The Dissertation Committee for Alice Zhao certifies that this is the approved version of the following dissertation:

Cytoplasmic foci at the crossroads of artifactual science and biological function

Committee:

Edward M. Marcotte, Supervisor

Andrew D. Ellington

Yan (Jessie) Zhang

Dean R. Appling

Vishwanath R. Iyer

Cytoplasmic foci at the crossroads of artifactual science and biological function

by

Alice Zhao, B.S.Bio

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2016

Dedication

To my mother.

The simplest lessons turn out to be the most important.

And to those who (have) dedicate(d)

their lives to open-access and transparency.

Acknowledgements

With great appreciation, I acknowledge the following individuals for their contributions to various aspects of my development as a scientist and a person. This body of research could not have been accomplished without their help, to each I owe my very deep gratitude.

I would like to thank members of the Marcotte lab, former and current, for all have helped or influenced me in some way. Through the comradery of tackling science, I have learned valuable lessons from each and every one of us. In particular, I thank Traver Hart for his excellent mentorship which recruited me into graduate studies. I thank the expert adviser Zhihua Li, for his wealth of knowledge on experimental techniques. I thank Mark Tsechansky for his persistent encouragement and earnest fascination with science. I thank Jagannath Swaminathan, Dan Boutz, Fan Tu, and Anna Battenhouse for their skillful help and expertise at critical points of each project. I also thank Jeremy O'Connell, Kevin Drew, and Chris Yellman, for their kind and patient guidance. I thank the core proteomics facility staff and the sequencing facility staff, who have all consistently and reliably aided me with my questions on sample processing and analysis. I thank my committee members, Dr. Andy Ellington, Dr. Yan Jessie Zhang, Dr. Dean Appling, Dr. Vishy Iyer, and Dr. Edward Marcotte for their courageous support, motivation, and valuable inputs. I especially thank Edward Marcotte for being an exceptional role model for me and for generously providing me with an amazing environment to tinker and explore.

I further extend my thanks to the students, post-docs, and faculty from various other labs and institutes who have given me their valuable advice, materials, and time. I am also thankful for the support and diligence from past and present administrative staff of the building and of the department, and for the cheery custodial team with whom I've spent many late evenings.

Finally, I thank my friends and family for all of their love and support over the years, without which this work would not have been possible.

Cytoplasmic foci at the crossroads of artifactual science and biological function

Publication No. _____

Alice Zhao, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Edward M. Marcotte

Deciphering protein interaction and compartmentalization is crucial to understanding the molecular mechanisms that drive biological processes. Using various high throughput approaches, we have managed to score subcellular dynamic protein re-organization into supramolecular structures and map physical association networks to discover protein complexes on a proteome-wide level. However, the case by case studies of some of these novel structures and interactions reveal difficulties in interpreting their biological basis. This study offers insights into limits inherent in the molecular techniques used to investigate subcellular structures and protein interactions, describing a set of cautionary tales and critical analysis for deciphering cases of confounding data from orthogonal approaches. This study also offers a new experimental technique for high-throughput imaging assays with mammalian cell lines.

Table of Contents

List of Tables

List of Figures

Introduction

The inner workings of a cell are driven by a choreographed dance of proteins organizing into dynamically assembling and dissembling complexes. These interactions underlie essentially all biological processes within the cell, and the spatiotemporal regulation of protein expression and partnerships are linked to each protein's unique function. Thus, in order to understand biological processes, we first must elucidate the behaviors of each protein. As such, fundamental questions to address are the identities of protein complexes, their compositions and characteristics of behavior, and ultimately their functional roles within the cell.

Proteins rarely act on their own and usually require the interactions of specific partner proteins to carry out their designated roles. Large-scale biochemical and curation efforts to map physical interaction and functional interaction networks have provided broad insight into new protein functions (Wan et al. 2015, Huttlin et al. 2015, Ewing et al. 2007, Ruepp et al. 2010, Szklarczyk et al. 2015, Hein et al. 2015). Extensive repositories of these interaction datasets consist of millions of protein interactions (Oughtred et al. 2016). Although these high-throughput protein interaction assays and databases suffer from both false-positive (false observations) and false-negative (true interactions not observed) errors (von Mering et al. 2002, Bader and Hogue 2002), they serve as valuable starting points for uncovering new protein functions through physical and functional linkages. Within this manuscript, I explore case examples of putative novel protein complexes supported from such high-throughput assays. A putative complex consisting of the purine biosynthesis

pathway explored in **Chapter 3** was proposed to underlie regulation of purine biosynthesis (and, as explained later, is subject to other interpretations). Furthermore, the prediction of an uncharacterized protein, TTC4, as a novel member of an established molecular cargo trafficking complex named the retromer is explored in **Appendix C** (and, as shown later, is likely involved in other complexes).

Cellular processes are often compartmentalized within the cell (for example, in eukaryotes, within membrane-bound organelles). Therefore, the subcellular localization of particular proteins or protein complexes can give clues as to their functional roles within the cell. As such, large-scale microscopy efforts to map global protein expression, subcellular localization (Uhlen et al. 2005, Mazumder et al. 2013, Huh et al. 2003), and context-specific re-organization of proteins (Narayanaswamy, Levy, et al. 2009, Narayanaswamy, Moradi, et al. 2009, Lam et al. 2014, Tkach et al. 2012) have provided insight into elucidating particular protein functions. Many of such studies rely on automated microscopy platforms that allow for systematic screening (Glory and Murphy 2007, Murphy 2008). In **Chapter 1**, I present a novel experimental method well-suited for high-throughput cell microscopy adapted from genomic-scale microscopy screens in yeast. The adapted technique involves printing cultured human cells onto spotted cell microarrays, or cell chips, and demonstrate the capacity to determine cellular state by reporting activation states, subcellular localizations, and/or relative abundances of target proteins.

The same biophysical properties that drive protein interactions crucial for biological processes can also give rise to deleterious protein mis-interactions.

