.10: Cellular PolySUMO-PolySIM Puncta Selectively Recruit Low Valency Clients



See also Figure 3.12 and Figure 3.11. (A) 60 nM of GFP-SUMO or GFP-SIM (green) was mixed with 12 μM of $(\text{SUMO})_{10}-(\text{SIM})_5$ (left) or $(\text{SUMO})_5-(\text{SIM})_{10}$ (right) (1 % RFP-tagged; magenta), and the resulting droplets were imaged for scaffold and client fluorescence. (B) *PCs* for both scaffold (black bars) and clients (white bars) from experiment in (A). Graphs show averages from triplicate experiments. Error bars represent SEM. Dotted line, $PC = 1$. (C) Live cell fluorescence images of YFP-SUMO or YFP-SIM (green) co-transfected with RFP-(SUMO) $_{10}$ -(SIM) $_6$ (left) or RFP-(SUMO) $_6$ -(SIM) $_{10}$ (right) (magenta) into HeLa cells. (D) *PCs* of scaffolds and client components calculated from cells in the experiment. Each symbol represents the average *PC* into all puncta (typically 1-3) in a given cell (12-35 cells per sample) when the indicated scaffold was co-transfected with YFP-SUMO (black circles) or YFP-SIM (white circles). Dotted line, $PC = 1$. Red + sign, median *PC*.

Figure 3.11: Cellular Poly-SUMO-PolySIM Puncta Selectively Recruit Low Valency Clients in 3T3 Cells

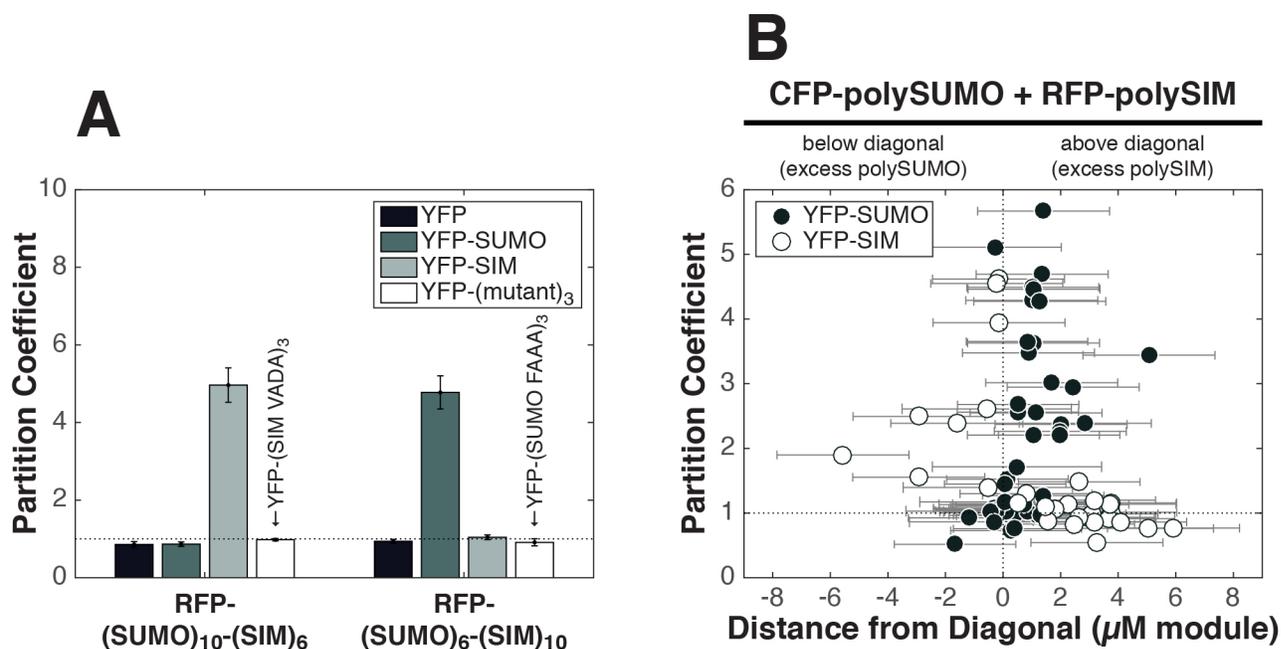


Related to Figure 3.10. Representative images of RFP-(SUMO)₁₀-(SIM)₆ (left) and RFP-(SUMO)₆-(SIM)₁₀ (right) co-transfected in 3T3 cells with GFP-SUMO, GFP-SIM, or GFP. SUMO/SIM clients were recruited in a selective manner.

in vitro at micromolar concentrations. (SUMO)₁₀-(SIM)₅ droplets enriched GFP-SIM ($PC = 4.7$), but not GFP-SUMO ($PC = 1.3$), and (SUMO)₅-(SIM)₁₀ showed the reverse ($PC = 2.8$ for GFP-SUMO; $PC = 1.2$ for GFP-SIM) (Figure 3.10A-B).

Allyson Rice (AR) then individually expressed RFP-(SUMO)₁₀-(SIM)₆ or RFP-(SUMO)₆-(SIM)₁₀ in 3T3 or HeLa cells, where they each formed spherical, micron-sized puncta in the cytoplasm. In live cells the puncta occasionally contacted each other and coalesced into larger structures, suggesting that they are phase separated liquids (data not shown). When co-transfected in HeLa cells with individual YFP-tagged clients, RFP-(SUMO)₁₀-(SIM)₆ puncta only concentrated YFP-SIM. Reciprocally, RFP-(SUMO)₆-(SIM)₁₀ puncta only concentrated YFP-SUMO. In both cases, neither YFP alone nor clients with mutations at their

Figure 3.12: *In cis* Scaffold Behavior is Analogous to the *in trans* Scaffold System (polySUMO and polySIM)



Related to Figure 3.10. (A) *PCs* of the binding site mutant clients YFP-(SIM VADA)₃ or YFP-(SUMO FAAA)₃ when co-transfected with either RFP-(SUMO)₁₀-(SIM)₆ or RFP-(SUMO)₆-(SIM)₁₀, respectively, as well as of YFP into both scaffolds, as in Figure 3.10D. Average *PC* from all analyzed cells (7-21 cells per sample) in each sample is plotted. Error bars represent SEM. Average partitioning of the YFP-SUMO and YFP-SIM clients (from data in Figure 3.10D) is shown alongside for comparison. Dotted line, *PC* = 1. (B) *PCs* for YFP-SUMO (black circles) or YFP-SIM (white circles) when co-transfected with both CFP-polySUMO and RFP-polySIM (*in trans* system), plotted versus distance from the phase diagram diagonal (based on total cellular concentrations of the scaffolds). Each symbol represents the average *PC* of all puncta (typically 1-3) in a given cell. 48 or 31 cells were analyzed for YFP-SUMO or YFP-SIM, respectively. Error bars are propagated errors from measurements and calibration using standard curves to determine the total cellular concentrations of CFP-polySUMO and RFP-polySIM in each cell (see Methods). Vertical dotted line, the diagonal; horizontal dotted line, *PC* = 1.

binding sites were enriched in cellular puncta (Figure 3.12A). I obtained analogous results in 3T3 cells (Figure 3.11). AR also obtained qualitatively analogous data using co-expression of *in trans* polySUMO and polySIM scaffolds along with YFP-tagged clients (Figure 3.12B). However, experimental uncertainties in the relative concentrations of the scaffold components made it difficult to assign cells confidently to one side the diagonal.

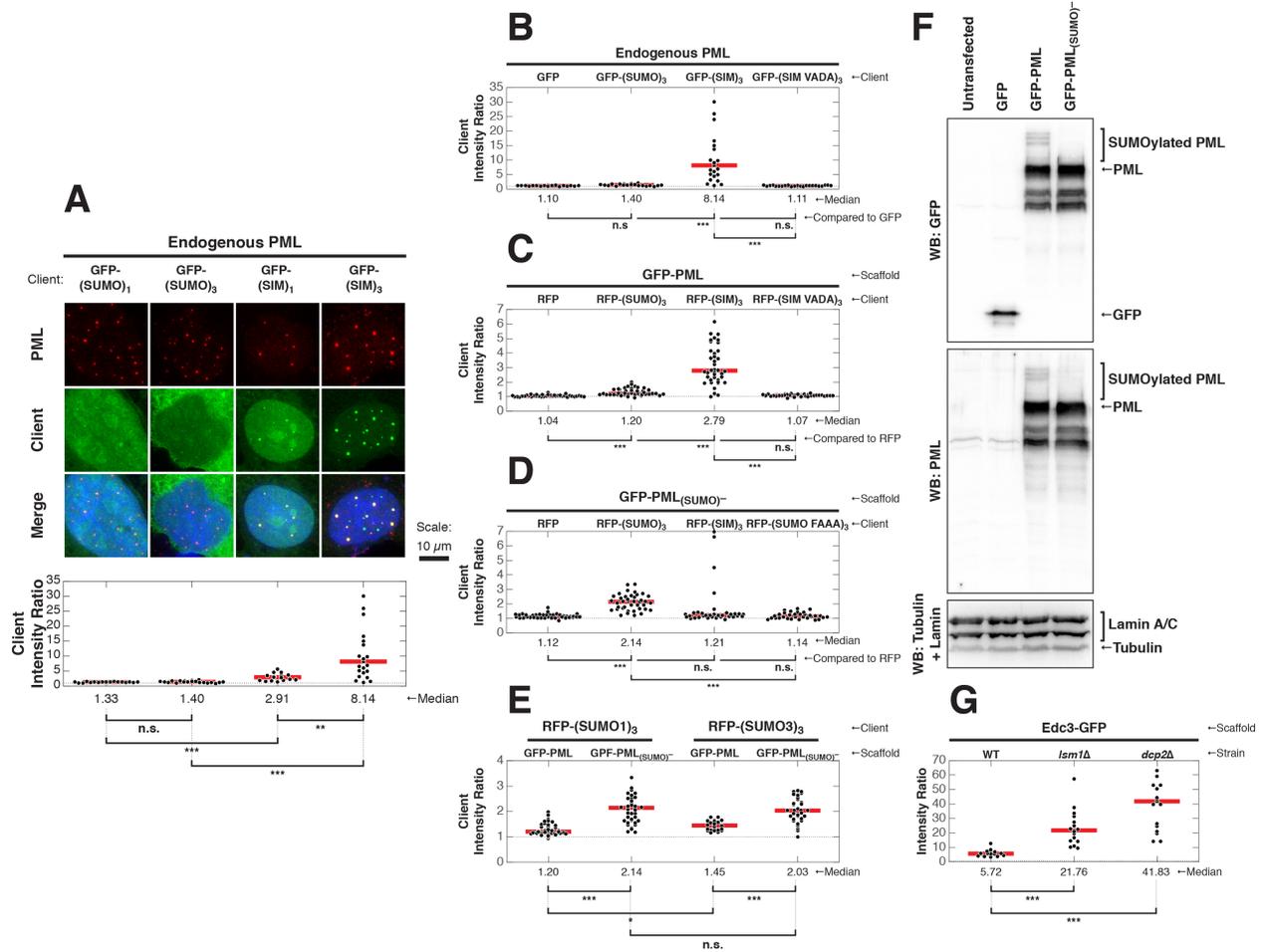
Taken together, our data suggest that mass action-based compositional control can be achieved as robustly in cells as *in vitro*.

3.7 Scaffold Stoichiometries Control Client Recruitment into Natural Cellular Bodies

We sought to determine whether natural cellular bodies could exhibit compositional control analogous to our model systems. We focused on two natural cellular bodies, PML NBs in mammalian nuclei and P bodies in the yeast cytoplasm, systems in which the interactions governing client recruitment were well-characterized and where their stoichiometries were experimentally perturbable.

PML NBs are micron-sized, membrane-less organelles in mammalian nuclei that are involved in processes including DNA damage repair, apoptosis and anti-viral responses [88]. The PML protein appears to be the primary scaffold for these bodies [73]. PML can self-assemble via elements within its Tripartite Motif (TRIM) [68, 2], and also via binding of its conserved SIM element to SUMOs conjugated at up to eight sites in the protein [106, 125]. Though not strictly required for body assembly [120, 16], SUMO-SIM interactions likely contribute substantially to body architecture, as deletion of the SIM motif or perturbations to PML SUMOylation via mutagenesis, viral infection, or knockdown/overexpression of SUMO ligases/proteases can cause changes in the size, number, morphology, or dynamics of PML NBs [103, 12, 125, 144, 62, 63]. SUMO-SIM interactions also appear to be critical for the recruitment of many PML NB clients (e.g. Daxx and Sp100) [94, 135, 150].

Figure 3.13: Partitioning of Clients into PML NBs Requires SUMO-SIM Binding and Depends on Levels of PML SUMOylation

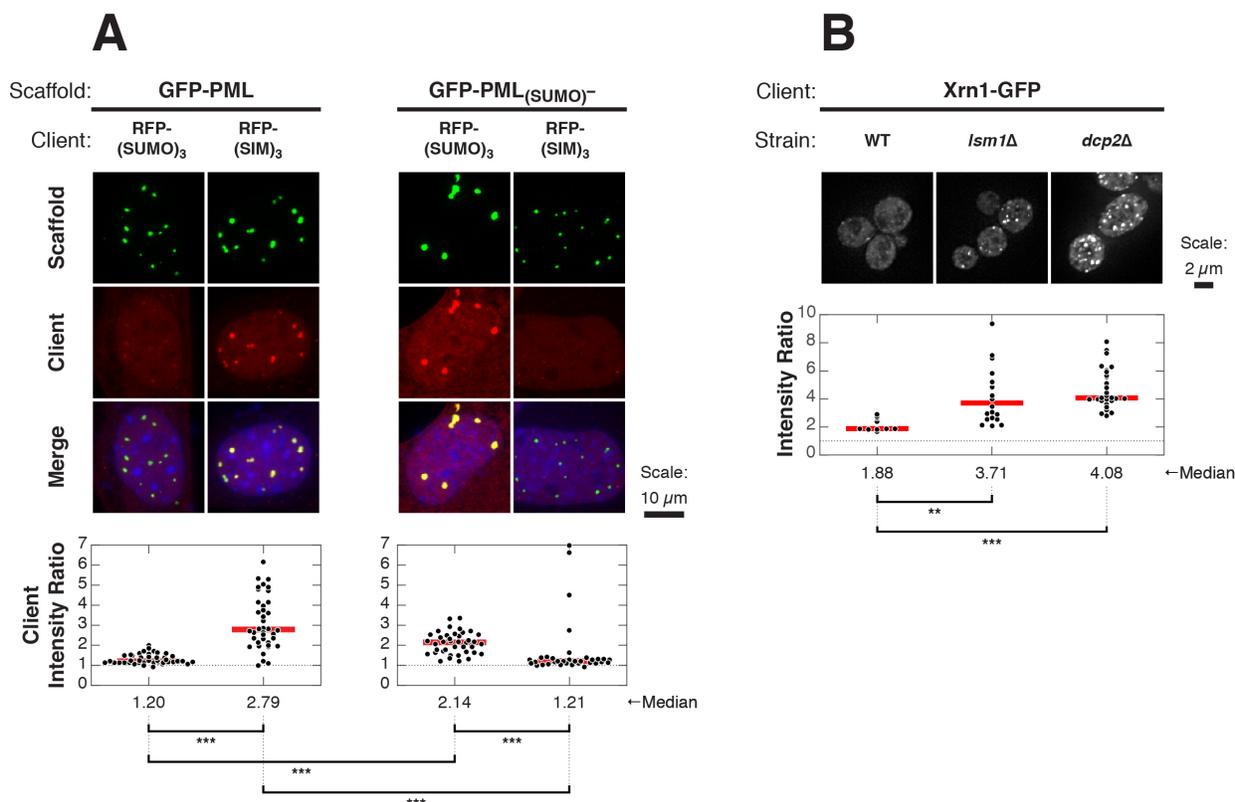


Related to Figure 3.14. (A) U2OS cells were transfected with GFP-low-valency clients (green), fixed, and stained with an antibody against PML (red) and Hoechst 33342 (blue) to visualize the nucleus. Representative images are shown in the top panel. Plots (bottom panel) show *IRs* from individual cells (16-23 cells per sample) (black dots) and median values (red horizontal line). Each symbol represents the average *IR* for all puncta (10 ± 7) in a given cell. Distributions were statistically compared using the Wilcoxon rank sum test followed by the Bonferroni correction for multiple comparisons to determine significance. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; n.s., not significant. Dotted line, $IR = 1$. (B) *IRs* of GFP and the binding site mutant client GFP-(SIM VADA)₃ into PML NBs in U2OS cells, analyzed and displayed as in (a). 17-23 cells were analyzed per sample. GFP-(SIM)₃ [reproduced here along with GFP-(SUMO)₃ from (A)] was likewise compared to GFP-(SIM)₃. (C) *IRs* of RFP and the binding site mutant client RFP-(SIM VADA)₃ when co-transfected with GFP-PML in *PML*^{-/-} MEFs, analyzed and displayed as in Figure 3.14A and panel (A). 36-44 cells were analyzed per sample, with an average of 16 puncta per cell. RFP-(SIM)₃ [reproduced here along with RFP-(SUMO)₃ from Figure 3.14A] was likewise compared to RFP-(SIM VADA)₃. (D) *IRs* of RFP and the binding site mutant client RFP-(SUMO FAAA)₃ when co-transfected with GFP-PML(SUMO)⁻ in *PML*^{-/-} MEFs, analyzed and plotted as in in Figure 3.14A and panel (A). 32-41 cells were analyzed per sample with an average of 5 puncta per cell. RFP-(SUMO)₃ [reproduced here along with RFP-(SIM)₃ from Figure 3.15B] was likewise compared to RFP-(SUMO FAAA)₃. (E) *IRs* of the RFP-(SUMO)₃ client constructed with the human SUMO3 paralog when co-transfected with GFP-PML or GFP-PML(SUMO)⁻ in *PML*^{-/-} MEFs, analyzed and plotted as in Figure 3.14A and panel (A). 24-44 cells were analyzed per sample, with an average of 3 or 8 puncta per cell for the two scaffold constructs, respectively. *IRs* for the RFP-(SUMO)₃ client constructed with the human SUMO1 paralog are reproduced here from Figure 3.14A for comparison. (F) Western blot analysis of *PML*^{-/-} MEFs, either untransfected or transfected with GFP, GFP-PML, or GFP-PML(SUMO)⁻. PML constructs were detected with anti-GFP (top) and anti-PML (middle) antibodies. Note that fewer and less intense higher molecular weight bands above the dominant PML band are apparent for the GFP-PML(SUMO)⁻ construct compared to the GFP-PML construct, indicative of decreased steady-state SUMOylation levels. Simultaneous blotting with anti-Tubulin and anti-Lamin (bottom) served as a loading control. (G) Distributions of Edc3-GFP *IRs*, displayed and analyzed as in 3.14B. 1-3 P bodies per cell were analyzed, from a set of 4-6 cells per sample.