Specifically, the mis-regulation of protein expression and/or the mis-folding of proteins can lead to their unruly accumulation and aggregation, becoming toxic species that cause cell damage and which can manifest into tissue and organ destruction. Such misfolding, aggregation, and accumulation of a wide range of proteins is the hallmark of a variety of "protein conformational diseases", encompassing highly prevalent illnesses such as Alzheimer's disease (β-amyloid), Huntington's disease (huntingtin and proteins with poly-glutamine expansions), Parkinson's disease (α-synuclein), Creutzfeldt–Jakob disease (PrP^{Sc}), amyotrophic lateral sclerosis (superoxide dismutase, TDP-43, and FUS), cataracts (crystallin), cystic fibrosis (CFTR), type II diabetes (amylin), and sickle cell anemia (mutant hemoglobin). Many of these disease-associated protein aggregates form visible intracellular deposits of various forms ranging from amorphous bodies like cytoplasmic foci (Krobitsch and Lindquist 2000, Wang et al. 2007) to ordered fibrils and fiber bundles (Cerami and Peterson 1975) which can be reconstituted *in vitro* (Vilar et al. 2008, Serpell 2000, Riesner 2003). For some cases, it has been both argued that these structures cause detriment to the cell and that they form as protective measures against smaller, more toxic species (Goure et al. 2014, Ahmed et al. 2010).

Our in-house genomic-scale yeast microscopy screen identified hundreds of new assemblies of cytoplasmic foci and fibers composed of metabolic proteins (Narayanaswamy, Levy, et al. 2009). The newly discovered wide-spread assembling of yeast proteins are morphologically similar with the aforementioned disease-associated protein aggregates, forming intracellular foci or fibers which, for many

cases (such as glutamine synthetase, cytidine triphosphate synthase, and inosine-5'-monophosphate dehydrogenase 2), have also been reconstituted *in vitro* (Juda et al. 2014, Ingerson-Mahar et al. 2010, Petrovska et al. 2014). Expounded upon in **Chapter 2**, such proteinaceous assemblies of metabolic proteins commonly serve as enzymatic compartmentalization in bacteria, yet its intriguing occurrence—and widespread extent—in eukaryotes largely remains a functional mystery. Case studies have suggested functionality of particular assemblies by their potential role as enzyme storage depots and thus a means for regulating enzymatic state (Petrovska et al. 2014, Barry et al. 2014), or as cytoskeletal elements regulating cell shape (Ingerson-Mahar et al. 2010). Other studies raise the possibility that a subset of these newly discovered protein assemblies represent stress-induced endemic misfolding and aggregation (O'Connell et al. 2014). **Chapter 3** is a case study of intracellular assemblies comprising of the purine biosynthesis pathway proteins, in which I present a cautionary tale detailing how methods used to examine these intracellular foci and fibers, such as experimental design or fluorescent tagging, could sway interpretations between functionality and aggregation.

A substantial set of these fluorescence microscopy imaging-based subcellular protein localization studies rely on libraries of cells expressing proteins modified with synthetic epitopes or fluorescent tags. While the fusion of these tags with a protein of interest have revolutionized the field of protein characterization (by facilitating visualization or immunologic assays such as purification), it can also lead to artifactual "discoveries" due to their interference with the labeled protein's endogenous behavior, such as inducing severe mis-localization (Ramanathan,

Ayyavoo, and Weiner 2001, Skube, Chaverri, and Goodson 2010, Landgraf et al. 2012, Palmer and Freeman 2004, Snapp 2005). A systematic analysis of >400 human proteins showed that one-fifth of subcellular distributions of fluorescently-fused proteins disagreed with that acquired by immunofluorescence against the native protein (Stadler et al. 2013). It has also been shown for a number of foci- and fiber-forming proteins that they formed such assemblies *only* when fluorescently tagged (Swulius and Jensen 2012, Landgraf et al. 2012, Margolin 2012, Snapp 2005). Others have demonstrated that differential tagging produced differential characteristic phenotypes of foci and fibers (Bak, Cutting, and Milewski 2007, Landgraf et al. 2012), even in the case of the highly-characterized CTP synthase fibers (Gou et al. 2014). A theme within the latter parts of this manuscript highlights how fluorescent tags used for live cell imaging, as well as small epitope tags used for purification and immunofluorescent localization, are a double-edged sword. Thus, interpretations in the literature of protein characterization experiments employing tags must be carefully considered.

Chapter 1: A human cell chip for high-throughput cell imaging

1.1 Abstract

Spotted cell chips are a high-density imaging technology for determining cellular state across arrays of cells subjected to chemical or genetic perturbation. Cells are grown and treated under standard tissue culture conditions before being fixed and printed onto replicate glass slides, effectively decoupling the experimental conditions from the assay technique. Each slide is then probed using immunofluorescence or other optical reporter and assayed by automated microscopy. We show potential applications of the cell chip by assaying cultured human cells for changes in protein abundance, subcellular localization, and activation state in response to treatment by several chemical effectors. We demonstrate scalability by printing a single chip with ~4,700 discrete samples. Coupling this technology to high-throughput methods for culturing and treating cell lines could enable researchers to examine the impact of exogenous effectors on the same population of experimentally treated cells across multiple reporter targets representing a variety of molecular systems. This allows the systematic production of a highly multiplexed datasets with minimized experimental variance and at reduced reagent costs compared to alternative techniques. The ability to prepare and store chips also allows researchers to follow up on observations gleaned from initial screens with maximal repeatability.

Chapter 1 has been previously published in: Hart, T., A. Zhao, A. Garg, S. Bolusani, and E. M. Marcotte. 2009. PLoS ONE 4 (10):e7088. doi: 10.1371/journal.pone.0007088.

This work was supported by grants from the United States National Science Foundation, National Institute of Health, the Welch Foundation (F-1515), and a Packard Fellowship to Edward M. Marcotte.

1.2 Introduction

1.2.1 Current platforms for compact high-content screening

Despite enormous progress in the postgenomic era, large-scale characterization of mammalian genes remains a daunting challenge. The vast availability of libraries consisting of thousands of constituents that can be used to perturb or examine cells (examples include small molecules, siRNAs, or expression vectors containing shRNAs, sgRNAs, or ORFs) has driven the need for new techniques for researchers to utilize these tools efficiently. Classical perturbation experiments have been enabled by the creation of RNA interference (RNAi) and chemical libraries, but there exist few platforms able to conduct cell-based experiments on the scale of mammalian genomes, especially when multiple reporters are required (Michnick 2004, Albeck et al. 2006). Plate-based assays can be used for high-content screening of cell populations (MacKeigan, Murphy, and Blenis 2005) or to capture detailed cell morphology and state information (Abraham, Taylor, and Haskins 2004) – in fact a number of dedicated commercial platforms are on the market (Starkuviene and Pepperkok 2007) – but these latter applications come at a high reagent cost relative to miniaturized assays. High-throughput flow cytometry probing immunolabeled phosphoproteins (Krutzik et al. 2008, Krutzik et al. 2004) allows multiparameter sampling of protein activation state across a large cell population, but requires serial analysis of samples, hence performing sequential assays of every experimental condition or timepoint – a key limitation when performing genome-scale screens.

Transfected cell microarrays (Ziauddin and Sabatini 2001, Wheeler, Carpenter, and Sabatini 2005, Castel et al. 2006, Wu, Bailey, and Sabatini 2002), where cells are grown over a glass slide printed with transfection constructs, allows screening of expression (Ziauddin and Sabatini 2001) or RNAi libraries (Wheeler, Carpenter, and Sabatini 2005) for functional genomics or drug screening (Castel et al. 2006, Wu, Bailey, and Sabatini 2002). Transfected cell arrays have been applied to various efforts from identifying genes involved in chromosome maintenance (Neumann et al. 2006) to detecting protein-protein interactions (Fiebitz and Vanhecke 2011). For transfection-based screens, however, the technique is limited to cell lines compatible with the transfection technique used, although Sabatini and colleagues created arrays based on lentiviral infection constructs to circumvent this problem (Bailey et al. 2006). Tissue microarray technology has enabled the multiplexed immunohistochemical analysis of tissue samples on a single array but thus far has been limited to tens or hundreds of samples per array (Camp, Neumeister, and Rimm 2008). Complementary technologies include small molecule microarrays which involves growing cells over a glass slide printed with molecules such as proteins. This has been previously used to measure the response of neural precursor cells to a variety of printed signaling molecules (Soen et al. 2006). Another similar technology is the cell lysate microarray, in which the protein repertoire of a number of cell populations are spotted in parallel on a slide and assayed for reporters of cellular state (Paweletz et al. 2001, Sevecka and MacBeath 2006), although all data are population averages rather than per-cell readouts. Thus, there is a clear role for a platform that enables analysis of multiple cell types and/or treatment

conditions in a manner that scales to thousands of samples, while minimizing reagent cost and experimental variance.

My contributions to the work described in this Chapter include conducting experiments and analyzing data, including performing cell culture, drug treatments, western blotting, printing the cell chip and conducting high-throughput microscopy, with guidance from Traver Hart. Traver designed the experiments and developed analysis techniques for the cell chip, and was absolutely instrumental in overcoming every technical challenge. I acknowledge Vishy Iyer for custom building of the microarray printer and Ankit Garg who worked alongside me.

1.2.2 Inspirations from yeast cell microarrays

Previously, we have used a spotted cell microarray technique to identify pheromone response genes (Narayanaswamy, Moradi, et al. 2009) and to examine changes in protein localization (Narayanaswamy, Levy, et al. 2009) of the GFP-tagged yeast strain collection (Huh et al. 2003). Building on the experience gained from applying cell chip technology to yeast functional genomic screens, we set out to develop a human cell chip: microsamples of cultured human cells printed in arrays onto glass slides. The human cell chip would allow for high-dimensional assays consisting of multiplexed cell lines, cell types, treatment conditions, and time-points on a single slide, as well as data integration across multiple cell chips. Given these features, we believe the cell chip provides a general platform for performing human cell functional genomic assays.

1.2.3 Advantages of the new platform of human cell microarrays

Here we describe the application of spotted cell microarrays to the study of human cell lines. Spotted cell microarrays, hereafter referred to as "cell chips," are constructed by growing and treating cells under normal tissue culture conditions, formaldehyde fixing, and printing microsamples of each culture onto replicate glass slides. Each slide is then assayed by immunofluorescence against a specific target and imaged by high-throughput microscopy. Entire collections of cells comprising hundreds to thousands of discrete samples can be assayed onto replicate slides. Each slide is probed with a single reporter in a single assay, reducing experimental variance compared to multi-well plate assays where each well is effectively an independent experiment. Reagent cost is similarly reduced, with less than 100 µL of diluted antibody sufficient to probe a slide, 10–100 fold less than that required for a single 96-well plate assay. Importantly, cells from a wide variety of cell types (including both suspension and adherent cells), growth conditions, and treatments, can be arrayed on a single slide. Replicate slides allow researchers to conduct multiple assays against samples drawn from the same collection of treated cells, and to probe multiple pathways elements from the same sample of cells. Finally, slides can be stored after printing to facilitate generating replicates and following up on observations gleaned from initial screens with subsequent assays against samples drawn from the same original population of cells.

1.3 Development of the cell printing process

To print human cell lines, we used custom microarray pins with blunt tips and wide slots, and after experimenting with other adhesion protocols, we determined that printing biotin-decorated cells on streptavidin-coated glass slides ensured cell adhesion and reproducibility. An overview of the process is shown in **Figure 1-1**. To demonstrate achievable array densities, we printed eight replicate HeLa cultures repeatedly onto a slide. A total of 4,608 spots were successfully printed on a single slide (**Figure 1-1B**), using eight spotting pins and a spot pitch of 400 µm. Chips of much higher density, exceeding 8,000 spots per slide, could be achieved by decreasing spot pitch ~10% and increasing to 12 spotting pins.

1.3.1 Adapting microarray printing components

Contact microarray technology, typically used to print DNA oligonucleotides or cDNA sequences for use in RNA hybridization assays, is optimized around printing the smallest spots that can be consistently delivered. Spotted cell microarrays were first developed and applied to functional genomic screens in Saccharomyces cerevisiae (Narayanaswamy et al. 2006) and bacteria (Zhao et al. 2008). To print yeast cell chips, we used a contact microarray printing robot to draw a microsample from a suspension of fixed cells in a 96-well microplate (the "source plate") and deposit it on a poly-L-lysine coated glass slide. During the development of the yeast spotted cell microarray technique it was observed that better performance was achieved using microarray pins that had been "blunted" by repeated use in printing cDNA arrays. The blunted pins gave a larger spot size, a greater volume of medium deposited and,

typically, a larger number of cells in the spot. However, the degree of blunting and therefore the quality of spots delivered varied widely among these well-used pins.