I initially examined partitioning of GFP-tagged SUMO/SIM clients into endogenous PML NBs in U2OS cells (Figure 3.13A). Immunofluorescence imaging using an antibody against PML revealed numerous PML NBs in nearly all cell nuclei. For each client I measured the ratio of GFP fluorescence intensity within the PML NBs to that in the bulk nucleoplasm (Intensity Ratio, $IR = \text{Intensity}_{\text{PML NB}} / \text{Intensity}_{\text{nucleoplasm}}$, see §B.3). GFP-SIM was enriched in these bodies with a median IR of 2.9. In contrast, as previously reported for monovalent SUMO clients [4], GFP-SUMO had little enrichment in the PML NBs (median $IR = 1.3$). Increasing valency increased the enrichment of the preferred client into PML NBs [median $IR = 8.1$ for GFP-(SIM)₃], but had no effect on the impartial client [median $IR = 1.4$ for GFP-(SUMO)₃]. Neither GFP alone nor a client with mutated SIM sites were enriched (Figure 3.13B).

The selective, valency-dependent partitioning into PML NBs is analogous to the behaviors of our polySUMO/polySIM model system on the polySUMO-enriched side of the phase diagram diagonal [i.e. to that of (SUMO)₁₀-(SIM)₅, above the diagonal]. Our model predicts that PML NBs on the opposite side of the phase diagram diagonal should exhibit inverted partitioning behavior with respect to SUMO versus SIM clients. To create such structures, I expressed either wild type (WT) GFP-PML or a PML mutant (GFP-PML_{(SUMO)-}) lacking some of the known SUMOylation sites in $PML^{-/-}$ mouse embryonic fibroblasts (MEFs) (Figure 3.14A). The mutant protein is SUMOylated in cells, but to a lesser degree than the wild type protein (Figure 3.13F). Both the WT and mutant scaffolds formed micron-sized puncta in nuclei. Like natural PML NBs, GFP-PML puncta substantially recruited RFP-(SIM)₃ (median $IR = 2.8$) but not RFP-(SUMO)₃ (median $IR = 1.2$). In reciprocal fashion, GFP-PML_{(SUMO)-} puncta efficiently recruited RFP-(SUMO)₃ (median $IR = 2.1$), but recruited RFP-(SIM)₃ poorly (median $IR = 1.2$). Neither scaffold could recruit RFP alone nor clients with mutations at the SUMO- or SIM-binding site (Figure 3.13C-D). Moreover, these results were independent of the SUMO paralog (SUMO1 versus

Figure 3.14: Client Recruitment Into Natural Cellular Bodies Is Affected By Scaffold Stoichiometries



See also Figure 3.13. (A) Images of RFP-SUMO or RFP-SIM (red) co-transfected with GFP-PML or GFP-PML_{(SUMO)-} (green) into *PML*^{-/-} MEFs (top panels); nuclear staining with Hoechst 33342 (blue). Plots (bottom panels) show *IRs* from individual cells (black dots) and median values (red horizontal lines). Each symbol represents the average *IR* (see Methods) for all puncta in a given cell. 32-44 cells were analyzed per sample, each on average containing 16 or 5 puncta per cell with GFP-PML or GFP-PML_{(SUMO)-}, respectively. Distributions were statistically compared using the Wilcoxon rank sum test followed by the Bonferonni correction for multiple comparisons to determine significance. ***, $p < 0.001$. Dotted line, $IR = 1$. (B) Representative images of WT, *lsm1* Δ , or *dcp2* Δ yeast strains carrying Xrn1-GFP (green) in their genomes (top panel). Distributions of Xrn1-GFP *IRs* (bottom panel), where each symbol represents *IR* corresponding to an individual P body. 1-3 P bodies per cell were analyzed, from a set of 4-10 cells per sample. Analysis for significance was performed as in (A). **, $p < 0.01$.

SUMO3) used to construct the client (Figure 3.13E).

These data suggest that decreasing the degree of PML SUMOylation can shift the bodies to a region analogous to the opposite side of the SUMO/SIM diagonal, where recruitment of SUMO-containing clients is favored.

In collaboration with the laboratory of Dr. Roy Parker, Ph.D., University of Colorado Boulder, Boulder Colorado), we next explored analogous compositional control in P bodies, protein- and mRNA-rich cellular bodies in the cytoplasm of eukaryotes that promote mRNA decay [109]. P bodies assemble through multivalent interactions of RNA binding proteins composed of modular RNA binding domains and self-associating disordered regions, and mRNA molecules [36]. mRNAs that have exited translation act as important P body scaffolds [134]. We thus asked whether modulating the levels of mRNA, thereby affecting the relative stoichiometry of an important scaffold component, could affect recruitment of clients into P bodies. Dr. Saumya Jain (SJ), Ph.D., used the *lsm1Δ* and *dcp2Δ* yeast strains, which are deficient in mRNA decapping and therefore accumulate deadenylated mRNAs that would otherwise be targets for degradation [109]. SJ then measured the IR of the P body client Xrn1 [74] fused to GFP (Xrn1-GFP), in the WT, *lsm1Δ*, or *dcp2Δ* strains. Xrn1, which contains RNA binding elements, is predicted to be recruited, at least in part, by interactions with RNA. Compared to its recruitment into P bodies in WT cells (median $IR = 1.88$), recruitment in the *lsm1Δ* and *dcp2Δ* strains increased ~ 2 -fold (median $IR = 3.71$ and 4.08 , respectively) (Figure 3.14B). This behavior was qualitatively consistent with the behaviors of our engineered clients (Figure 1). The recruitment of the P body scaffold Edc3 [74] also increased in the two deletion strains concomitant with the increase in deadenylated mRNAs, consistent with the increased partitioning scaffolds that SJ observed in certain regimes of the phase diagram when the concentration of the scaffold binding partner increased (Figure 3.5B). Thus, these data suggest that, like PML NBs, compositional control can be achieved in P bodies by modulation of scaffold stoichiometries.

Collectively these data indicate that the compositions of both of these natural cellular bodies can be controlled by modulation of stoichiometries of scaffold elements, analogous to the behaviors observed in our model systems. This behavior suggests simple cellular mechanisms to rapidly and dramatically alter the composition, and thus function, of cellular bodies in response to stimuli.

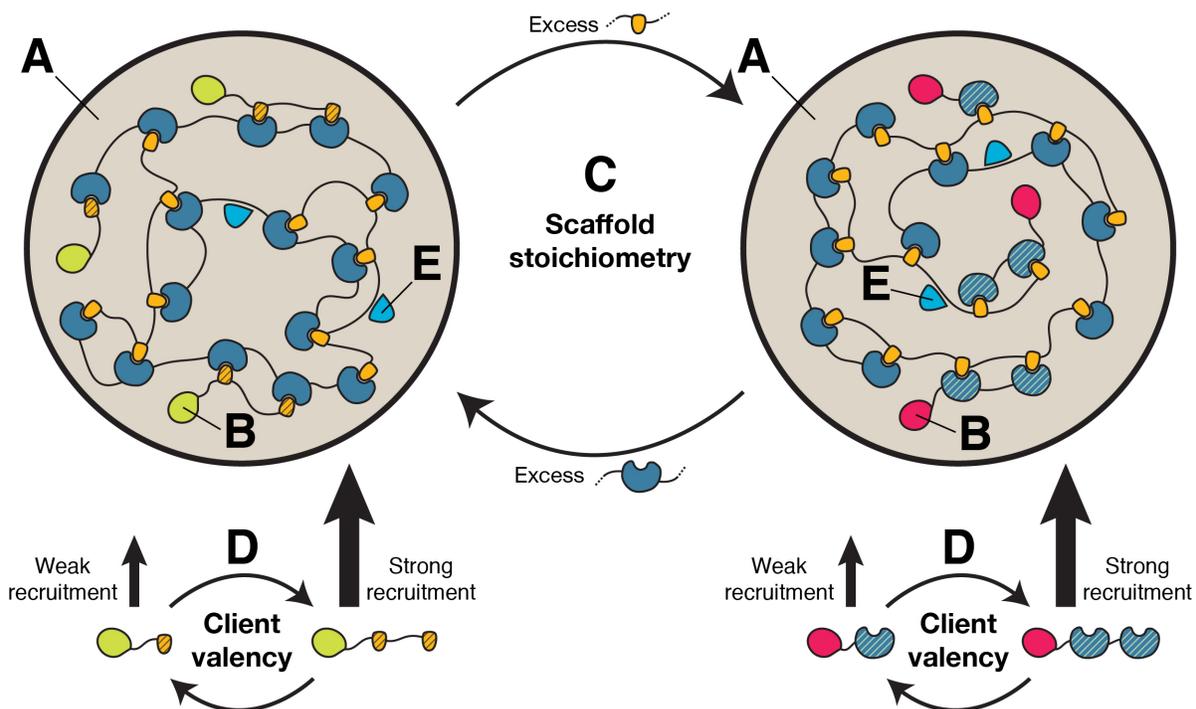
3.8 Discussion

3.8.1 Hierarchical Organization of Cellular Bodies

We propose a hierarchical model for the composition of cellular bodies (Figure 3.15). The model has several key features. First, scaffolds self-associate by multivalent heterotypic interactions and undergo assembly-driven phase separation [91], forming a condensed phase (Figure 3.15A)—i.e., the cellular body. Second, clients partition into bodies by interacting with scaffolds, often utilizing the same types of interacting elements as those between scaffolds (Figure 3.15B). The typically lower valency (and therefore lower apparent affinity) of clients minimizes their competition with the higher affinity scaffold-scaffold interactions. As a result, clients are recruited by binding only to excess, or free, scaffold sites. Thus, their recruitment will be governed by the stoichiometric ratios of the scaffolds (Figure 3.15C). Third, since the enrichment of excess sites switches sharply across the phase diagram diagonal from one class of sites to the other, bodies can change compositions in a switch-like manner as a function of phase diagram position. Fourth, since scaffold and client valencies can affect position on the phase diagram and the degree of client partitioning, respectively, covalent modifications that change valency can be used to rapidly switch between compositional states (Figure 3.15D).

We note that clients that bind regions of the scaffold that are not involved in heterotypic assembly will be recruited, but remain relatively insensitive to changes in scaffold stoichiometries (Figure 3.15E). Moreover, molecules with otherwise appropriate physico-

Figure 3.15: A Model for Compositional Control of Cellular Bodies



Multivalent scaffold molecules (high valency blue and yellow molecules) assemble and phase separate to form the body (A). Many client molecules (low valency blue and yellow molecules, with additional domains) are enriched in the body through binding to free cognate sites in the scaffold that is in stoichiometric excess (B). Client modules have a hatched pattern to distinguish them from scaffold modules. Stoichiometric excess of the scaffold modules can be changed either through changes in the scaffold concentrations (C) or through changes in the scaffold valency (not shown). Since stoichiometric excess of the scaffolds in droplet (A) and bulk (not shown) changes sharply across the phase diagram diagonal, the nature of the clients also switches sharply across the diagonal. Higher valency promotes stronger recruitment of the clients (D). Molecules that bind to other regions of the scaffolds (E, light blue trianguloids) will be recruited independently of the scaffold stoichiometry. Natural bodies are composed of more complicated molecules, with multiple types of interaction elements, but should follow this same logic.

chemical properties (e.g. complementary charge) may also partition into droplets due to non-specific interactions [91]. Clients containing multiple types of interaction elements, some mirroring scaffold-scaffold interactions and others not, could show complex behaviors that are essentially superpositions of these individual effects. This reasoning may explain the recruitment of Xrn1-GFP into P bodies without perturbation of mRNA content (i.e. on what may be the non-cognate side of the phase diagram diagonal), as well as the enhanced recruitment when cellular mRNA is increased, as observed in Figure 3.14B.

3.8.2 Complexities of Natural Cellular Bodies

Although natural cellular bodies are appreciably more complicated than our engineered model systems, their compositions may still be understood with simple extensions to the framework we present here. First, cellular bodies may have multiple scaffolds held together by different types of multivalent interactions. For example, RNA granules likely have multiple scaffolds with contributions from both low complexity sequence elements as well as RNA and RNA-binding domains. PML NBs likely assemble by a combination of TRIM and SUMO-SIM interactions. Multiple types of scaffolds and scaffold interactions may cooperate to synergistically promote polymerization and phase separation, as suggested previously [95]. Moreover, clients may also possess multiple classes of low valency elements that can each interact with scaffolds. Nevertheless, in the absence of cooperativity, one can think of the different interaction motifs independently. For any given class, the corresponding free sites in a scaffold will dictate partitioning of clients that can bind to those sites. Indeed, perturbing one type of interaction motif within PML NBs or P bodies had strong effects on the recruitment of clients that bound to that motif (Figure 3.14).

Second, for some systems, the distinction between scaffolds and clients may be less stark than in our engineered systems (see also §3.4.5). As client valency approaches that of the scaffolds, this distinction breaks down, and the client begins to compete with scaffold-

scaffold interactions. For such clients, the change in partitioning across the diagonal is likely to be less sharp, as we observe for scaffolds (Figure 3.5A-B). Further investigations are needed to understand the distribution of scaffolds, clients, and such intermediate molecules in the various known natural bodies.

Finally, several cellular bodies contain subcompartments (condensed phases within the primary condensed phase) and thus are not simple droplet/bulk systems [18, 75, 138, 48]. Each subcompartment can have a unique composition organized by a distinct set of scaffold molecules. A client may bind to free sites in any or all of the body's subcompartments. Partition coefficients between any two sub-compartments or between a sub-compartment and the surrounding bulk will still result from mass action driven by the corresponding free site ratios.

Thus, despite the inherent complexities of natural cellular bodies, they may still be understood through the lens of our simple model.

3.8.3 Biological Mechanisms of Regulating Body Composition

Biological processes could regulate the composition of cellular bodies by acting on either scaffolds or clients on timescales ranging from physiologic to evolutionary. On the most rapid timescales (seconds to minutes), covalent modifications could change valencies of the scaffold components, shifting the position of the system within the phase diagram. They could also change valencies and affinities of the clients, influencing their degree of partitioning, as suggested here (Figure 3.3), and previously [59, 87]. On slower timescales (hours to days), the scaffold concentrations could change via regulation of expression levels, or their valencies could be changed by alternative splicing. On evolutionary timescales, changes in gene sequences could change the affinity of scaffold components for each other or for clients, shifting composition and function in a more permanent sense.

Some of these processes can be observed in PML NBs. For example, the SUMOyla-

tion of PML is substantially decreased during mitosis concomitant with loss of some SIM-containing clients [45, 38], and phosphorylation of the SIM in PML increases its affinity for SUMO [28]. Similarly, phosphorylation of the SIMs of PML NB clients, including Daxx [31], increases their interactions with the bodies.

3.8.4 Changes in Body Composition may Dictate Changes in Function

Unlike macromolecular machines, cellular bodies continuously rearrange the bonding interactions and organization of their constituent parts, and thus are not stereochemically defined across their lengths. Their functions, therefore, cannot be controlled by allosteric transitions between conformational states, as often occurs with macromolecules. Instead, transitions between compositional states are likely to be key determinants of body function. The differential partitioning of molecules in different regions of the phase diagram implies that it may be most appropriate to consider a given type of cellular body as a distribution of entities (likely defined by a limited number of scaffolds) that lie on a continuum of compositions subject to cellular control. This idea was suggested previously for RNA based bodies based on the related compositions of P bodies, stress granules and RNA transport granules [23]. Similarly, in the case of PML NBs, a variety of structures in different cell types and cell states have been characterized, unified by their enrichment of the PML protein but varying in their composition of other components [38, 97]. Our data suggest that this behavior may be generally applicable to many cellular bodies.

Since function is dictated by composition, this reasoning implies that cellular bodies may exhibit a continuum of functions, rather than a limited set of discrete functions as seen for macromolecular machines in different conformations. Even though cellular body function may be more continuous than discrete, our data suggest that mechanisms could exist, as they do in canonical macromolecular machines, to mediate sharp switches between different functional regimes. Moreover, we and others [115, 142, 102, 110], speculate that pathological

states of cellular bodies may also lie on the same compositional and functional continuum. As such, manipulation or depletion of certain scaffolds may be a promising approach to mitigate the toxicities associated with these pathological granules. Indeed, toxicities due to TDP-43 aggregation in models of ALS can be alleviated by removal of the Ataxin-2 scaffold [44].

3.8.5 Implications for Cellular Body Function

Precise control of client partitioning could mediate colocalization of reaction partners to accelerate reaction rates and increase reaction specificity. For example, polySUMOylation of cellular substrates was recently demonstrated to activate the ubiquitin E3 ligase RNF4, a process that could be driven or enhanced by such compartmentalization [119]. Similarly, metabolic flux could be controlled by colocalization-mediated substrate channeling [129] or the colocalization of a branch point enzyme and downstream molecules in a pathway [29]. Compositional control may also help regulate en masse reactions such as SUMOylation of many cellular factors at PML NBs, which, analogous to DNA repair foci [114], colocalize not only enzymes of the SUMOylation cascade but also several SUMOylation substrates [135]. Partitioning into a cellular body could also serve to sequester components away from their cellular targets, as has been proposed in the regulation of Daxx [88] and the priming of RNA Polymerase II prior to transcription initiation [87]. Indeed, strong depletion of clients from bulk solution through dramatically high PC is consistent with behaviors we observe in our mass action model (Figures 3.5F and 3.6E-F and §3.4.2).

3.8.6 Conclusion

We demonstrate how cellular body assembly, when driven by heterotypic polymerization and concomitant phase separation, naturally leads to a simple and predictive model for compositional control of these structures. Our model suggests how bodies could be switched

sharply between distinct compositional (and thus functional) states on a range of biological timescales. Moreover, it suggests that superficially similar cellular bodies composed of a given set of scaffolds may be markedly different in their composition and function, depending on the relative scaffold stoichiometries. Thus, a complete understanding of cellular bodies may require knowing relative scaffold amounts in addition to scaffold identities.

Our studies thus provide a mechanistic framework for studying the biochemical and regulatory function of cellular bodies owing to properties not attributable to any individual molecule, but rather to those intrinsic to the macroscopic structure itself.

Chapter 4

Reconstitution of Functional Cellular Bodies

4.1 Rationale

We have proposed a model for the formation and regulation of cellular bodies that is based on the phase separation of multivalent interacting macromolecules [91, 7]. Moreover, we have proposed that their compositions, and therefore functions, can be regulated by modulating the relative stoichiometries (or valencies) of scaffold components [7]. However, the extent to which multivalency-driven phase separation can account for the formation and function of natural cellular bodies remains to be tested.

PML NBs are micron-sized structures in mammalian nuclei that are involved in DNA damage repair, apoptosis and anti-viral responses[88]. They exhibit the liquid-like behaviors of coalescence [39, 33] and rapid rearrangements of component molecules [144]. In contrast to most other bodies, to our knowledge, PML NBs are organized by only a single known scaffold molecule: the PML protein [73]. Moreover, the PML protein exhibits multivalency, suggesting it acts as a multivalent scaffold that phase separates to organize the body. As mentioned, PML can self-assemble via elements within its TRIM [68, 2], as well as via interactions between its SIM and up to eight conjugated SUMOs [106, 125]. As such, PML NBs are ideally suited to test whether multivalency-driven phase separation is fundamental to the organization of natural cellular bodies.

As discussed, SUMOylation is not strictly required for body assembly [120, 16]. However, SUMO-SIM interactions are required for the proper assembly of the body [103, 12, 125, 144, 62, 63]. Moreover, many PML body components are modified by SUMO and/or contain

SIMs [135] and PML SUMOylation is required for proper recruitment of components (e.g. Daxx and Sp100 [94, 150]). Thus, SUMO-SIM interactions appear to be necessary for proper body assembly.

These properties suggest that multivalent SUMO-SIM interactions may account in substantial part for PML NB organization. But what is the relative role of SUMO-SIM versus TRIM interactions in the organization of the body? Are SUMO-SIM interactions sufficient to form functional nuclear bodies?

Teasing apart the relative contribution of SUMO-SIM versus TRIM interactions to PML NB formation is experimentally challenging in the context of the natural PML protein in cells, as perturbing TRIM interactions may also impact PML SUMOylation [83]. However, I previously showed that engineered polySUMO/polySIM proteins form punctate structures in mammalian cells that resemble cellular bodies (Figures 3.11 and 3.10B). Here, we begin to test whether synthetic, phase separated polySUMO/polySIM structures can recapitulate PML NBs, both structurally and functionally. I show that these synthetic structures can recruit many of the PML NB components in a manner that depends on the SUMO:SIM stoichiometry in the scaffold. We also show preliminary work that aims find a robust functional readout of PML NB function in mammalian cells to eventually test the whether synthetic polySUMO/polySIM puncta recapitulate PML NB functions. These studies hold promise to address whether multivalency-driven phase separation is an organizing principle in PML NBs, as well as to elucidate the relative role of SUMO-SIM versus TRIM interactions in the PML protein in the assembly of functional PML NBs.

4.2 Recruitment of PML NB Components

We first asked whether polySUMO/polySIM puncta could recruit known PML NB components [105, 135, 88]. I transfected RFP-tagged WT PML (RFP-PML) into *PML*^{-/-} MEFs. When transfected alone, RFP-PML formed numerous micron-sized puncta in the

nucleus (not shown).

I next co-transfected GFP-tagged components with RFP-PML and tested for colocalization (Figure 4.1, left panel). GFP-tagged Daxx, HIPK2, Sp100, and Ubc9 all colocalized with RFP-PML, consistent with their previously reported designations as PML NB components. However, MMS21, SENP2, Aosl-Uba2, ATRX, Hp1 α , p53, and Rb did not show colocalization, in contrast to previous reports [97, 96, 65, 105, 135, 112, 1].

We next asked whether these components would localize to polySUMO/polySIM puncta. I used the *in cis* polySUMO/polySIM scaffolds for these experiments. These scaffolds readily form puncta in mammalian cells (Figures 3.11 and 3.10B). To form the synthetic bodies in the nucleus, I tagged the engineered scaffolds with a C-terminal nuclear localization (NLS; based on a sequence derived from the SV40 T-antigen [80]). All clients that colocalized with RFP-PML also did so with RFP-(SUMO)₁₀-(SIM)₅-NLS, but not with RFP-(SUMO)₅-(SIM)₁₀-NLS (Figure 4.1, middle and right panels). Finally, SENP2, which did not colocalize with RFP-PML, did so with both of the synthetic scaffolds.

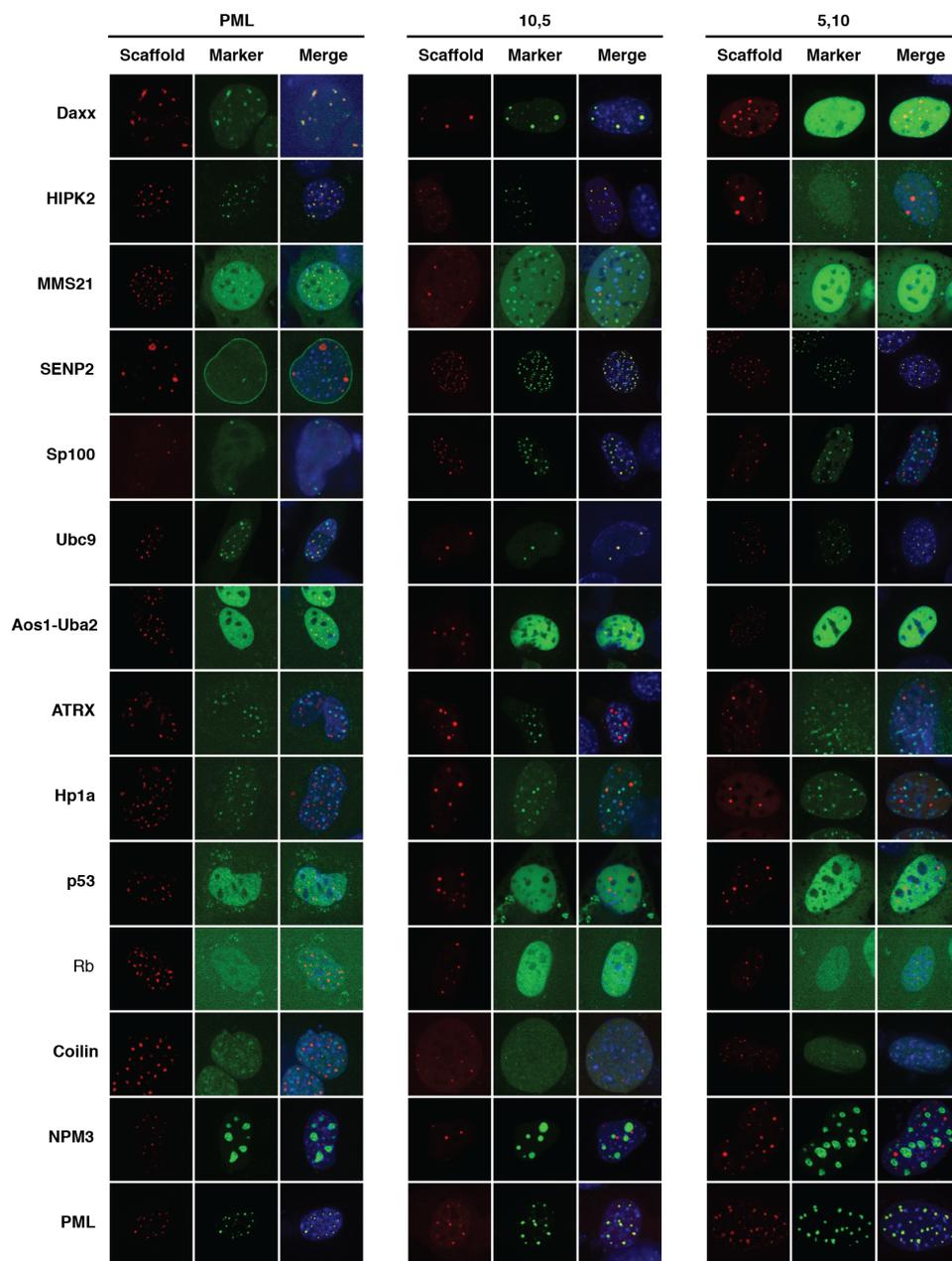
As expected, neither RFP-PML nor the synthetic scaffolds recruited Coilin or NPM3, known components of Cajal bodies and the nucleolus, respectively [15, 40]. Moreover, both of the engineered scaffolds colocalized with GFP-PML, suggesting that the engineered structures were miscible with PML NBs.

Together these data suggest that phase separated structures driven by multivalent SUMO-SIM interactions are sufficient to recapitulate the structural and compositional features of PML NB assembly.

4.3 Cellular Functions of PML

We next sought to ask whether our synthetic bodies could recapitulate aspects of PML NB function. Such experiments require a robust readouts of cellular functions of PML

Figure 4.1: Synthetic PolySUMO/polySIM Nuclear Bodies Recruit Natural PML NB Components



RFP-PML (left), RFP-(SUMO)₁₀-(SIM)₅-NLS (middle), or RFP-(SUMO)₅-(SIM)₁₀-NLS (right) were co-transfected with various reported components of PML NBs, tagged with GFP. The scaffolds were also co-transfected with Coilin or NPM3, components of Cajal bodies and nucleoli, respectively.

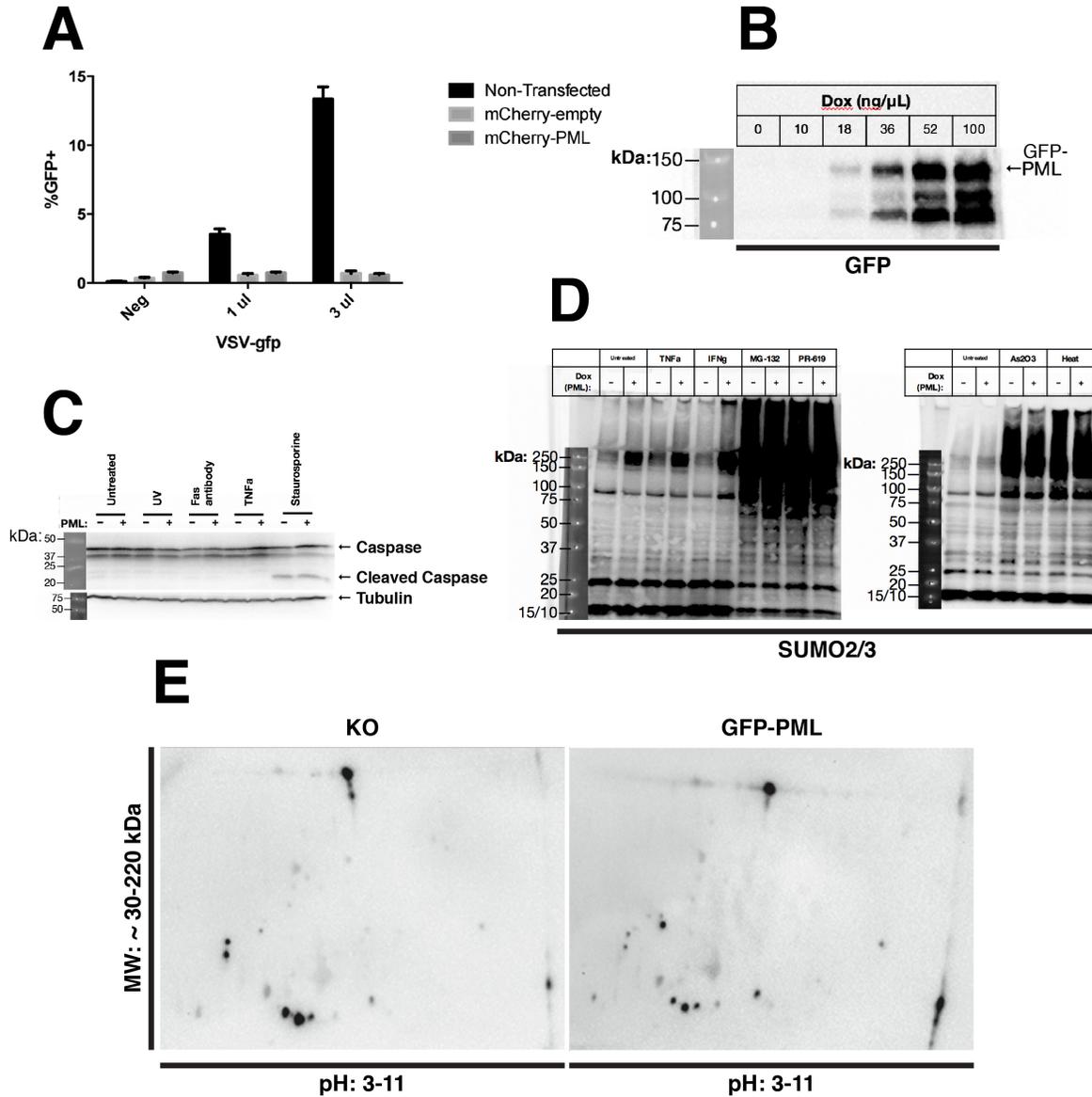
NBs. Specifically, these experiments require an assay for PML function(s) that (i) scores positive in convenient WT cell lines (e.g. MEFs); (ii) scores negative in those lines when PML is deleted (e.g. in $PML^{-/-}$ MEFs); and (iii) can be rescued in those cell lines by re-expression of PML. Such a readout could then be tested for rescue by the engineered polySUMO/polySIM scaffolds alongside the re-expressed PML.