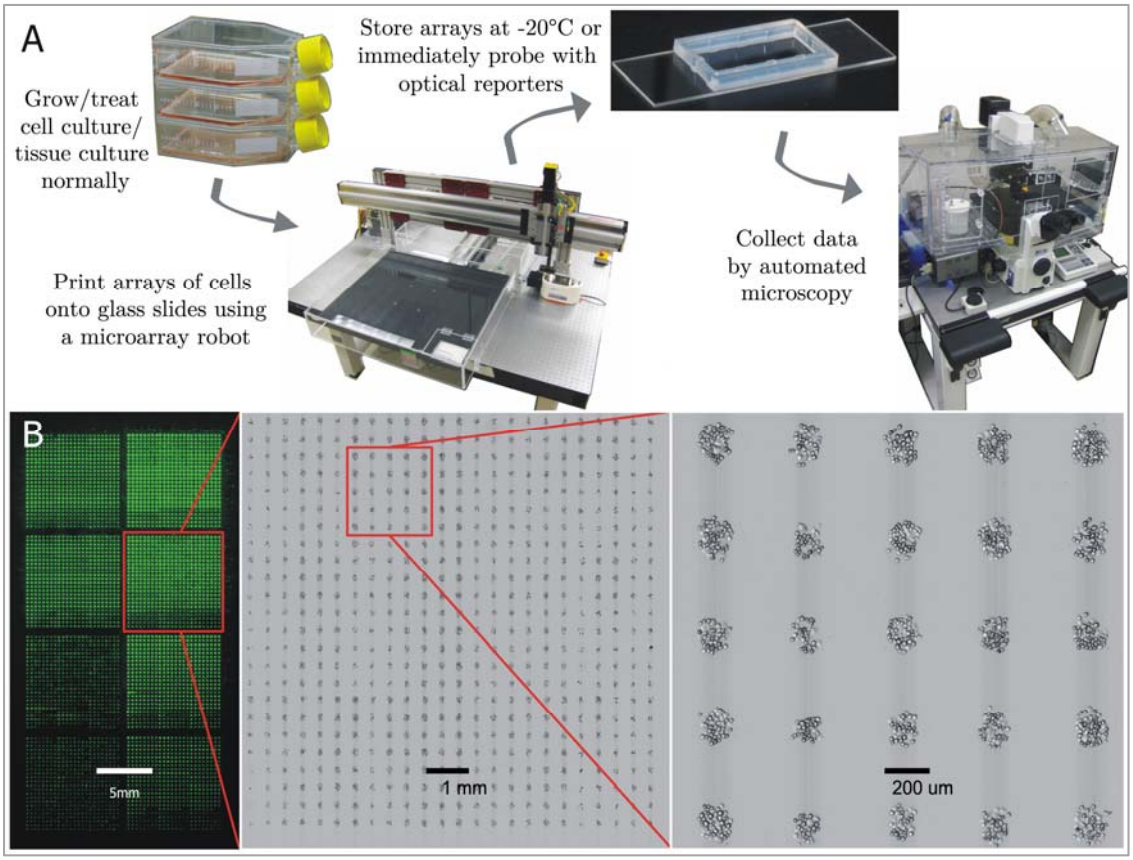

Figure 1-1. Overview of spotted cell chip process.

(**A**) Cells are grown and treated under normal cell culture conditions. Our recommended protocol (less successful early protocols are discussed in the text and in **Figure 1-2**) involves trypsinizing adherent cells and fixing with formaldehyde, decorating with WGA-biotin, and permeabilizing in −20°C methanol. Cells can be stored for several weeks in this state before resuspending in PBS and transferring to source plate for printing. Using a robotic microarray spotting device, cells are printed onto streptavidin-coated slides. To assay, each slide is probed by immunofluorescence against the target of choice and imaged by automated microscopy. (**B**) A high-density cell chip. An 8-pin print of 4,608 replicate spots, each containing a microsample of HeLa cells, is shown by imaging with a microarray scanner (left; green signal is light scattering in the fluorescent channel off freshly printed spots) and by stitching together multiple 10x microscope images (center and right).

To adapt cell chips to human cells, we initially used the same microarray pins as in the yeast cell chip, and printed on poly-L-lysine (poly-K) coated slides. Early testing was conducted using the Jurkat T-cell leukemia cell line, as these cells are easy to grow in large quantities and a successful cell chip would provide a new platform for assaying suspension cells. We immediately observed that the larger human cells – which are typically spheroids 10–20 µm in diameter, many times larger than ovoid yeast cells that measure 3–8 µm on the long axis – did not print consistently onto poly-K coated slides, and that the inconsistency was in part attributable to how deformed the microarray pins were. To address this issue in a more systematic manner we acquired microarray pins with sharp or blunt tips in different sizes (Majer Precision MicroQuill 2000, part nos. 11077-1 and 11077-3). The 11077-1 pins were sharp and yielded spots <100 µm in diameter, while the -3 pins had the largest blunt area and gave spots ~200 µm across. Number of cells deposited per contact was further improved by using custom pins, based on the 11077-3 form factor, but with a slot width of 0.003 inch (76 µm) vs. the standard 0.0015 inch (38 µm). The smaller slot is only 2–3 cell diameters in width and may have induced shear effects and clumping as cells were loaded and deposited by the pins; these effects appear to have been largely mitigated by using the wider slots. The custom 11077-3 pin with 0.003 inch slot width consistently delivers a spot ~200 µm in diameter and was used for all subsequent human cell chip prints.

Although we achieved regularity in spot sizes by selecting the appropriate microarray pins, the number of spots delivered was found to be highly dependent on the concentration of cells in the 384-well source plate. Depositing 50 cells in a spot

~1 nl in volume implies a concentration of ~50,000 cells/µl, or 10^6 cells in 20 µl suspension in each well of the source plate. However, during the time required to print ~100 samples onto each of 10–20 slides – roughly 30 minutes – the cell suspension settles into a loose pellet at the bottom of the well. In an effort to maintain the cells in suspension during printing, we increased the viscosity of print media using glycerol (15–50%) and sucrose (30–50%).

1.3.2 Proof of concept: detecting cellular states with high density cell chip-based immunofluorescence

We tested the cell chip's ability to detect cellular state by inducing apoptosis in Jurkat cells. We grew the cells under normal tissue culture conditions. Separate cultures were treated with staurosporine, a potent inhibitor of protein kinase C and other essential cellular kinases, and fixed with formaldehyde after 1, 2, or 4 hours. Treated and untreated cells were collected in several wells of a 384-well plate at a concentration of $>10^5$ cells/µl and printed on poly-L-lysine coated slides such that each sample was printed several times on each of several replicate slides.

Immediately after printing, slides were imaged with transmitted light to analyze print quality; printed spots were discrete and typically contained 20–50 cells. Three slides were then probed for signs of apoptosis by immunofluorescence with antibodies against cleaved caspase 3, cleaved caspase 9, and cleaved PARP. Each slide was also labeled with a nuclear stain, and each spot was imaged using

automated microscopy. Images of Jurkat cells immediately after printing, and of labeled cells after probing for cleaved caspase 3, are shown in **Figure 1-2**.