To our knowledge, most known PML NB functions have been demonstrated in non-MEF primary cells or cell lines, and none of the functions has been shown to be rescuable by re-expression of PML. Rather, PML NB functions have been demonstrated either by comparing WT and $PML^{-/-}$ cells (e.g. [41, 141]) or by over-expressing PML in a WT background (e.g. [32]). We thus sought to find readout for PML NB function rescuable by re-expression in $PML^{-/-}$ MEFs, the $PML^{-/-}$ cell line most immediately accessible to me at the time these experiments were performed. Since PML has been reported to play roles in antiviral defense, apoptotic signaling, and SUMOylation, and transcriptional regulation [88], I began probing for cellular readouts for and defects in these functions in $PML^{-/-}$ MEFs.

4.3.1 Antiviral Response

PML is an interferon-regulated gene and has been implicated as playing a role in innate antiviral responses [130, 88]. Moreover, it was shown that PML confers resistance to vesicular stomatitis virus and rabies virus in MEFs [41, 13]. In collaboration with the laboratory of Dr. John Schoggins, Ph.D. [University of Texas Southwestern Medical Center (UTSW), Dallas, TX], we sought to test whether PML re-expression could inhibit infection by vesicular stomatitis virus (VSV) carrying a GFP reporter [122]. Jennifer Eitson (JE) measured infectivity by flow cytometry, by calculating the fraction of RFP-positive cells that were also GFP-positive. Untransfected cells became infected by VSV in a dose-dependent manner (Figure 4.2A). Though co-transfection with RFP-PML inhibited VSV infectivity, the inhibition was indistinguishable from co-transfection with RFP alone, suggesting that the

Figure 4.2: Reported Cellular Functions of PML NBs Fail to Recapitulate in $PML^{-/-}$ MEFs



(A) Triplicate flow cytometry experiments showing infectivity of $PML^{-/-}$ MEFs by VSV stock (carrying a GFP reporter) in untransfected cells or in cells transfected with RFP (mCherry) or RFP-PML. (B) Expression of GFP-PML in an inducible stable cell line expressing GFP-PML in response to doxycycline (Dox) treatment for 12 hours at the indicated concentrations, probed by an anti-GFP antibody. (C) Western blot testing the initiation of apoptosis in the inducible cell line, as indicated by cleaved Caspase (≈ 17 kDa) in response to various stimuli—UV irradiation (8 mJ/cm^2); activation of the Fas receptor by a crosslinking antibody ($0.5 \text{ }\mu\text{g/mL}$); activation of the TNF α receptor by TNF α (10 ng/mL); or Staurosporine ($1 \text{ }\mu\text{M}$). Tubulin, loading control. (D) Western blot showing SUMOylation, probed by an anti-SUMO2/3 antibody, in whole cell lysates (with or without GFP-PML induction) in response to TNF α (10 ng/mL), IFN γ (100 ng/mL), MG-132 ($10 \text{ }\mu\text{M}$), PR-619 ($20 \text{ }\mu\text{M}$), As $_2$ O $_3$ ($1 \text{ }\mu\text{M}$), or heat stress (incubation at $43 \text{ }^\circ\text{C}$). Analogous results were obtained with an anti-SUMO1 antibody (not shown). (E) 2D gel electrophoresis in $PML^{-/-}$ MEFs or a derivative cell line stably expressing GFP-PML in a constitutive manner. Western blot was performed with a mixture of anti-SUMO1 and anti-SUMO2/3 antibodies.

antiviral effect may be nonspecific. These preliminary data suggest that our experimental conditions were not suitable for measuring the antiviral properties of PML NBs.

4.3.2 Apoptotic Sensitization

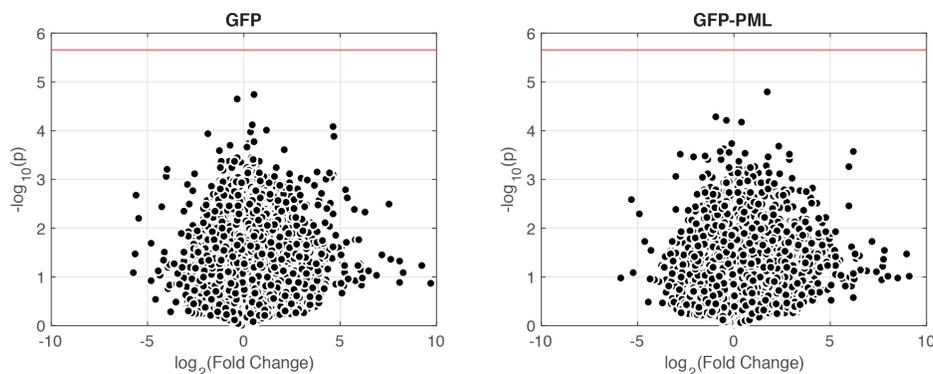
In collaboration with the laboratory of Dr. Pier Scaglioni, M.D. (UTSW), we next tested whether re-expression of PML could sensitize cells to apoptosis, a reported function of PML NBs [88]. Previous studies have showed that deletion of PML made primary thymocytes and splenocytes resistant to the induction of apoptosis by Fas receptor activation or ionizing radiation [141]. This suggests that re-expression of PML in $PML^{-/-}$ cells should sensitize cells to apoptotic stimuli. To avoid the potential for apoptotic sensitization by PML during culturing, for these experiments we used $PML^{-/-}$ MEFs stably expressing GFP-PML under the control of a tetracycline-dependent promoter. GFP-PML expression in these cells was dose-dependent on doxycycline, and a substantial expression signal was observable by western blot after 12 hours of induction with 100 ng/ μ L of doxycycline (Figure 4.2B). With assistance from Dr. Andrea Rabellino, Ph.D., I probed for apoptosis using an antibody that can detect cleavage of Caspase3, a marker of the initiation of apoptosis. These cells initiated apoptosis, with or without GFP-PML re-expression, in response to Staurosporine, a non-specific kinase inhibitor that potently induces apoptosis [9] and is often used as a positive control in apoptosis assays. However, these cells failed to initiate apoptosis, in response to three different apoptotic stimuli—UV irradiation, activation of the Fas receptor, or activation of the TNF α receptor—that have been reported to promote apoptosis in the presence of intact PML-NBs [10]. These results suggest that, at least in this experimental context, apoptosis sensitization is not recapitulated in $PML^{-/-}$ MEFs with GFP-PML re-expression.

4.3.3 Regulation of Cellular SUMOylation

PML NBs enrich SUMOylated proteins in cells [4]. Both PML as well as numerous PML NB clients are known to be SUMOylated [135, 106], and SUMOylation appears to be important for the proper assembly and composition of PML NBs, as discussed. Moreover, enzymes in the SUMOylation cascade are reported to be PML NB components [65, 135]. Thus, it has been suggested that PML NBs are cellular hubs for SUMOylation [88, 135]. As such, we sought to test whether the global SUMOylation status of cells could be a rescuable cellular readout of PML NB function. I probed for SUMOylation in whole cell lysates of the inducible cell lines, with or without GFP-PML induction, with anti-SUMO1 (not shown) or anti-SUMO2/3 antibodies (Figure 4.2D). Induction of GFP-PML expression resulted in no gross changes in cellular SUMOylation (lanes 1-2 and 11-12), except the appearance of a cluster of bands in the 150-250 kDa range, likely corresponding to SUMOylated GFP-PML (as seen in Figure 3.13F). I next probed global SUMOylation in response to $\text{TNF}\alpha$ and $\text{IFN}\gamma$, treatments that mimic apoptotic or viral stimuli, respectively. Neither of these stimuli resulted in significant changes in the SUMOylation pattern in whole cell lysates (lanes 3-4 and 5-6). I next exposed cells to arsenic (As_2O_3) or heat stresses, both reported to induce cellular SUMOylation [121, 64]. Both stimuli increased cellular SUMOylation compared to untreated cells, but the effect was similar for both induced and uninduced cells (lanes 13-14 and 15-16 vs. 11-12). Similarly, I also failed to detect SUMOylation differences in low molecular weight species between induced and uninduced cells in response to the proteasome or deSUMOylase inhibitors (MG-132 and PR-619, respectively [64]), both of which increased cellular SUMOylation (lanes 7-8 and 9-10).

We reasoned that perhaps PML-specific differences in SUMOylation were hidden in our experiments within the poorly resolved cluster of bands in the 50-250 kDa range. We thus attempted to resolve these bands by performing 2D gel analysis (with the Protein Chemistry Lab, Texas A&M University, College Station, TX). However, under these conditions the

Figure 4.3: **Global Gene Expression Profiling in $PML^{-/-}$ MEFs Fails to Detect Significant Changes in Expression Due to PML**



Global gene expression analysis in $PML^{-/-}$ MEFs with constitutively expressed GFP (left) or GFP-PML (right) compared to parent cell line. Average fold changes from duplicate experiments are plotted, where each dot corresponds to a unique gene. Fold changes were tested for statistical significance (t -test) and the resulting p values are plotted for each gene on the ordinate axis. The red horizontal line depicts the Bonferroni-adjusted p value threshold for statistical significance, $p_{\text{adj}} = 2.2 \times 10^{-6}$.

detection sensitivity by 2D gel was significantly reduced compared to 1D gel analysis (Figure 4.2E). The total number of spots observed with 2D gel analysis was much less than the total number of bands observed with 1D gel analysis (compare Figure 4.2E with 4.2D). Together these preliminary data suggest that further experimental optimization is required to robustly detect differences in cellular SUMOylation that result from the presence of intact PML NBs.

4.3.4 Regulation of Transcription

Finally, since many PML NB components are known transcriptional regulators [101], we asked whether PML re-expression in $PML^{-/-}$ MEFs results in any functional changes in the cellular transcription state. To this end, we performed RNASeq analysis (with the UTSW Genomics and Microarray Core) to compare the global steady state expression levels of stable cell lines that constitutively express either GFP or GFP-PML, and compared their levels to the parent $PML^{-/-}$ MEF line (fold change for gene i , $FC_i = \frac{\langle \text{expression}_i \rangle_{\text{GFP or GFP-PML}}}{\langle \text{expression}_i \rangle_{\text{parent}}}$).

No significant changes in gene expression were detected in either the GFP or GFP-PML expressing cell line compared to the parent knockout line, perhaps owing, in part, to a low number of replicates used. However, the expression of 273 genes increased [$\log_2(FC) \geq 2$] and 169 genes decreased [$\log_2(FC) \leq -2$] concomitant with the expression of GFP-PML but not with that of GFP. Though not statistically significant (for all genes, $p > p_{\text{adj}}$, where p_{adj} is the Bonferonni-adjusted p -value threshold for statistical significance using the t -test), these preliminary data suggest that finding robust expression changes owing to PML ought to be feasible with better experimental design.

4.4 Discussion

4.4.1 Reconstitution of PML NB Composition

Our data suggest that our synthetic nuclear bodies composed of engineered poly-SUMO/polySIM scaffolds are remarkably good at recapitulating the composition of natural PML NBs. Not all of the reported PML NB clients I tested were recruited to the PML puncta in our experiments. Recruitment of ATRX, HP1 α , Rb, and MMS21 were reported in non-MEF cell types[97, 1, 112], which may account for the differences in recruitment. Moreover, the recruitment of p53 may require stimulation by ionizing radiation, as reported [96]. Nevertheless, all of our tested PML NB clients that were recruited to PML puncta were also recruited to synthetic puncta formed with (SUMO)₁₀-(SIM)₅. Taken together with studies that show that PML SUMOylation is crucial for client recruitment [150, 94], and that SUMOylation appears to be a hallmark of PML NB components [135], our data suggest an important role for SUMO-SIM interactions in maintaining the proper composition of PML NBs.

Of these clients, none except SENP2 and Ubc9 were recruited to puncta formed by (SUMO)₅-(SIM)₁₀, suggesting that SIMs are perhaps the primary molecular determinants for clients to get recruited to PML NBs. This behavior was analogous to that of the engineered

GFP-SIM client, which was recruited to polySUMO/polySIM droplets *in vitro* and in cells when the SUMO:SIM stoichiometry was in favor of SUMO, and to PML NBs when PML SUMOylation was intact (but not when it was perturbed) (Figure 3.14). This suggests a potential role for PML SUMOylation in regulating the composition of PML NBs, and supports the idea that compositional regulation of PML NBs occurs by modulating scaffold free sites (see Figure 3.15 and Chapter 3). Indeed, PML SUMOylation levels decrease in mitosis, resulting in changes in the client composition of PML NBs [46, 38]. This regulation could in principle be conducted by the balance of SUMO ligase cascade versus the SUMO proteases [56]. SENP2 and Ubc9 may not be subject to this mode of regulation by changes in SUMO:SIM stoichiometries, as they bind to SUMO at sites other than the SIM-binding site on SUMO [117, 118, 27]. However, that they were both strongly recruited to the engineered puncta suggests that they may help mediate this regulatory effect in PML NBs; the lack of SENP2 recruitment to PML puncta in our experiments was likely due to over expression and complete deSUMOylation of the co-transfected PML.

Finally, PML colocalized with both of the engineered scaffolds. This is consistent, within our framework of compositional control (Figure 3.15), with the notion that PML behaves as a scaffold. Low valency clients (e.g., GFP-SIM or Daxx) cannot compete occupied scaffold sites and are subject to regulation by scaffold stoichiometries, scaffolds (e.g., polySUMO/polySIM or PML) due to their higher valency are strongly recruited to the bodies at all stoichiometries and in turn govern the recruitment of clients (see Chapter 3). This further supports the notion of the client and scaffold framework as a principle in the organization of PML NBs.

4.4.2 Reconstitution of PML NB Function

There still remains a need for a robust functional readout to test whether multivalent SUMO-SIM interactions suffice to form functional nuclear bodies. Our experiments failed to

detect functional readouts for a variety of reasons. First, it is possible that the antiviral effect of RFP alone was due to a nonspecific triggering of the interferon pathway by transfected DNA. Indeed it has been reported that plasmid DNA can initiate the interferon response [92]. Thus, conducting these experiments with our stable cell lines may be a way to address this issue.

Second, lack of apoptotic sensitization may be due to the fact that the sensitizing ability of PML is absent or was lost in our cultured cell lines. To our knowledge, most reported effects on apoptotic sensitization by PML were performed using primary cells (e.g. splenocytes and thymocytes). It is possible that in our cell lines, MEFs immortalized by continuous culturing, the ability of PML to induce apoptosis in response to stress may have been selected against during the immortalization process. Using cell lines more similar to the ones where this function has been reported may allow us to recapitulate this effect with PML re-expression. Third, though there are no gross differences in cellular SUMOylation in MEFs, there may be a small number key factors whose SUMOylation levels change in response to PML. More sensitive proteomic approaches (e.g. mass spectrometry) may be required to detect such changes.

Further, our experimental design for the RNASeq experiments was apparently suboptimal. The noise in our measurements was such that duplicate experiments were insufficient to determine statistically significant changes in gene expression. Increasing the number of replicates may address this issue. Direct comparisons of gene expression to WT cell lines may also control for some of our observed variability and increase sensitivity for detecting specific changes due to PML deletion.

Finally, PML bodies colocalize with many components of the DNA damage response, undergo changes in their structure and composition in response to DNA damage, and confer an increased efficiency of homology-directed repair [88, 96, 14]. Thus DNA damage repair may be an additional promising functional readout for PML NBs.

4.4.3 Conclusion

Our results indicate that the synthetic polySUMO/polySIM bodies in large part recapitulate the structure and composition of PML NB. Further work is required to determine a robust cellular readout for PML NB function. The next step would then be to test whether synthetic polySUMO/polySIM bodies recapitulate such PML NB functions.

Chapter 5

Concluding Remarks

This chapter contains material adapted from a manuscript in preparation [6].

Cellular bodies are complex macroscopic assemblies that likely carry out important functions in cells. First observed in as early as the 1800s, cellular bodies have long been studied by cell biologists. They are unique among cellular structures in their ability to achieve compartmentalization without membranes along with rapid dynamics of their constituent molecules. These observed properties and behaviors are only recently coming to be understood at a molecular level. The recent strides in understanding cellular bodies have come in a large part by applying concepts from polymer chemistry and soft matter physics and by efforts to biochemically reconstitute and study model cellular bodies *in vitro*.