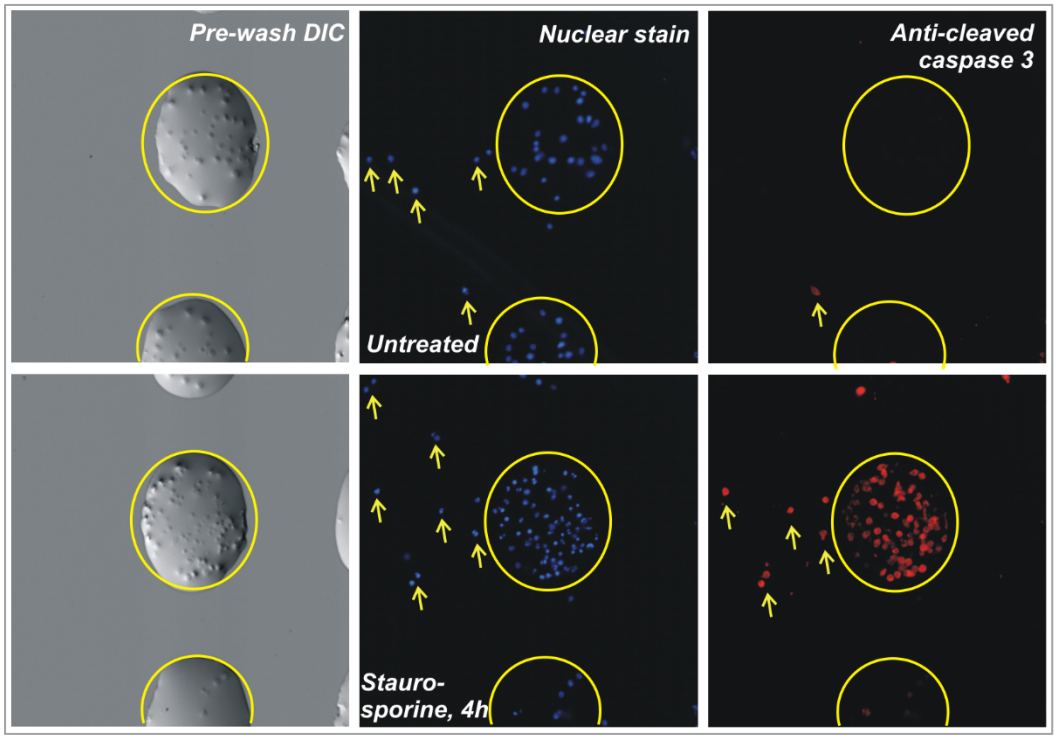

Figure 1-2. Early print of Jurkat cells onto poly-L-lysine slides.

A 30% sucrose print buffer prevented complete liquid evaporation after printing. Other panels show nuclear stain (center) and immunofluorescence (right) against cleaved caspase 3, an indicator of apoptosis. Left panels: DIC images acquired immediately after printing, before commencing immunofluorescence assay. Top row, untreated; bottom row, treated with staurosporine, 4 h. Yellow circles indicate one printed spot; arrows indicate cells outside the circle translocated during the immunofluorescence protocol. As a consequence, all subsequent prints were conducted with WGA-biotin-decorated cells printed on streptavidin-coated slides to ensure cell immobilization.

1.3.3 Troubleshooting and optimization to increase robustness against assay-related mechanical insults

Although the immunofluorescence data supported the prototype cell chip's ability to detect cellular state, we observed that a significant number of cells – perhaps 10% of the cells in some spots — had shifted on the slide during the wash steps of the immunofluorescence protocol. This translocation is evident in **Figure 1-2** when the pre-probe DIC images are compared to the fluorescent images (see arrows in figure). Given the relatively small numbers of cells in each spot, cross-contamination of even individual cells could dramatically reduce the dynamic range of the cell chip as an assay tool. To alleviate this problem, we tested an alternate adherence technique involving an adaptor molecule instead of relying on electrostatic interaction. After fixation, we decorated cells with a biotinylated lectin, wheat germ agglutinin (WGA-biotin), and printed the cells on streptavidin-coated slides. Under this protocol, increased print buffer viscosity is not required; cells were resuspended at 10^6 cells in 20 µl PBS (without glycerol or sucrose) in each well of the 384-well source plate and allowed to settle into a loose pellet. The microarray robot was calibrated to dip the pins into the pellet during loading. We printed WGA-biotin-labeled Jurkat and DG-75 suspension cells as well as trypsinized HeLa and HEK293 adherent cells onto replicate chips. Using the wide-slot pins and a standard wash cycle between loads, we observed neither cell clumping in the pins nor cross-contamination of cells into adjacent spots. After printing on a streptavidin-coated slide and allowing the print to dry, we observed no cell translocation throughout many repeated washing steps. The WGA-biotin/streptavidin slide combination was used for all subsequent prints.

1.4 Demonstration of cell chip multiplexing and utility

To demonstrate the multiplex capability of the cell chip, we printed chips with both A549 non-small-cell lung cancer cells and HeLa cervical cancer cells. Each cell line was divided into three cultures: one treated with anisomycin (1 µM, 30 minutes), one with tumor necrosis factor alpha (TNFα; 10 ng/ml, 60 minutes), and one untreated control. Anisomycin, a translation inhibitor, activates (by phosphorylation) both the p38 and c-Jun N-terminal kinase (JNK) stress kinases. Among the effects of TNFα exposure are JNK activation and NFκB translocation to the nucleus. NFκB is maximally concentrated in the nucleus at about an hour after TNFα exposure (Nelson et al. 2004), while JNK activation peaks after about 15 minutes and degrades to background levels about half an hour later (Reuther-Madrid et al. 2002). Multiple replicate chips were printed, each carrying all six conditions printed in multiple replicate spots.

Individual chips were probed for phospho-p38 kinase, phospho-JNK, and the p65/RelA subunit of NFκB. Each slide was counterstained with Hoechst 33342 nucleic acid stain and a high resolution image of each spot was captured in the corresponding fluorescent wavelengths. **Figure 1-3A** illustrates representative nuclear stain and immunofluorescence images from the chip probed for phospho-p38; the two spots show the increase in signal in HeLa cells treated with anisomycin compared to controls. The translocation of NFκB-p65 subunit into the nucleus in response to TNFα in both cell lines was apparent in the images (**Figure 1-3B** shows p65 translocation in HeLa cells).