5.1 Practical Considerations in Studying Phase Separated Structures

Major advances in cell biology have often involved the confluence of genetics, cell biology and biochemistry. We anticipate that, likewise, a complete understanding of cellular bodies will continue to require both cellular and biochemical approaches. However, cellular bodies represent a unique class of biological structures in many respects, being unlike both macromolecules as well as canonical membrane-bound organelles. As such, their study by existing experimental approaches poses unique practical challenges.

5.1.1 Biochemical considerations

Obtaining robust biochemical models for cellular bodies has proved challenging. Unlike macromolecules or membrane-bound organelles, cellular bodies are difficult to purify from cells as intact, functional structures. Being liquid-like, cellular bodies exhibit rapid molecular rearrangements. Thus, when necessarily diluted during standard purification procedures, we expect the thermodynamically favored disassembly of these structures under these conditions to occur rapidly. Where purification has been attempted, this approach generally results in static structures that likely do not recapitulate the biochemical environment of native bodies *in vivo* (e.g. [89]). As mentioned, it is known that several types of cellular bodies (and perhaps all of them) require ATP-dependent processes to maintain their dynamics in liquid-like states [18, 86, 21], and likely transition from liquid into solid-like states without ATP or outside the cellular environment during purification. In addition, many bodies may have solid-like cores within a liquid-like shell [138, 75, 18]. Though this feature may prove fortuitous for proteomic studies (e.g. [75]) since the structure remains intact during purification, additional or alternative procedures may be required to ensure that the purified material retains physiological functionality and the the dynamic, liquid-like biochemical environment.

A promising alternative strategy for obtaining biochemical models for cellular bodies is the reconstitution of phase separated structures from a minimal set of natural or engineered protein and/or RNA molecules. Reconstitution may provide the advantage of reducing the complexity of natural cellular bodies, which could contain up to 100s of components [52, 23], down to a minimal functional set. Indeed, phase separated droplets can be generated with as few as 1-3 molecules [91, 95, 107, 100, 59, 81, 87]. However, at least two challenges face the reconstitution approach. First, the ultimate success of this strategy will hinge on identifying the key minimal sets of components that can recapitulate phase separation as well as physiological function. Second, proteins can occupy a diversity of states of matter,

from dynamic, liquid-like phases, to fibrils, to unfolded precipitates [136, 47], each producing macroscopic structures visible by techniques commonly used for these studies, such as light microscopy and light scattering. The success of reconstitution approaches will thus require scrutiny in distinguishing physiologically relevant structures from artifactual, precipitated structures.

5.1.2 Genetic considerations

The cellular effects of perturbing key cellular body components required for body assembly may provide clues about the function of the body. A functional understanding of cellular bodies requires the discovery of functions that arise specifically from organizing molecules into a body, which are distinct from the functions of the molecules individually. However, typical genetic manipulations involving gene knock-outs or knock-downs of key body components may be confounded by both of these classes of function. Experimental perturbations that uncouple the functional contribution of a body from the intrinsic functions of the molecules within it will therefore prove valuable. For example, mutations could be introduced that affect phase separation propensity by perturbing solubility (e.g. mutations at surface sites) without affecting intrinsic molecular activity. Additionally, genetically fusing well-characterized domains that either favor or disfavor phase separation may also influence body formation without influencing molecular activity. Modulation of additional parameters that affect phase separation, such as temperature or volume, may also be useful approaches. During germline development in *C. elegans*, for example, the total amount of nucleolar material stays constant while the cellular (and nuclear) volume decreases [143]. Thus, the nucleolar volume fraction gradually increases during development. Mapping this developmental progression to a potential functional output of the nucleolus may prove to be a powerful way to determine the specific functional advantage of compartmentalizing nucleolar components. Moreover, genetically altering the degree of partitioning of components to

the body (e.g. by changing their valencies) could be a way to isolate the effects of cellular bodies that are independent of the molecules. Finally, generating synthetic/engineered cellular bodies into which cellular components can be targeted may provide yet another approach for studying the function of compartmentalization independent of molecular activity.

5.2 Remaining Questions

Research in the last several years has made significant strides toward understanding the molecular mechanisms that underlie the formation, regulation, and function of cellular bodies and their relationship to disease. It appears that many cellular bodies form through LLPS, driven by interactions of multivalent molecules. This mechanism naturally leads to routes to control the assembly and disassembly, composition, and physical properties of cellular bodies. These routes in turn have implications for the biochemistry that occurs within bodies and thus their cellular functions. However, many important questions remain.

5.2.1 The Function of Cellular Bodies

Perhaps most importantly, we do not understand in most cases what biochemical or cellular functions uniquely emerge from organizing molecules into cellular bodies. In many cases we can infer function from the collection of body components, but we do not understand how the activities of those components change by virtue of being in the body rather than being more uniformly distributed in the cell. When examined, the phenotypes resulting from disruption of bodies are relatively subtle and bodies do not appear to be essential for the viability of cells or organisms (e.g. [10, 36]). Yet, they are conserved over evolution, suggesting they do play important functional roles, perhaps in response to particular stimuli or stresses. We have suggested various possibilities for the biochemical functions for cellular bodies based on a small number of experiments and physical intuition, but in most cases this remains an open question.

We also do not understand the relationship between the microscopic properties of the component molecules and the macroscopic properties of the bodies. We also do not understand how the latter relate to biochemical and cellular functions or if cells regulate these properties to functional effect. As discussed above, it seems likely that the role of energy consuming processes is closely related to these questions, though the exact molecules and mechanisms remain unknown.

5.2.2 Structural Organization of Cellular Bodies

Although at low resolution many bodies appear to be homogeneous, electron microscopy and super-resolution light microscopy have both indicated that many bodies in fact contain internal organization at multiple length scales [15, 18, 75]. For example, micron scale nucleoli contain 100-nanometer scale subcompartments, which show 10-nanometer scale material heterogeneity. Does this organization occur in other bodies and in general how does it come to be? Is it dynamically controlled? Is it functionally important?

5.2.3 Cellular Body Composition

What are all of the factors that control the composition of a given body? Recent work has demonstrated the importance of direct binding interactions and electrostatic effects, but are there other considerations? What do we need to know about a body (or even a simplified phase separated droplet) to predict how other molecules will partition into it? How is composition finely tuned so that distinct bodies can coexist in a cell with shared components but functional differences, and how does composition respond to external stimuli?

5.2.4 The Relationship of Cellular Bodies to Disease

Does the idea that cellular bodies are generated through phase separation and multivalent assemblies have implications for disease, and could this enable novel clinical oppor-

tunities? Existing data suggests that bodies may lie across a continuum of material and compositional states. Moreover, aberrations in this natural spectrum, some of which may involve misregulation of fiber formation, are implicated in neurodegeneration. How do these aberrations affect cell physiology? Could this understanding present new possibilities in characterizing and treating neurodegenerative disease. This is probably only one of many instances where a mechanistic understanding of cellular bodies could have medical implications.

5.2.5 Generality of Phase Separation in Biology

Finally, what other cellular structures might be organized by phase separation. Cells contain numerous types of membraneless macroscopic structures with rapid molecular dynamics. Indeed, components of centrosomes and the mitotic spindle can undergo phase separation *in vitro* [77, 146]. In principle, any system composed of interactions between multivalent entities should have the propensity to phase separate under appropriate solvent conditions.

Addressing these questions will likely require new technologies and new conceptual approaches, drawing on disciplines ranging from genetics to biochemistry to physics. Their answers promise to explain how nanometer-scale molecules can give rise to micron-scale cellular organization and the function of this organization in biology.

5.3 Conclusion

Cellular bodies are exemplars of emergent behaviors of molecules [70]. That is, they demonstrate how the connection between angstrom-scale properties of a collection of molecules and the resulting effects on the macroscopic structure they create can be under-

stood. A major question that remains for cellular bodies is how their properties functionally influence the pathways with which their constituents are associated. It remains an important challenge to tease apart the functional contribution of the macroscopic structure from that of the component molecules. Cellular bodies may have evolved to afford cells additional regulatory capacity controlling reaction rates, specificity, or pathway flux. Elucidating their influence on specific pathways may provide a deeper understanding of how cells regulate signaling or metabolic pathways.

Appendices

Appendix A

Experimental Procedures

This appendix contains material adapted from a manuscript in press [7].

A.1 Genes, RNA, and Plasmids

polyPRM, polySH3, PTB, and the polyUCUCU RNA were described previously [91]. The RNA client UCUCU-AF647 (5'-UCUCUAAAAA-3'; 3'-labeled with AF647), as well as (SUMO)₅ and (SIM)₅ as synthetic genes were purchased from Integrated DNA Technologies. Decavalent, fused, and low valency SUMO/SIM constructs were constructed from (SUMO)₅ and (SIM)₅ by PCR. To prevent conjugation and proteolysis, we mutated the C-terminal di-glycine motif in all SUMO proteins (see §A.2). The RRM client was constructed from the first RRM domain of PTB. The mCherry, mEGFP, mVenus, and mCerulean (referred to as RFP, GFP, YFP, and CFP, respectively) fusion proteins were produced by cloning into corresponding vectors (Clontech). Coding sequences for natural PML NB clients were obtained from Open Biosystems (GE Dharmacon), Addgene, or the Ultimate ORF Clones collection (Thermo Fisher) and cloned into the mEGFP mammalian expression vector (Clontech). Sequences of molecules used in this study (except the natural PML NB clients) are listed in Table A.1.

A.2 Modules and Proteins

Sequences for the modules and proteins used in this study are shown in Table A.1. In all engineered proteins, modules were separated by intervening Gly- and Ser-rich flexible

Table A.1: Sequences of Modules and Proteins Used in this Study

Module or Protein	Sequence
SUMO module (SUMO3)	MSEEK P KEGVKTENDHINLKVAGQDGSVVQ FKIK RHTPLSKLMKAYCERQGLSMRQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQT VV
SUMO module (SUMO1)	MSDQEAKPSTEDLGDKKEGEYIKLKVIGQDSSEI HFKVK MTTHLKKLKESY CQRQGVPMNSLRFLFEGQRIADNHTPKELGMEEEDVIEVYQEQT VV
SIM module (PIAS _x)	KVD VIDL TISSSSDEEEDPPAKR
SH3 module (Nck)	DLNMPAYVKFNMAEREDLSLIKGTKVIVMEKSSDGGWWRGSYNGQVGFPSNYVTEEGDSPL
PRM module (ABL1)	KKKKTAPTPPKR
RNA module	5'-UCUCU-3'
RRM module (PTB)	KGDSRSAGVPSRVIIHRKLPIDVTEGEVISLGLPFGKVTNLLMLKGKNQAF IEMNTEEAANTMVNYTTSVTPVLRGQPIYIQFSNHKELKTDSSPNQARAQA ALQAVNSVQSGNL
PTB	MDGIVPDIAGVTKRGSDELSTCVTNGPFIMSSNSASAANGNDSKKFKGDS RSAGVPSRVIIHRKLPIDVTEGEVISLGLPFGKVTNLLMLKGKNQAFIEMN TEEAANTMVNYTTSVTPVLRGQPIYIQFSNHKELKTDSSPNQARAQAALQA VNSVQSGNLALAASAAAVDAGMAMAGQSPVLR IIVENLFYPVTLDLVHQIF SKFGTVLKIITFTKNNQFQALLQYADPVSAQHAKLSLDGQNIYNACCTLRI DFKLTSLNKYNNDKSRDYTRPDLPSGDSQPSLDQTMAAFGLSVPNVHG ALAPLAIPSAAAAAAAAAAGRIAIPGLAGAGNSVLLVSNLNPERVTPQSLFIL FGVYGDVQRVKILFNKKNALVQMADGNQAQLAMSHLNHKLHGKPIRITL SKHQNVQLPREGQEDQGLTKDYGNSPLHRFKKPGSKNFQNIFFPSATLHLS NIPPSVSEEDLKVLFSSNGGVVKGFKFFQKDRKMALIQMGSVEEAVQALID LHNHDLGENHHLRVSFSKSTI
PML	MEPAPARSPRPQQDPARPQEPTMPPPETPSEGRQPSPPSPPTERAPASEEE FQFLRCQQCQAEAKCPKLLPCLHTLCSGCLEASGMQCPICQAPWPLGADTP ALDNVFFESLQRRLSVYRQIVDAQAVCTRCKESADFWCFECEQLLCAKCFE AHQWFLKHEARPLAELRNQSVREFLDGTRKTNNIFCSNPNHRTPTLTSIYC RGCSKPLCCSCALLDSSHSELKCDISAEIQQRQEELDAMTQALQEQDSAFG AVHAQMHAAVGQLGRARAETEELIRERVRQVVAHVRAQERELLEAVDARYQ RDYEEMASRLGRLDAVLQRIRTGSALVQRMKCYASDQEVLDMHGFLRQALC RLRQEEPQSLQAAVRTDGFDEFKVRQLDLSSCITQGKDAAVSKKASPEAAS TPRDPIDVDLDVSNTTTAQRKRCSTQCPRKVIKMESEEGKEARLARSSPE QPRPSTSKAVSPPHLDGPPSPRSPVIGSEVFLPNSNHVASGAGEAEERVVV ISSSESDAENSSSRELDSSSESSDLQLEGPSTLRVLDENLADPQAEDRP LVFFDLKIDNESGFSWGYPHFLI

Sequences mutated in the binding site mutants [FKIK or FKVK (SUMO) and VIDL (SIM)]; the mutated di-Gly motif in SUMO (now di-Val); and the 3 Lys residues mutated in PML_(SUMO)- are shown in bold.

linkers. N-terminal **Trp** and/or C-terminal **Cys** residues were added to the sequences for purposes of quantitation and maleimide-based labeling with fluorophores, respectively. Modules in the engineered RNA were separated by intervening 5'-AAAAA-3' linkers. The SH3/PRM and PTB/RNA systems were described in detail previously Li:2012fx. SH3 and PRM modules were derived from Nck and ABL1, respectively. We used the human SUMO3 paralog for all SUMO-containing scaffolds in this study. Clients for *in vitro* experiments were constructed with SUMO3, and for cellular experiments with human SUMO1 (except in Figure 3.13E). To prevent conjugation/proteolysis, all SUMO modules had the C-terminal di-Gly motif [56] mutated to di-Val (except the first two SUMO modules in the (SUMO)₉-(SIM)₈ scaffold, where the di-Gly motifs were left intact (Figure 3.8C)). We used the SIM from PIASx [127] for all SIM-containing scaffolds and clients. Binding-site mutants were designed using previously reported SUMO-SIM structures [128, 123] and generated by mutating the FKIK (SUMO3) or FKVK (SUMO1) sequence in each SUMO module to FAAA, and the VIDL sequence in each SIM to VADA. We used the PML-IVa isoform in our cellular experiments [76]. PML_{(SUMO)-} was generated by mutating K65, K160, and K442 to Arg.

A.3 Protein Expression and Purification

All purified proteins were expressed and purified similarly. Proteins were expressed in *E. coli* strain BL21 DE3T1R by induction with 1 mM IPTG. Proteins were purified with Ni-NTA Agarose Resin (Qiagen) or Amylose Resin (NEB), followed by ion exchange [Source 15Q and/or Source 15S (GE Healthcare)] and size exclusion chromatographies using a Superdex 200 or Superdex 75 gel filtration columns (GE Healthcare). Proteins were labeled using maleimide-conjugated Alexa dyes (Life Technologies) following the manufacturers protocol.

Procedures dealing with polyPRM, polySH3, and PTB were described previously [91]. The RFP-RRM client was expressed and purified similarly to PTB. PolySUMO, polySIM, and all SUMO/SIM clients were cloned into a pET-derived vector (Novagen) containing an N-

terminal (His)₁₀ tag and a C-terminal (Arg-Lys)₅ tag, each separated from the protein coding sequence by a Tobacco Etch Virus Protease (TEV)-cleavable site. All constructs containing only SIMs (without SUMOs or fluorescent proteins) also contained a Maltose Binding Protein (MBP) tag downstream of the (His)₁₀ tag but prior to the TEV site. (SUMO)₁₀-(SIM)₅ and (SUMO)₅-(SIM)₁₀ proteins (with or without fluorescent protein tags) were cloned into a pMal-C2-derived vector (NEB) with the addition of a C-terminal (His)₆ tag.