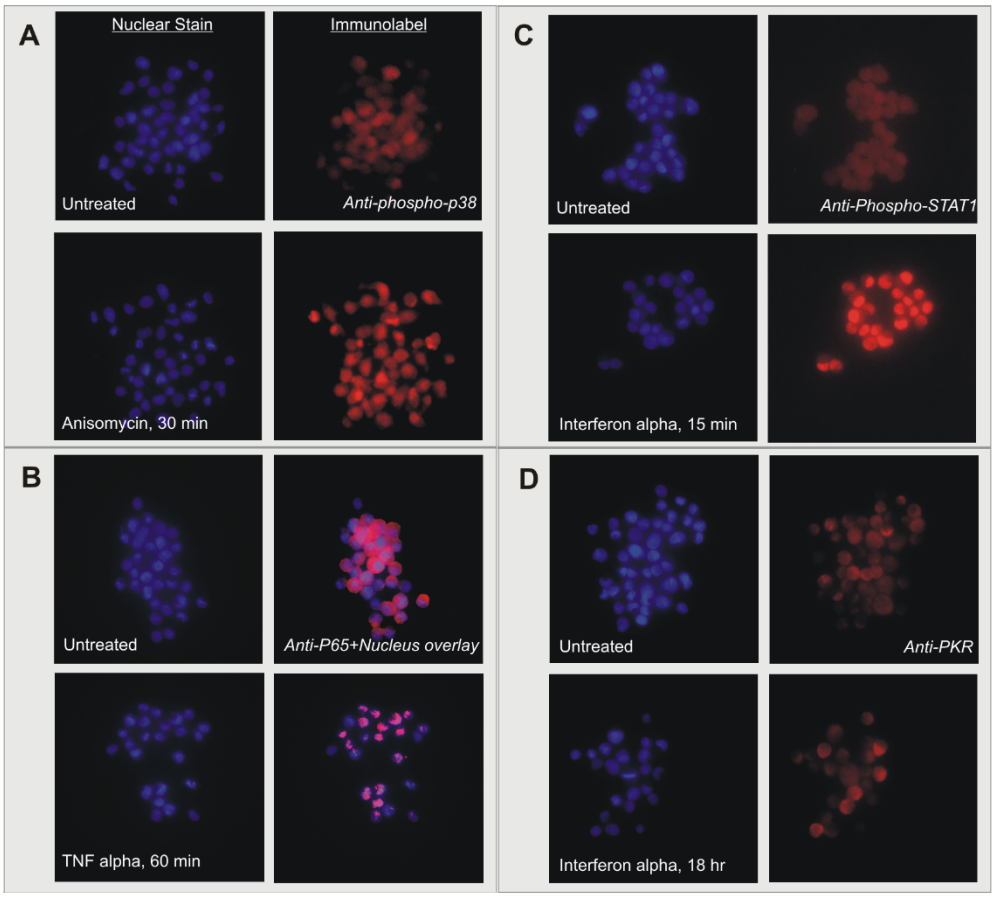

Figure 1-3. Cell chips can be printed with multiple cell types, treatments, and time points on the same arrays, enabling extensive multiplexing of experiments.

HeLa and A549 cells were grown in T-75 flasks collected after treatment with anisomycin, TNFα, or no treatment, and printed on replicate slides. (**A**) A slide was probed for phosphorylated p38 kinase; images show Hoechst 33342 nuclear stain (left) and IF spots of treated and untreated HeLa cells. (**B**) A slide from the HeLa/A549 print was probed for the p65/RelA subunit of NFκB. Right panels show overlay of brightest pixels from IF images onto nuclear stain, demonstrating cytoplasmic localization in untreated cells and nuclear translocation in response to TNFα. (**C**) A549 cells were grown in T-75 flasks and collected without treatment and after treatment with interferon-alpha (1000 U/ml) for 15 minutes or 18 hours. A chip was probed for phospho-STAT1; sample images of one 15-minute timepoint and one untreated spot show strong difference in signal. (**D**) A slide from the same print was probed for PKR; a weak signal is detected in the 18 hour timepoint, which corresponds to a ~3-fold increase in protein level as detected by western blotting (data not shown).

We analyzed the set of treated spots from each cell line for an increase in signal relative to that cell type's control spots on the same slide by comparing the set of mean bias-corrected signal intensities of each set of spots (two-sample one-tailed T-test; see Methods). Calculated p-values are shown in **Table 1-1**. JNK was phosphorylated in response to anisomycin treatment in both cell lines but TNFα-treated cells showed a weak response only in HeLa, consistent with the expected dynamics of TNF-induced JNK activation and deactivation. Anisomycin also activated p38 in HeLa cells, as expected, but surprisingly the response was much weaker in A549s; the p-value of 0.01 is not significant after multiple-hypothesis correction. P65 translocation to the nucleus is represented as an increase in nuclear signal in HeLa cells.

Probe	A549								HeLa							
	Control		Anisomycin			TNF-alpha			Control		Anisomycin			TNF-alpha		
	<I>	SE	<I>	SE	Pval	<I>	SE	Pval	<I>	SE	<I>	SE	Pval	<I>	SE	Pval
p-P38	0 (n=10)	397	491 (n=5)	294	0.01	252 (n=4)	384	0.13	0 (n=10)	225	2256 (n=5)	679	$<10^{-3}$	234 (n=5)	414	0.15
p-JNK	0 (n=10)	241	3897 (n=5)	1561	0.002	-292 (n=5)	804	0.77	0 (n=10)	374	2701 (n=5)	1119	0.002	605 (n=5)	280	0.003
P65	0 (n=8)	356	-505 (n=5)	725	0.90	-76 (n=4)	700	0.58	0 (n=9)	359	-1051 (n=5)	1151	0.95	969 (n=5)	423	0.002

Table 1-1. Drug response measured from printed cells.

Cellular response to drug treatments (anisomycin, 1 µM, 30 min; TNFα, 10 ng/ml, 60 min) on A549 and HeLa cells printed on the same chips. <I>, bias-corrected mean signal intensity (see Methods) for a given condition, with number of spots of that condition on the chip. SE, standard error of mean signal intensity for a given condition. Pval, p-value of difference between condition <I> and control <I>, measured by one-tailed, two-sample T-test.