All *in trans* SUMO/SIM scaffold proteins as well as the corresponding clients were expressed and purified similarly. Proteins were expressed in *E. coli* strain BL21(DE3)-T1R. Transformed bacteria were grown to an optical density (measured at 600 nm) of 0.6-0.8 and then induced with 1 mM IPTG at 37 °C for 4 hours for constructs containing only SUMOs (without SIMs or fluorescent proteins) or at 18 °C for 16 hours for all other constructs. Cells were lysed by cell disruption (Emulsiflex-C5, Avestin), and lysates were cleared by centrifugation. Proteins were affinity-purified with Ni-NTA Agarose Resin (Qia-gen), followed by cation exchange chromatography using Source 15S Resin (GE Healthcare). Purified fractions were pooled and cleaved with TEV protease and flowed through Source 15S resin to remove uncleaved proteins. Cleaved protein products were purified further by anion exchange chromatography using Source 15Q Resin (GE Life Sciences), followed by size exclusion chromatography using a Superdex 200 gel filtration column (GE Healthcare) equilibrated with the corresponding experimental buffer (see below).

The *in cis* scaffold proteins, (SUMO)₁₀-(SIM)₅ and (SUMO)₅-(SIM)₁₀, were cloned into the pMal-C2-derived vector described above. These were affinity purified first on Ni-NTA Agarose Resin and then on Amylose Resin (NEB), followed by TEV protease cleavage and cation exchange chromatography with Source 15Q. Purified fractions were pooled and loaded onto a Superdex 200 gel filtration column equilibrated with the corresponding experimental buffer (see below).

All non-fluorescent proteins were quantified using absorbance at 280 nm, with extinc-

tion coefficients calculated from primary sequences (ProtParam). Fluorescent protein-tagged proteins were quantified by absorption at the excitation maxima of the specific fluorescent protein used, with previously reported extinction coefficients [85].

A.4 Phase Diagram Mapping

PolySUMO and polySIM proteins were mixed into the corresponding experimental buffer (see below) at varying concentrations in a total solution volume of 40 μ L. Mixtures were prepared directly in non-binding surface-treated, half-area 96-well plates (Fisher) and incubated for 24 hours at room temperature. The presence or absence of droplets was scored by brightfield observation of wells at 10 \times magnification using an Olympus SZX10 microscope.

A.5 Partitioning Assays

Scaffold molecules (1 % Alexa-labeled) were mixed with GFP- or RFP-tagged or Alexa-labeled clients in wells of chambered cover glass (GraceBiolabs) or 384-well plates (Sigma) passivated with 30 mg/mL BSA (Sigma). Mixtures were incubated for 2-4 hours for SH3/PRM and PTB/RNA experiments and 20-26 hours for SUMO/SIM experiments and imaged at 20x magnification.

Purified scaffold and client proteins were mixed together at desired concentrations and incubated, as described. Partitioning assays with the PRM/SH3 system were performed in 150 mM KCl; 1 mM MgCl₂; 10 mM Imidazole, pH 7.0; 1 mM EGTA; and 100 mg/ml BSA. Partitioning assays with the PTB/RNA system were performed in 100 mM NaCl; 1 mM MgCl₂; 20 mM Imidazole, pH 7.0; 50 mM Sodium Phosphate, pH 6.0; and 100 mg/ml Polyethylene Glycol 3350 (Sigma). Partitioning assays with the *in trans* SUMO/SIM system were performed in 150 mM KCl; 1 mM MgCl₂; 20 mM HEPES, pH 7.0; 1 mM EGTA; and 1 mM DTT.

For partitioning assays with the *in cis* (fused) SUMO/SIM system, scaffold and client components were gel filtered and mixed into 50 mM NaCl; 10 mM Tris, pH 8.0; and 1 mM DTT. Phase separation of SUMO/SIM scaffold proteins (which occurs robustly at pH 7.0 but not at pH 8.0) was induced by adding to this mixture a small volume (10 %) of a high-concentration stock solution of NaCl; Imidazole, pH 7.0; and DTT such that the final solution composition after mixing became 150 mM NaCl; 100 mM Imidazole, pH 7.0; 10 mM Tris, pH 8.0; and 1 mM DTT, with all proteins at the desired final concentrations. Addition of an excess of Imidazole over Tris changes the solution pH closer to 7.0, inducing phase separation. Time lapse assays with the *in cis* system also contained 150 mg/mL BSA in solution.

A.6 Image acquisition and analysis

Yeast cells were imaged using DeltaVision Elite microscope at 100x magnification using a sCMOS camera. In all other experiments, imaging was performed using spinning disk confocal microscopes equipped with EMCCD cameras at 20x or 100x magnification for *in vitro* or cellular experiments, respectively. Images were analyzed using ImageJ or MATLAB (Mathworks) (see §B). Fluorescence intensities were calibrated to concentrations using standard solutions of purified client molecules or corresponding fluorescent proteins, whose concentrations were independently determined. When possible, care was taken, to circumvent effects of the PSF in concentration determination (see §B.2-B.3).

A.7 Isothermal Titration Calorimetry

ITC assays were performed using a VP-ITC 200 calorimeter (GE Healthcare). Proteins were degassed and diluted to appropriate concentrations with degassed SUMO/SIM experimental buffer. Baseline corrections and calculations of integrated heats from the obtained thermogram were conducted using the NITPIC software [82]. Values for apparent

dissociation constants (K_d , in units of module concentration) were obtained with the Sedphat software [66] by fitting the integrated heats to a 1:1 heteromeric binding model and using experimental module concentrations of the titrant and titrand as concentrations of ligand and receptor, respectively.

A.8 Mass Action Model for Client Partitioning

Measured concentrations (from imaging) and affinities (from ITC) of polySUMO and polySIM were used to calculate the free sites concentrations in droplet and bulk phases. An equilibrium mass action model was created to describe our systems as two compartments with unequal concentrations of receptors (free scaffold sites) and a permeable ligand (client). The model was numerically solved using MATLAB and predicted PC s were calculated as ratio of the total ligand concentration between the two compartments (see §3.4.1 for details).

A.9 Fluorescence Recovery After Photobleaching

FRAP assays were performed using a Zeiss 780 laser scanning confocal microscope. A region of interest surrounding ~ 20 μm droplets was photobleached with a 488 nm Argon laser. Intensities were normalized to pre-bleach intensities of the droplet and scaled with the intensities of an unbleached control region to correct for acquisition-induced photobleaching during the recovery phase. The resulting scaled relative intensities as a function of time, $I(t)$, were fit to the single exponential model

$$I(t) = I_\infty + (I_0 - I_\infty)e^{-kt}, \tag{A.1}$$

where I_0 is the fractional intensity that remains unbleached at the start of the experiment, I_∞ the fractional recovery, and k the exponential constant. The characteristic timescale of recovery, τ , was found by calculating the reciprocal of k ,

$$\tau = 1/k. \tag{A.2}$$

A.10 Immunofluorescence Staining of Endogenous PML

U2OS cells were fixed in 4 % paraformaldehyde and stained with an anti-PML primary antibody (H-238; Santa Cruz Biotechnologies) followed by an AF647-labeled secondary antibody (Life Technologies); cells were also stained with Hoechst 33342 (Life Technologies) using the manufacturers instructions.

A.11 Western Blotting of Transfected PML

PML^{-/-} MEFs were transfected (using Lipofectamine 2000) with GFP, GFP-PML, or GFP-PML_(SUMO)⁻. Cells were lysed 20 hours post-transfection into 8 M Urea to minimize deSUMOylation in lysates. Immunoblotting was performed with antibodies against PML (Santa Cruz); GFP (Abcam); or Tubulin (Sigma) and Lamin A/C (Cell Signaling).

A.12 Viral Assays

PML^{-/-} MEFs were transfected (using Lipofectamine 2000) and infected with VSV carrying a GFP reporter, as described [122]. The percentage of transfected cells (RFP⁺) that were also infected (GFP⁺) cells was quantified by flow cytometry.

A.13 Apoptosis Assays

PML^{-/-} MEFs stably expressing tetracycline-inducible GFP-PML were treated with doxycycline for 12 hours, then treated with the apoptotic stimuli for 8 hours (except Staurosporine, 2 hours). Cells were analyzed by western blotting using an anti-Caspase3 antibody (Cell Signaling).

A.14 2D Gel Electrophoresis

PML^{-/-} MEFs stably expressing constitutive GFP or GFP-PML were submitted to the Texas A&M Protein Chemistry Lab for 2D gel analysis. Cells were lysed in 7 M Urea, 2 M Thiourea, and 4 % CHAPS, followed by methanol precipitation and rehydration. Whole cell protein samples were then subject to isoelectric focusing (pH 3-10 NL) and SDS PAGE steps using the IPGPHor (GE Healthcare) and Hoefer SE600 systems, respectively. Western blotting was performed using a mixture of anti-SUMO1 and anti-SUMO2/3 antibodies (Cell Signaling).

A.15 RNASeq

PML^{-/-} MEFs and derived lines stably expressing constitutive GFP or GFP-PML were submitted to the UTSW Genomics & Microarray Core for RNA extraction [using Trizol reagent (Thermo Fisher)], library preparation, and RNASeq (50 bp, single-end reads). Reads were mapped using CLC-Bio and expression levels were quantified as reads per (kilobase of transcript) per (million reads) (RPKM).

Appendix B

Image Analysis and Quantification of Partitioning

This appendix contains material adapted from a manuscript in press [7].

B.1 Preprocessing

Image analysis was performed using either ImageJ or custom scripts in MATLAB. In general, all images were background-corrected by subtracting dark images. Spinning disk images were also corrected for nonuniform illumination by normalizing to images of uniformly fluorescent samples prepared with $\sim 1 \mu\text{M}$ fluorophore corresponding to each channel [147].

B.2 Identification of Droplet and Bulk Regions

For *in vitro* assays, droplet and bulk regions were identified by thresholding the fluorescence intensities from the AF-labeled or fluorescent protein-tagged scaffold. To mitigate the effects of the PSF on intensities, regions within $2 \mu\text{m}$ of droplet/bulk edges were excluded from the analysis. Moreover, droplets smaller than $\approx 9 \mu\text{M}$ or larger than $\approx 28 \mu\text{M}$ were also excluded from analysis to minimize the influence of the PSF or pinhole crosstalk (in spinning disk imaging), respectively.

For cellular experiments, puncta and cyto- or nucleoplasmic regions were likewise identified by thresholding the fluorescence intensities from the transfected fluorescent protein-tagged scaffold. For P body experiments in yeast, the fluorescence signal from Edc3-GFP was used for this purpose. To circumvent the effects of the PSF on intensities of small objects, whenever possible, only structures ≥ 2 -fold larger than the axial PSF were included in the

analysis. For all PML and P body experiments, however, cellular puncta were generally smaller than the corresponding PSF cutoff. In these cases we only analyzed puncta within a preset size window (100-300 nm for PML bodies and 200-400 nm for P bodies).

B.3 Calculation of PC or IR

For *in vitro* experiments, intensities from all included regions of droplets or bulk phases were averaged and used to calculate molecule and module concentrations in the two phases using an intensity vs. concentration standard curve. Standard curve solutions were prepared alongside experimental samples using fluorescent protein-tagged or AF-labeled molecules in the corresponding experimental buffer supplemented with 0.1-0.2 mg/mL BSA to prevent fluorophore adsorption.

Intensity values in cellular experiments were likewise converted to concentrations using intensity vs. concentration standard curves. We used the average of maximum intensities of all the included puncta as well as the average intensity in the cytoplasm in each cell to obtain puncta and bulk concentrations, respectively.

For the PML and P body experiments, accurate concentration determinations were impractical due to the small size of the puncta relative to the PSF. We therefore opted to use a more convenient, albeit less accurate, surrogate metric for PC that reports the ratio of the intensity values without converting to concentrations (IR , as described). We used the average of maximum intensities of all the included puncta as well as the average intensity in the nucleo- or cytoplasm in each cell to determine this ratio.

In the experiments with endogenous PML, recruitment could not be detected for any client above a particular client expression level (assessed by mean nuclear intensity of the client), possibly due to saturation of binding sites within the bodies resulting from client overexpression. We therefore filtered out all cells with client expression levels above this *ad*

hoc cutoff (preset to be the same for all clients) and only analyzed the remaining 36 % of imaged cells.

Bibliography

- [1] M Alcalay, L Tomassoni, E Colombo, S Stoldt, F Grignani, M Fagioli, L Szekely, K Helin, and P G Pelicci. The promyelocytic leukemia gene product (PML) forms stable complexes with the retinoblastoma protein. *Molecular and cellular biology*, 18(2):1084–1093, February 1998.
- [2] Francesco Antolini, Mario Lo Bello, and Marco Sette. Purified promyelocytic leukemia coiled-coil aggregates as a tetramer displaying low alpha-helical content. *Protein expression and purification*, 29(1):94–102, May 2003.
- [3] N Asherie, J Pande, A Lomakin, O Ogun, S R Hanson, J B Smith, and G B Benedek. Oligomerization and phase separation in globular protein solutions. *Biophysical chemistry*, 75(3):213–227, December 1998.
- [4] Ferhan Ayaydin and Mary Dasso. Distinct in vivo dynamics of vertebrate SUMO paralogues. *Molecular biology of the cell*, 15(12):5208–5218, December 2004.
- [5] M M Babu, R W Kriwacki, and R V Pappu. Versatility from Protein Disorder. *Science (New York, NY)*, 337(6101):1460–1461, September 2012.
- [6] Salman F Banani, Hyun O Lee, Anthony A Hyman, and Michael K Rosen. Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology*, 18(5):285–298, May 2016.
- [7] Salman F Banani, Allyson Rice, William Peeples, Yuan Lin, Saumya Jain, Roy Parker, and Michael K Rosen. Compositional control of phase-separated cellular bodies. *Cell*, 166(3):651–663, July 2016.

- [8] Sudeep Banjade and Michael K Rosen. Phase transitions of multivalent proteins can promote clustering of membrane receptors. *eLife*, 3, 2014.
- [9] C A Belmokhtar, J Hillion, and E Ségal-Bendirdjian. Staurosporine induces apoptosis through both caspase-dependent and caspase-independent mechanisms. *Oncogene*, 20(26):3354–3362, June 2001.
- [10] R Bernardi, A Papa, and P P Pandolfi. Regulation of apoptosis by PML and the PML-NBs. *Oncogene*, 27(48):6299–6312, October 2008.
- [11] Joel Berry, Stephanie C Weber, Nilesh Vaidya, Mikko Haataja, and Clifford P Brangwynne. RNA transcription modulates phase transition-driven nuclear body assembly. *Proceedings of the National Academy of Sciences*, 112(38):E5237–45, September 2015.
- [12] Jennifer L Best, Soula Ganiatsas, Sadhana Agarwal, Austin Changou, Paolo Salomoni, Orian Shirihai, Pamela B Meluh, Pier Paolo Pandolfi, and Leonard I Zon. SUMO-1 protease-1 regulates gene transcription through PML. *Molecular cell*, 10(4):843–855, October 2002.
- [13] Danielle Blondel, Sabrina Kheddache, Xavier Lahaye, Laurent Dianoux, and Mounira K Chelbi-Alix. Resistance to rabies virus infection conferred by the PMLIV isoform. *Journal of virology*, 84(20):10719–10726, October 2010.
- [14] Sergei Boichuk, Liang Hu, Kathleen Makielski, Pier Paolo Pandolfi, and Ole V Gjoerup. Functional connection between Rad51 and PML in homology-directed repair. *PLoS ONE*, 6(10):e25814, 2011.
- [15] François-Michel Boisvert, Silvana van Koningsbruggen, Joaquín Navascués, and Angus I Lamond. The multifunctional nucleolus. *Nature Reviews Molecular Cell Biology*, 8(7):574–585, July 2007.

- [16] Peter Brand, Thorsten Lenser, and Peter Hemmerich. Assembly dynamics of PML nuclear bodies in living cells. *PMC biophysics*, 3(1):3, 2010.
- [17] Clifford P Brangwynne, Christian R Eckmann, David S Courson, Agata Rybarska, Carsten Hoegel, Jöbin Gharakhani, Frank Jülicher, and Anthony A Hyman. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science (New York, NY)*, 324(5935):1729–1732, June 2009.
- [18] Clifford P Brangwynne, Timothy J Mitchison, and Anthony A Hyman. Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proceedings of the National Academy of Sciences*, 108(11):4334–4339, March 2011.
- [19] Clifford P Brangwynne, Peter Tompa, and Rohit V Pappu. Polymer physics of intracellular phase transitions. *Nature Physics*, 11(11):899–904, November 2015.
- [20] CP Brangwynne. Soft active aggregates: mechanics, dynamics and self-assembly of liquid-like intracellular protein bodies. *Soft Matter*, 2011.
- [21] J Ross Buchan, Regina-Maria Kolaitis, J Paul Taylor, and Roy Parker. Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell*, 153(7):1461–1474, June 2013.
- [22] J Ross Buchan, Denise Muhlrads, and Roy Parker. P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. *The Journal of cell biology*, 183(3):441–455, November 2008.
- [23] J Ross Buchan and Roy Parker. Eukaryotic stress granules: the ins and outs of translation. *Molecular cell*, 36(6):932–941, December 2009.
- [24] Kathleen A Burke, Abigail M Janke, Christy L Rhine, and Nicolas L Fawzi. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Molecular cell*, 60(2):231–241, October 2015.