To explore the utility of cell chip technology for pathway analysis, we examined the chip's ability to recapitulate the interferon response of A549 cells. Exposure to interferon activates the JAK/STAT signal transduction cascade, resulting in up-regulation of interferon response genes, including dsRNA-activated protein kinase (PKR), the 2'-5'-oligoadenylate synthetases, and the myxovirus resistance gene Mx (Garcia-Sastre and Biron 2006). We chose two assay targets, PKR and phospho-STAT1, to further validate the accuracy of the cell chip technology and to explore its dynamic range. We grew cells and exposed them to IFN-α (1000 U/ml) for 15 minutes or 18 hours before trypsinizing and formaldehyde-fixing, along with untreated control cells. Technical and biological repeats were printed on the same slide ('print 1'). At the same time, an equal number of cultures were prepared but stored in -20°C methanol for seven weeks before printing in an identical manner ('print 2'). After printing, one slide from each print was immunoprobed for phosphorylated-STAT1, counterstained with nuclear stain, and each spot was imaged at 40X. A second slide from each print was probed for PKR.

Figure 1-3C shows representative nuclear stain and immunofluorescence images of an individual spot from a slide probed for phospho-STAT1. Results of quantitative analysis are shown in **Table 1-2**. A fifteen-minute interferon exposure gave very strong signal for both prints, as well as weaker signal after 18 hours, indicating no loss of signal due to storage of fixed cells prior to printing. A third chip was probed after 30 days of storage at 4°C; it showed less overall signal strength across all spots but also less variance, resulting in p-values nearly identical to the other two chips. The slides probed for PKR showed a small increase in signal at the

18-hour timepoint (**Figure 1-3D; Table 1-2**), but, as expected, no response in the 15-minute samples. The small increase probably reflects the fact that PKR under these conditions is only up-regulated ~3-fold (measured by western blot; data not shown). This low relative signal may bound the sensitivity of the current state of this technology.

Probe	Slide	Control		Interferon-alpha (1000U/mL)					
				15 min			18 hr		
		<I>	SE	<I>	SE	Pval	<I>	SE	Pval
p-STAT1	1	0 (n=20)	47	860 (n=10)	278	$<10^{-5}$	115 (n=9)	143	0.02
	2	0 (n=19)	102	3100 (n=10)	1056	$<10^{-5}$	153 (n=10)	71	$<10^{-4}$
	2*	0 (n=19)	64	423 (n=10)	118	$<10^{-6}$	88 (n=10)	57	$<10^{-3}$
PKR	1	0 (n=17)	246	83 (n=9)	232	0.2	620 (n=8)	414	0.002
	2	0 (n=20)	278	156 (n=10)	148	0.03	416 (n=10)	228	$<10^{-3}$

Table 1-2. Cellular response to interferon treatment and signal variance due to experimental methods.

P-values of each condition are shown. Slide 1: Normal cell preparation followed by immediate printing and assay. Slide 2: Cells were fixed and stored for seven weeks at -20°C before printing. Slide 2*: Printed cell chip was stored for 30 days at 4°C before assaying. For column headers, see legend for **Table 1-1**.

The presence of multiple cell lines and treatment conditions on the same slide can be exploited as internal controls for both experimental conditions and probes. In the anisomycin/TNFα chips, for example, p38 kinase showed lower response to anisomycin treatment in A549 cells than in HeLa. The anti-phospho-p38 antibody

gave the expected response for the HeLa cells on the same slide, which serves as a positive control for the probe. JNK kinase responded to anisomycin in both the A549 and HeLa cells, which are drawn from the same population as those probed for phospho-p38, indicating the drug treatment worked properly. Therefore it is reasonable to conclude that the observed difference in p38 activation reflects a biological phenomenon rather than an experimental artifact. This property of multiplex controls can in principle be applied to much larger screens including a wider variety of experimental conditions and probes.

1.5 Discussion and Remarks

We have demonstrated the capability of the cell chip to probe multiple aspects of cellular state using a variety of cell types and treatment conditions. Since cells are grown and treated under standard tissue culture conditions, treatment protocols such as transfections or drug exposure times can be optimized for each set of samples individually without affecting the assay. Furthermore, the adaptor molecule used to bind cells to the slide, WGA-biotin, targets a broad spectrum of human and other cell lines, but even this step could be optimized on a per-cell-line basis by using a specific biotinylated lectin. Also, the ability to store cells prior to printing allows researchers to perform large library transfections or other treatments asynchronously rather than all at once immediately before printing. Finally, although we analyzed images to gather population data across the cells in a printed spot, it is clear that single-cell data could be gleaned using more sophisticated image processing techniques (Carpenter et al. 2006, Glory and Murphy 2007). For transfection

experiments, adding optical reporters (e.g., green fluorescent protein) on the same expression vector as the clone of interest or by co-transfection allows the measurement of cellular response exclusively on successfully transfected cells (Bailey et al. 2006), mitigating the signal loss encountered when low transfection efficiency is averaged over a population.

The cell chip is complementary to, and in some cases an advance over, current high-throughput cell-based assay technologies. It differs from transfected cell arrays in that it allows the analysis of multiple cell types and multiple growth and treatment conditions on a single slide, and it offers an order of magnitude increase in sample density over existing tissue microarray technology. Finally, in probing samples from diverse populations for a single reporter, the cell chip represents an orthogonal assay to the single-population, many-reporter gene expression DNA microarray.

1.5.1 Constraints of the human cell chip

A key constraint of the technology as described here is the considerable manual effort required to prepare a source plate for printing. While we consider it a major advantage of the cell chip that cells are grown under normal tissue culture conditions, pin-based printing requires very high cell density in the source plate, which is not easily achieved by automated cell handling techniques. The use of other printing technologies, for example inkjet or other microspray methods, might ease this burden and make microplate-based cell culture growth and treatments more compatible with cell chip printing.

1.5.2 Applicability of the technology across fields of study

Such development to overcome the key constraints of the cell chip technology could make it readily applicable to functional genomics and chemogenomics. With the ability to probe an array of cells for target protein abundance, activation state, and subcellular localization, libraries of small molecule effectors could be screened for their impact on a variety of cellular systems. Furthermore, the technology could in principle be adapted for on-chip fluorescence *in situ* hybridization (FISH) assays against nucleic acid targets. The fact that multiple dosages and time-points can be printed on the same set of cell chips increases the depth to which researchers can investigate the impact of chemical libraries. Likewise, whereas typical functional genetic screens are designed around a single reporter or phenotype, the cell chip allows a different reporter for each replicate slide. Thus, each genetic perturbation could be assayed for impact on multiple cellular subsystems and/or for multiple reporters within the same system, greatly multiplying the data "bang" for the experimental "buck." Taking into account all these features, we believe the cell chip offers a useful and general approach for medium- to large-scale cell-based assays.