- [25] David N Cacace and Christine D Keating. Biocatalyzed mineralization in an aqueous two-phase system: effect of background polymers and enzyme partitioning. *Journal of Materials Chemistry B*, 1(13):1794, 2013.
- [26] Li-Heng Cai, Sergey Panyukov, and Michael Rubinstein. Mobility of Nonsticky Nanoparticles in Polymer Liquids. *Macromolecules*, 44(19):7853–7863, October 2011.
- [27] Allan D Capili and Christopher D Lima. Structure and analysis of a complex between SUMO and Ubc9 illustrates features of a conserved E2-Ubl interaction. *Journal of molecular biology*, 369(3):608–618, June 2007.
- [28] Laurent Cappadocia, Xavier H Mascle, Véronique Bourdeau, Samuel Tremblay-Belzile, Malik Chaker-Margot, Mathieu Lussier-Price, Junya Wada, Kazuyasu Sakaguchi, Muriel Aubry, Gerardo Ferbeyre, and James G Omichinski. Structural and functional characterization of the phosphorylation-dependent interaction between PML and SUMO1. *Structure (London, England : 1993)*, 23(1):126–138, January 2015.
- [29] Michele Castellana, Maxwell Z Wilson, Yifan Xu, Preeti Joshi, Ileana M Cristea, Joshua D Rabinowitz, Zemer Gitai, and Ned S Wingreen. Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nature biotechnology*, 32(10):1011–1018, October 2014.
- [30] Katerina Chalupníková, Simon Lattmann, Nives Selak, Fumiko Iwamoto, Yukio Fujiki, and Yoshikuni Nagamine. Recruitment of the RNA helicase RHAU to stress granules via a unique RNA-binding domain. *The Journal of biological chemistry*, 283(50):35186–35198, December 2008.
- [31] Che-Chang Chang, Mandar T Naik, Yen-Sung Huang, Jen-Chong Jeng, Pei-Hsin Liao, Hong-Yi Kuo, Chun-Chen Ho, Yung-Lin Hsieh, Chiou-Hong Lin, Nai-Jia Huang, Nandita M Naik, Camy C-H Kung, Shu-Yu Lin, Ruey-Hwa Chen, Kun-Sang Chang, Tai-

- Huang Huang, and Hsiu-Ming Shih. Structural and functional roles of Daxx SIM phosphorylation in SUMO paralog-selective binding and apoptosis modulation. *Molecular cell*, 42(1):62–74, April 2011.
- [32] Mounira K Chelbi-Alix, Frédérique Quignon, Luis Pelicano, Marcel H M Koken, and Hugues de Thé. Resistance to virus infection conferred by the interferon-induced promyelocytic leukemia protein. 1998.
- [33] Yi-Chun M Chen, Constantin Kappel, Joel Beaudouin, Roland Eils, and David L Spector. Live cell dynamics of promyelocytic leukemia nuclear bodies upon entry into and exit from mitosis. *Molecular biology of the cell*, 19(7):3147–3162, July 2008.
- [34] Inn Chung, Heinrich Leonhardt, and Karsten Rippe. De novo assembly of a PML nuclear subcompartment occurs through multiple pathways and induces telomere elongation. 124(Pt 21):3603–3618, November 2011.
- [35] Christine M Clemson, John N Hutchinson, Sergio A Sara, Alexander W Ensminger, Archa H Fox, Andrew Chess, and Jeanne B Lawrence. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Molecular cell*, 33(6):717–726, March 2009.
- [36] Carolyn J Decker and Roy Parker. P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harbor perspectives in biology*, 4(9):a012286, September 2012.
- [37] Carolyn J Decker, Daniela Teixeira, and Roy Parker. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *The Journal of cell biology*, 179(3):437–449, November 2007.

- [38] G Dellaire, CH Eskiw, H Dehghani, RW Ching, and DP Bazett-Jones. Mitotic accumulations of PML protein contribute to the re-establishment of PML nuclear bodies in G1. *Journal of cell science*, 119(6):1034–1042, 2006.
- [39] Graham Dellaire, Reagan W Ching, Hesam Dehghani, Ying Ren, and David P Bazett-Jones. The number of PML nuclear bodies increases in early S phase by a fission mechanism. *Journal of cell science*, 119(Pt 6):1026–1033, March 2006.
- [40] Miroslav Dundr, Michael D Hebert, Tatiana S Karpova, David Stanek, Hongzi Xu, Karl B Shpargel, U Thomas Meier, Karla M Neugebauer, A Gregory Matera, and Tom Misteli. In vivo kinetics of Cajal body components. *The Journal of cell biology*, 164(6):831–842, March 2004.
- [41] Faten El Asmi, Mohamed Ali Maroui, Jacques Dutrieux, Danielle Blondel, Sébastien Nisole, and Mounira K Chelbi-Alix. Implication of PMLIV in both intrinsic and innate immunity. *PLoS Pathogens*, 10(2):e1003975, February 2014.
- [42] Shana Elbaum-Garfinkle and Clifford P Brangwynne. Liquids, Fibers, and Gels: The Many Phases of Neurodegeneration. *Developmental Cell*, 35(5):531–532, December 2015.
- [43] Shana Elbaum-Garfinkle, Younghoon Kim, Krzysztof Szczepaniak, Carlos Chih-Hsiung Chen, Christian R Eckmann, Sua Myong, and Clifford P Brangwynne. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proceedings of the National Academy of Sciences*, May 2015.
- [44] Andrew C Elden, Hyung-Jun Kim, Michael P Hart, Alice S Chen-Plotkin, Brian S Johnson, Xiaodong Fang, Maria Armakola, Felix Geser, Robert Greene, Min Min Lu, Arun Padmanabhan, Dana Clay-Falcone, Leo McCluskey, Lauren Elman, Denise

- Juhr, Peter J Gruber, Udo Rüb, Georg Auburger, John Q Trojanowski, Virginia M-Y Lee, Vivianna M Van Deerlin, Nancy M Bonini, and Aaron D Gitler. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*, 466(7310):1069–1075, August 2010.
- [45] R D Everett, W C Earnshaw, A F Pluta, T Sternsdorf, A M Ainsztein, M Carmena, S Ruchaud, W L Hsu, and A Orr. A dynamic connection between centromeres and ND10 proteins. *Journal of cell science*, 112 (Pt 20):3443–3454, October 1999.
- [46] R D Everett, P Lomonte, T Sternsdorf, R van Driel, and A Orr. Cell cycle regulation of PML modification and ND10 composition. *Journal of cell science*, 112 (Pt 24):4581–4588, December 1999.
- [47] M Fändrich, M A Fletcher, and C M Dobson. Amyloid fibrils from muscle myoglobin. *Nature*, 410(6825):165–166, March 2001.
- [48] Marina Feric, Nilesh Vaidya, Tyler S Harmon, Diana M Mitrea, Lian Zhu, Tiffany M Richardson, Richard W Kriwacki, Rohit V Pappu, and Clifford P Brangwynne. Co-existing Liquid Phases Underlie Nucleolar Subcompartments. *Cell*, May 2016.
- [49] R Ferrari, D Kapogiannis, E D Huey, and P Momeni. FTD and ALS: a tale of two diseases. *Current Alzheimer research*, 8(3):273–294, May 2011.
- [50] Paul J Flory. Thermodynamics of High Polymer Solutions. *Journal Of Chemical Physics*, 10(1):51, 1942.
- [51] PJ Flory. Molecular Size Distribution in Three Dimensional Polymers. I. Gelation1. *Journal of the American Chemical Society*, 63(11):3083–3090, 1941.
- [52] K w Fong, Y Li, W Wang, W Ma, K Li, R Z Qi, D Liu, Z Songyang, and J Chen. Whole-genome screening identifies proteins localized to distinct nuclear bodies. *The Journal of cell biology*, 203(1):149–164, October 2013.

- [53] Cheryl T S Wong Po Foo, Ji Seok Lee, Widya Mulyasasmita, Andreina Parisi-Amon, and Sarah C Heilshorn. Two-component protein-engineered physical hydrogels for cell encapsulation. *106(52):22067–22072*, December 2009.
- [54] Simon A Fromm, Julia Kamenz, Erik R Nöldeke, Ancilla Neu, Georg Zocher, and Remco Sprangers. In vitro reconstitution of a cellular phase-transition process that involves the mRNA decapping machinery. *Angewandte Chemie (International ed. in English)*, 53(28):7354–7359, July 2014.
- [55] Ken Fujimura, Fumi Kano, and Masayuki Murata. Identification of PCBP2, a facilitator of IRES-mediated translation, as a novel constituent of stress granules and processing bodies. *RNA (New York, N.Y.)*, 14(3):425–431, March 2008.
- [56] Jaclyn R Gareau and Christopher D Lima. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nature Reviews Molecular Cell Biology*, 11(12):861–871, December 2010.
- [57] Tomáš Grousl, Pavel Ivanov, Ivana Frýdlová, Pavla Vasicová, Filip Janda, Jana Vojtová, Katerina Malínská, Ivana Malcová, Lenka Nováková, Dana Janosková, Leos Valásek, and Jirí Hasek. Robust heat shock induces eIF2alpha-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, *Saccharomyces cerevisiae*. *Journal of cell science*, 122(Pt 12):2078–2088, June 2009.
- [58] Randal Halfmann. A glass menagerie of low complexity sequences. *Current Opinion In Structural Biology*, 38:9–16, June 2016.
- [59] Tina W Han, Masato Kato, Shanhai Xie, Leeju C Wu, Hamid Mirzaei, Jimin Pei, Min Chen, Yang Xie, Jeffrey Allen, Guanghua Xiao, and Steven L McKnight. Cell-free

- formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell*, 149(4):768–779, May 2012.
- [60] Momoyo Hanazawa, Masafumi Yonetani, and Asako Sugimoto. PGL proteins self associate and bind RNPs to mediate germ granule assembly in *C. elegans*. *The Journal of cell biology*, 192(6):929–937, March 2011.
- [61] Korie E Handwerger, Jason A Cordero, and Joseph G Gall. Cajal bodies, nucleoli, and speckles in the *Xenopus* oocyte nucleus have a low-density, sponge-like structure. *Molecular biology of the cell*, 16(1):202–211, January 2005.
- [62] Neil Hattersley, Linnan Shen, Ellis G Jaffray, and Ronald T Hay. The SUMO protease SENP6 is a direct regulator of PML nuclear bodies. *Molecular biology of the cell*, 22(1):78–90, January 2011.
- [63] Xingyue He, Jessica Riceberg, Sai M Pulukuri, Steve Grossman, Vaishali Shinde, Pooja Shah, James E Brownell, Larry Dick, John Newcomb, and Neil Bence. Characterization of the loss of SUMO pathway function on cancer cells and tumor proliferation. *PLoS ONE*, 10(4):e0123882, 2015.
- [64] Ivo A Hendriks, Rochelle C J D’Souza, Bing Yang, Matty Verlaan-de Vries, Matthias Mann, and Alfred C O Vertegaal. Uncovering global SUMOylation signaling networks in a site-specific manner. *Nature Structural & Molecular Biology*, 21(10):927–936, October 2014.
- [65] T G Hofmann and H Will. Body language: the function of PML nuclear bodies in apoptosis regulation. *Cell death and differentiation*, 10(12):1290–1299, December 2003.
- [66] Jon C D Houtman, Patrick H Brown, Brent Bowden, Hiroshi Yamaguchi, Ettore Appella, Lawrence E Samelson, and Peter Schuck. Studying multisite binary and ternary

- protein interactions by global analysis of isothermal titration calorimetry data in SED-PHAT: application to adaptor protein complexes in cell signaling. *Protein science : a publication of the Protein Society*, 16(1):30–42, January 2007.
- [67] Nathaniel P Hoyle, Lydia M Castelli, Susan G Campbell, Leah E A Holmes, and Mark P Ashe. Stress-dependent relocalization of translationally primed mRNPs to cytoplasmic granules that are kinetically and spatially distinct from P-bodies. *The Journal of cell biology*, 179(1):65–74, October 2007.
- [68] Shu-Yu Huang, Mandar T Naik, Chi-Fon Chang, Pei-Ju Fang, Ying-Hui Wang, Hsiu-Ming Shih, and Tai-Huang Huang. The B-box 1 dimer of human promyelocytic leukemia protein. *Journal of Biomolecular NMR*, 60(4):275–281, October 2014.
- [69] M L Huggins. *Solutions of long chain compounds*. The Journal of chemical physics, 1941.
- [70] Anthony A Hyman. Whither systems biology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1584):3635–3637, December 2011.
- [71] Anthony A Hyman and Kai Simons. Cell biology. Beyond oil and water–phase transitions in cells. *Science (New York, NY)*, 337(6098):1047–1049, August 2012.
- [72] Anthony A Hyman, Christoph A Weber, and Frank Jülicher. Liquid-liquid phase separation in biology. *Annual review of cell and developmental biology*, 30:39–58, 2014.
- [73] A M Ishov, A G Sotnikov, D Negorev, O V Vladimirova, N Neff, T Kamitani, E T Yeh, J F Strauss, and G G Maul. PML is critical for ND10 formation and recruits the PML-interacting protein daxx to this nuclear structure when modified by SUMO-1. *The Journal of cell biology*, 147(2):221–234, October 1999.

- [74] Saumya Jain and Roy Parker. The discovery and analysis of P Bodies. *Advances in experimental medicine and biology*, 768:23–43, 2013.
- [75] Saumya Jain, Joshua R Wheeler, Robert W Walters, Anurag Agrawal, Anthony Barsic, and Roy Parker. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell*, 164(3):487–498, January 2016.
- [76] K Jensen, C Shiels, and P S Freemont. PML protein isoforms and the RBCC/TRIM motif. *Oncogene*, 20(49):7223–7233, October 2001.
- [77] Hao Jiang, Shusheng Wang, Yuejia Huang, Xiaonan He, Honggang Cui, Xueliang Zhu, and Yixian Zheng. Phase Transition of Spindle-Associated Protein Regulate Spindle Apparatus Assembly. *Cell*, 163(1):108–122, September 2015.
- [78] Mathias Jucker and Lary C Walker. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature*, 501(7465):45–51, September 2013.
- [79] Trish E Kaiser, Robert V Intine, and Miroslav Dundr. De novo formation of a subnuclear body. *Science (New York, NY)*, 322(5908):1713–1717, December 2008.
- [80] D Kalderon, B L Roberts, W D Richardson, and A E Smith. A short amino acid sequence able to specify nuclear location. *Cell*, 39(3 Pt 2):499–509, December 1984.
- [81] Masato Kato, Tina W Han, Shanhai Xie, Kevin Shi, Xinlin Du, Leeju C Wu, Hamid Mirzaei, Elizabeth J Goldsmith, Jamie Longgood, Jimin Pei, Nick V Grishin, Douglas E Frantz, Jay W Schneider, She Chen, Lin Li, Michael R Sawaya, David Eisenberg, Robert Tycko, and Steven L McKnight. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*, 149(4):753–767, May 2012.

- [82] Sandro Keller, Carolyn Vargas, Huaying Zhao, Grzegorz Piszczek, Chad A Brautigam, and Peter Schuck. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Analytical chemistry*, 84(11):5066–5073, June 2012.
- [83] Young-Eui Kim, Dong-Yeon Kim, Jang-Mi Lee, Seong-Tae Kim, Tae-Hee Han, and Jin-Hyun Ahn. Requirement of the coiled-coil domain of PML-RARalpha oncoprotein for localization, sumoylation, and inhibition of monocyte differentiation. *Biochemical and biophysical research communications*, 330(3):746–754, May 2005.
- [84] Oliver D King, Aaron D Gitler, and James Shorter. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain research*, 1462:61–80, June 2012.
- [85] Gert-Jan Kremers, Sarah G Gilbert, Paula J Cranfill, Michael W Davidson, and David W Piston. Fluorescent proteins at a glance. *Journal of cell science*, 124(Pt 2):157–160, January 2011.
- [86] Sonja Kroschwald, Shovamayee Maharana, Daniel Mateju, Liliana Malinovska, Elisabeth Nüske, Ina Poser, Doris Richter, and Simon Alberti. Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *eLife*, 4:e06807, 2015.
- [87] Imin Kwon, Masato Kato, Siheng Xiang, Leeju Wu, Pano Theodoropoulos, Hamid Mirzaei, Tina Han, Shanhai Xie, Jeffrey L Corden, and Steven L McKnight. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell*, 155(5):1049–1060, November 2013.
- [88] Valérie Lallemand-Breitenbach and Hugues de Thé. PML nuclear bodies. *Cold Spring Harbor perspectives in biology*, 2(5):a000661, May 2010.

- [89] Yun Wah Lam, Carol E Lyon, and Angus I Lamond. Large-scale isolation of Cajal bodies from HeLa cells. *Molecular biology of the cell*, 13(7):2461–2473, July 2002.
- [90] Marion Lang, Thibaud Jegou, Inn Chung, Karsten Richter, Sandra Münch, Anikó Udvarhelyi, Christoph Cremer, Peter Hemmerich, Johann Engelhardt, Stefan W Hell, and Karsten Rippe. Three-dimensional organization of promyelocytic leukemia nuclear bodies. 123(Pt 3):392–400, February 2010.
- [91] Pilog Li, Sudeep Banjade, Hui-Chun Cheng, Soyeon Kim, Baoyu Chen, Liang Guo, Marc Llaguno, Javoris V Hollingsworth, David S King, Salman F Banani, Paul S Russo, Qiu-Xing Jiang, B Tracy Nixon, and Michael K Rosen. Phase transitions in the assembly of multivalent signalling proteins. *Nature*, 483(7389):336–340, March 2012.
- [92] Shulin Li, Miles Wilkinson, Xueqing Xia, Michael David, Lihong Xu, Amy Purkel-Sutton, and Anjana Bhardwaj. Induction of IFN-regulated factors and antitumoral surveillance by transfected placebo plasmid DNA. *Molecular therapy : the journal of the American Society of Gene Therapy*, 11(1):112–119, January 2005.
- [93] Yun R Li, Oliver D King, James Shorter, and Aaron D Gitler. Stress granules as crucibles of ALS pathogenesis. *The Journal of cell biology*, 201(3):361–372, April 2013.
- [94] Ding-Yen Lin, Yen-Sung Huang, Jen-Chong Jeng, Hong-Yi Kuo, Che-Chang Chang, Ting-Ting Chao, Chun-Chen Ho, Yun-Ching Chen, Tong-Ping Lin, Hsin-I Fang, Chih-Chang Hung, Ching-Shu Suen, Ming-Jing Hwang, Kun-Sang Chang, Gerd G Maul, and Hsiu-Ming Shih. Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. *Molecular cell*, 24(3):341–354, November 2006.

- [95] Yuan Lin, David S W Protter, Michael K Rosen, and Roy Parker. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular cell*, September 2015.
- [96] Igal Louria-Hayon, Tamar Grossman, Ronit Vogt Sionov, Osnat Alsheich, Pier Paolo Pandolfi, and Ygal Haupt. The promyelocytic leukemia protein protects p53 from Mdm2-mediated inhibition and degradation. *The Journal of biological chemistry*, 278(35):33134–33141, August 2003.
- [97] Judith J Luciani, Danielle Depetris, Yves Usson, Catherine Metzler-Guillemain, Cecile Mignon-Ravix, Micheal J Mitchell, Andre Megarbane, Pierre Sarda, Huseyin Sirma, Anne Moncla, Jean Feunteun, and Marie-Genevieve Mattei. PML nuclear bodies are highly organised DNA-protein structures with a function in heterochromatin remodelling at the G2 phase. *Journal of cell science*, 119(Pt 12):2518–2531, June 2006.
- [98] Yuntao S Mao, Hongjae Sunwoo, Bin Zhang, and David L Spector. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nature cell biology*, 13(1):95–101, 2011.
- [99] Yuntao S Mao, Bin Zhang, and David L Spector. Biogenesis and function of nuclear bodies. *Trends in Genetics*, 27(8):295–306, August 2011.
- [100] Diana M Mitrea, Jaclyn A Cika, Clifford S Guy, David Ban, Priya R Banerjee, Christopher B Stanley, Amanda Nourse, Ashok A Deniz, and Richard W Kriwacki. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *eLife*, 5, 2016.
- [101] Nurul Mohamad and Mikael Bodén. The proteins of intra-nuclear bodies: a data-driven analysis of sequence, interaction and expression. *BMC Systems Biology*, 4(1):44, 2010.

- [102] Amandine Molliex, Jamshid Temirov, Jihun Lee, Maura Coughlin, Anderson P Kanagaraj, Hong Joo Kim, Tanja Mittag, and J Paul Taylor. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell*, 163(1):123–133, September 2015.
- [103] S Müller and A Dejean. Viral immediate-early proteins abrogate the modification by SUMO-1 of PML and Sp100 proteins, correlating with nuclear body disruption. *Journal of virology*, 73(6):5137–5143, June 1999.
- [104] Widya Mulyasmita, Ji Seok Lee, and Sarah C Heilshorn. Molecular-Level Engineering of Protein Physical Hydrogels for Predictive Sol–Gel Phase Behavior. *Biomacromolecules*, 12(10):3406–3411, October 2011.
- [105] D Negorev and G G Maul. Cellular proteins localized at and interacting within ND10/PML nuclear bodies/PODs suggest functions of a nuclear depot. *Oncogene*, 20(49):7234–7242, October 2001.
- [106] Sébastien Nisole, Mohamed Ali Maroui, Xavier H Mascle, Muriel Aubry, and Mounira K Chelbi-Alix. Differential Roles of PML Isoforms. *Frontiers in oncology*, 3:125, 2013.
- [107] Timothy J Nott, Evangelia Petsalaki, Patrick Farber, Dylan Jervis, Eden Fussner, Anne Plochowitz, Timothy D Craggs, David P Bazett-Jones, Tony Pawson, Julie D Forman-Kay, and Andrew J Baldwin. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Molecular cell*, 57(5):936–947, March 2015.
- [108] Chi W Pak, Martyna Kosno, Alex S Holehouse, Shae B Padrick, Anuradha Mittal, R Ali, Ali A Yunus, David R Liu, Rohit V Pappu, and Michael K Rosen. Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Model Disordered Protein. *Molecular Cell*, 63(1):72–85, 2016.

- [109] Roy Parker. RNA degradation in *Saccharomyces cerevisiae*. *Genetics*, 191(3):671–702, July 2012.
- [110] Avinash Patel, Hyun O Lee, Louise Jawerth, Shovamayee Maharana, Marcus Jahnel, Marco Y Hein, Stoyno Stoyanov, Julia Mahamid, Shambaditya Saha, Titus M Franzmann, Andrej Pozniakovski, Ina Poser, Nicola Maghelli, Loic A Royer, Martin Weigert, Eugene W Myers, Stephan Grill, David Drechsel, Anthony A Hyman, and Simon Alberti. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*, 162(5):1066–1077, August 2015.
- [111] R D Phair and T Misteli. High mobility of proteins in the mammalian cell nucleus. *Nature*, 404(6778):604–609, April 2000.
- [112] Patrick Ryan Potts and Hongtao Yu. The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. *Nature Structural & Molecular Biology*, 14(7):581–590, June 2007.
- [113] S B Prusiner. Neurodegenerative diseases and prions. *New England Journal of Medicine*, 2001.
- [114] Ivan Psakhye and Stefan Jentsch. Protein group modification and synergy in the SUMO pathway as exemplified in DNA repair. *Cell*, 151(4):807–820, November 2012.
- [115] Mani Ramaswami, J Paul Taylor, and Roy Parker. Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell*, 154(4):727–736, August 2013.
- [116] Martin A M Reijns, Ross D Alexander, Michael P Spiller, and Jean D Beggs. A role for Q/N-rich aggregation-prone regions in P-body localization. *Journal of cell science*, 121(Pt 15):2463–2472, August 2008.

- [117] David Reverter and Christopher D Lima. Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature*, 435(7042):687–692, June 2005.
- [118] David Reverter and Christopher D Lima. Structural basis for SENP2 protease interactions with SUMO precursors and conjugated substrates. *Nature Structural & Molecular Biology*, 13(12):1060–1068, December 2006.
- [119] Alejandro Rojas-Fernandez, Anna Plechanovova, Neil Hattersley, Ellis Jaffray, Michael H Tatham, and Ronald T Hay. SUMO chain-induced dimerization activates RNF4. *Molecular cell*, 53(6):880–892, March 2014.
- [120] Umut Sahin, Omar Ferhi, Marion Jeanne, Shirine Benhenda, Caroline Berthier, Florence Jollivet, Michiko Niwa-Kawakita, Orestis Faklaris, Niclas Setterblad, Hugues de Thé, and Valérie Lallemand-Breitenbach. Oxidative stress-induced assembly of PML nuclear bodies controls sumoylation of partner proteins. *The Journal of cell biology*, 204(6):931–945, March 2014.
- [121] H Saitoh and J Hinchev. Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *The Journal of biological chemistry*, 275(9):6252–6258, March 2000.
- [122] John W Schoggins, Donna A MacDuff, Naoko Imanaka, Maria D Gainey, Bimmi Shrestha, Jennifer L Eitson, Katrina B Mar, R Blake Richardson, Alexander V Ratushny, Vladimir Litvak, Rea Dabelic, Balaji Manicassamy, John D Aitchison, Alan Aderem, Richard M Elliott, Adolfo García-Sastre, Vincent Racaniello, Eric J Snijder, Wayne M Yokoyama, Michael S Diamond, Herbert W Virgin, and Charles M Rice. Pan-viral specificity of IFN-induced genes reveals new roles for cGAS in innate immunity. *Nature*, 505(7485):691–695, January 2014.

- [123] Naotaka Sekiyama, Takahisa Ikegami, Tsutomu Yamane, Mitsunori Ikeguchi, Yasuhiro Uchimura, Daichi Baba, Mariko Ariyoshi, Hidehito Tochio, Hisato Saitoh, and Masahiro Shirakawa. Structure of the small ubiquitin-like modifier (SUMO)-interacting motif of MBD1-containing chromatin-associated factor 1 bound to SUMO-3. *The Journal of biological chemistry*, 283(51):35966–35975, December 2008.
- [124] AN Semenov and M Rubinstein. Thermoreversible gelation in solutions of associative polymers. 1. Statics. *Macromolecules*, 31(4):1373–1385, 1998.
- [125] Tian Huai Shen, Hui-Kuan Lin, Pier Paolo Scaglioni, Thomas M Yung, and Pier Paolo Pandolfi. The Mechanisms of PML-Nuclear Body Formation. *Molecular cell*, 24(3):331–339, November 2006.
- [126] Sergey P Shevtsov and Miroslav Dundr. Nucleation of nuclear bodies by RNA. *Nature cell biology*, 13(2):167–U134, 2011.
- [127] Jing Song, Linda K Durrin, Thomas A Wilkinson, Theodore G Krontiris, and Yuan Chen. Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 101(40):14373–14378, October 2004.
- [128] Jing Song, Ziming Zhang, Weidong Hu, and Yuan Chen. Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. *The Journal of biological chemistry*, 280(48):40122–40129, December 2005.
- [129] P A Srere. Complexes of sequential metabolic enzymes. *Annual review of biochemistry*, 56:89–124, 1987.
- [130] M Stadler, M K Chelbi-Alix, M H Koken, L Venturini, C Lee, A Saïb, F Quignon, L Pelicano, M C Guillemin, and C Schindler. Transcriptional induction of the PML

- growth suppressor gene by interferons is mediated through an ISRE and a GAS element. *Oncogene*, 11(12):2565–2573, December 1995.
- [131] WH Stockmayer. Molecular distribution in condensation polymers. *Journal of polymer science*, 9(1):69–71, October 1952.
- [132] Christopher A Strulson, Rosalynn C Molden, Christine D Keating, and Philip C Bevilacqua. RNA catalysis through compartmentalization. *Nature chemistry*, 4(11):941–946, October 2012.
- [133] Xiaolei Su, Jonathon A Ditlev, Enfu Hui, Wenmin Xing, Sudeep Banjade, Julia Okrut, David S King, Jack Taunton, Michael K Rosen, and Ronald D Vale. Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science (New York, NY)*, April 2016.
- [134] Daniela Teixeira, Ujwal Sheth, Marco A Valencia-Sanchez, Muriel Brengues, and Roy Parker. Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA (New York, N.Y.)*, 11(4):371–382, April 2005.
- [135] Ellen Van Damme. A manually curated network of the PML nuclear body interactome reveals an important role for PML-NBs in SUMOylation dynamics. *International journal of biological sciences*, 6(1):51–67, 2010.
- [136] Peter G Vekilov. Phase transitions of folded proteins. *Soft Matter*, 6(21):5254, 2010.
- [137] Huiyuan Wang and Sarah C Heilshorn. Adaptable hydrogel networks with reversible linkages for tissue engineering. *Advanced materials (Deerfield Beach, Fla.)*, 27(25):3717–3736, July 2015.
- [138] Jennifer T Wang, Jarrett Smith, Bi-Chang Chen, Helen Schmidt, Dominique Rasoloson, Alexandre Paix, Bramwell G Lambrus, Deepika Calidas, Eric Betzig, and Geral-

- dine Seydoux. Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in *C. elegans*. *eLife*, 3:e04591, 2014.
- [139] Li Wang, Stefan L Oliver, Marvin Sommer, Jaya Rajamani, Mike Reichelt, and Ann M Arvin. Disruption of PML Nuclear Bodies Is Mediated by ORF61 SUMO-Interacting Motifs and Required for Varicella-Zoster Virus Pathogenesis in Skin. *PLoS Pathogens*, 7(8):e1002157, August 2011.
- [140] Ying Wang and Onofrio Annunziata. Liquid-liquid phase transition of protein aqueous solutions isothermally induced by protein cross-linking. *Langmuir : the ACS journal of surfaces and colloids*, 24(6):2799–2807, March 2008.
- [141] Z G Wang, D Ruggero, S Ronchetti, S Zhong, M Gaboli, R Rivi, and P P Pandolfi. PML is essential for multiple apoptotic pathways. *Nature genetics*, 20(3):266–272, November 1998.
- [142] Stephanie C Weber and Clifford P Brangwynne. Getting RNA and protein in phase. *Cell*, 149(6):1188–1191, June 2012.
- [143] Stephanie C Weber and Clifford P Brangwynne. Inverse size scaling of the nucleolus by a concentration-dependent phase transition. *Current biology : CB*, 25(5):641–646, March 2015.
- [144] Stefanie Weidtkamp-Peters, Thorsten Lenser, Dmitri Negorev, Norman Gerstner, Thomas G Hofmann, Georg Schwanitz, Christian Hoischen, Gerd Maul, Peter Dittrich, and Peter Hemmerich. Dynamics of component exchange at PML nuclear bodies. 121(Pt 16):2731–2743, August 2008.
- [145] Denis Wirtz. Particle-tracking microrheology of living cells: principles and applications. *Annual review of biophysics*, 38:301–326, 2009.

- [146] Jeffrey B Woodruff, Oliver Wueseke, Valeria Viscardi, Julia Mahamid, Stacy D Ochoa, Jakob Bunkenborg, Per O Widlund, Andrei Pozniakovsky, Esther Zanin, Shirin Bahmanyar, Andrea Zinke, Sun Hae Hong, Marcus Decker, Wolfgang Baumeister, Jens S Andersen, Karen Oegema, and Anthony A Hyman. Centrosomes. Regulated assembly of a supramolecular centrosome scaffold in vitro. *Science (New York, NY)*, 348(6236):808–812, May 2015.
- [147] Jian-Qiu Wu, Chad D McCormick, and Thomas D Pollard. Chapter 9: Counting proteins in living cells by quantitative fluorescence microscopy with internal standards. *Methods in cell biology*, 89:253–273, 2008.
- [148] Siheng Xiang, Masato Kato, Leeju C Wu, Yi Lin, Ming Ding, Yajie Zhang, Yonghao Yu, and Steven L McKnight. The LC Domain of hnRNPA2 Adopts Similar Conformations in Hydrogel Polymers, Liquid-like Droplets, and Nuclei. *Cell*, 163(4):829–839, November 2015.
- [149] Huaiying Zhang, Shana Elbaum-Garfinkle, Erin M Langdon, Nicole Taylor, Patricia Occhipinti, Andrew A Bridges, Clifford P Brangwynne, and Amy S Gladfelter. RNA Controls PolyQ Protein Phase Transitions. *Molecular cell*, 60(2):220–230, October 2015.
- [150] S Zhong, S Müller, S Ronchetti, P S Freemont, A Dejean, and P P Pandolfi. Role of SUMO-1-modified PML in nuclear body formation. *Blood*, 95(9):2748–2752, May 2000.