1.6 Materials and Methods

1.6.1 Cell culture

A549, HEK293, HeLa, DG-75, and Jurkat cell lines were acquired from ATCC. HeLa and HEK293 cells were maintained in DMEM medium, A549 in F12-K, and DG-75 and Jurkat in RPMI, each supplemented with 10% FBS (Gibco/Invitrogen). Cells were treated in culture flasks with staurosporine (Sigma), TNF-alpha (Sigma), anisomycin (Sigma), or interferon-alpha, as described.

1.6.2 Printing cell arrays

To prepare cells for printing, adherent cell lines were washed with PBS, trypsinized (0.25% trypsin in EDTA, Invitrogen) until detached, and resuspended in PBS. Adherent and suspension cells were fixed in ¼ vol 10% fresh paraformaldehyde (Sigma) (~2% final concentration) for 10 min, then washed in fresh PBS and transferred to 1.5 ml tubes for ease of handling.

After fixation, cells were treated with biotinylated wheat germ agglutinin (WGA-biotin; Biomeda) at 4 µg/ml final concentration for 15 minutes, then washed 3x in PBS. Finally, cells were pelleted, media was removed, and cells were resuspended in 100% methanol at −20°C. Cells were incubated at least 10 minutes, and can be stored at this stage for at least several weeks at −20°C.

Immediately prior to printing, cells were pelleted, methanol was removed, and cells were resuspended in a minimum volume of PBS. A 20–50 µl of high-density cell suspension was transferred to a 384-well plate. Cells were allowed to settle to the bottom of the well, forming a loose pellet, before printing was initiated. Cells were printed on streptavidin-coated slides (TeleChem) with a custom-built DNA microarray printing robot (DeRisi, Iyer, and Brown 1997) using blunt-tipped, slotted steel custom microarray pins (Majer Precision, part no. 11077-3 with a custom 0.003 inch slot width). Printed slides were assayed immediately or stored at 4° or −20°C for several weeks.

1.6.3 Assaying printed slides

1.6.3.1 Immunofluorescence and imaging

Immunofluorescence labeling of target proteins was performed using antibodies against cleaved caspase 3, phospho-STAT1, phospho-JNK, phospho-p38, cleaved PARP, phospho-p65 (RelA) (Cell Signaling Technologies), p65/RelA, and histone H3 (Santa Cruz Biotechnology). First, a 16-well rubber gasket (Grace Biosystems) was trimmed to provide a single large reservoir on the slide around the array. The cells were blocked in 5% goat serum (Sigma) in PBS (30 min, ambient temperature), washed once with PBS, and incubated 2 hr (ambient temperature) or overnight (4°C) with primary antibody diluted to manufacturer's specification in PBS + 0.2% Triton X-100. Slides were washed in PBS 3x for 5 min each in coplin staining jars, then incubated with goat anti-rabbit (or anti-mouse, as necessary) IgG conjugated with Alexa Fluor 594 (Invitrogen; diluted 1:1000) for 60 min at ambient

temperature in the dark. Slides were washed again 3x in PBS, with the final wash containing a 1:10,000 dilution of Hoechst 33342 (Invitrogen). Slides were air-dried, mounted with ProLong Gold mounting medium and coverslipped before imaging.

Imaging was performed on a Nikon TE2000 microscope with motorized XY stage and Z objective and a Photometrics Cascade II 16-bit CCD camera. Using the NIS Elements controller software, we generated a script that would automatically visit each spot, autofocus (in DIC), and capture fluorescent images for Hoechst 33342 (nuclear) and immunostained labels. Imaging was performed using a 40x (0.95 NA) dry objective, and exposure times were selected to minimize the occurrence of saturated pixels under normal assay conditions.

1.6.3.2 Image signal quantitation

Quantitation was carried out using Matlab and the Image Processing Toolkit. For each spot, depending on the region of interest (nuclear or whole-cell) for the given probe, either the image of the nuclear label or the one of the immunolabel was background corrected and converted to a binary pixel mask. The mean signal intensity in the immunofluorescence channel of all pixels within the mask is calculated and recorded for each spot, along with the number of pixels in the mask.

We discovered a pronounced linear bias in average pixel intensity that correlated with the number of pixels in the mask – which itself reflects the number of cells deposited in the spot. A correction for the size bias was applied by

normalizing signal to that found with wildtype (untreated) cells. A linear regression is applied to signal derived from wildtype cells, and the signal of a given spot is then measured as a distance from the regression line at the same mask size of that spot. This method is effective as long as the range of wildtype mask sizes meets or exceeds the range of mask sizes for treated cells; this was true for all cases discussed here.

The bias-corrected mean signal intensity was determined for each spot of treated cells. These were compared to spots of untreated cells of the same type on that slide, yielding a distinct set of data for each cell type, even where multiple cell types are arrayed on a slide (bias-corrected mean signal intensity data are available online (Hart et al. 2009)). Treated and untreated spots were compared by unpaired one-tailed T-test for samples with different variances. By definition, the bias-corrected mean sample intensity of control spots is zero.

Chapter 2: Reorganization of metabolic enzymes into megastructures and compartments

2.1 Abstract

Large-scale imaging-based studies have revealed that hundreds of metabolic enzymes across diverse organisms form large intracellular megastructures, or bodies. These proteinaceous bodies—such as those formed by enzymes of nitrogen and carbon utilization and of nucleotide biosynthesis—range in form from intracellular foci and fibers in eukaryotes to high-density packings inside bacterial micro-compartments. Although many enzymes clearly form functional mega-assemblies, it is not yet clear for many recently discovered cases whether they represent functional entities, storage bodies, or aggregates. In this Chapter, we survey intracellular protein bodies formed by metabolic enzymes, asking when and why such bodies form and what their formation implies for the functionality—and dysfunctionality—of the enzymes that comprise them. The panoply of intracellular protein bodies also raises interesting questions regarding their evolution and maintenance within cells. We speculate on models for how such structures form in the first place and why they may be inevitable.

Parts of Chapter 2 have been previously published in: O'Connell, J. D., A. Zhao, A. D. Ellington, and E. M. Marcotte. 2012. Annu Rev Cell Dev Biol 28:89-111. doi: 10.1146/annurev-cellbio-101011-155841.

This work was supported by grants from the United States National Science Foundation, National Institutes of Health, US Army Research (58343-MA), Cancer Prevention Research Institute of Texas, and the Welch Foundation (F-1515).

2.2 Introduction

In 2010, three research groups reported the startling discovery that cytidine triphosphate (CTP) synthase, the enzyme that catalyzes the last and rate-limiting step of *de novo* CTP biosynthesis, can reorganize into extended intracellular fibers in bacterial, fungal, and animal cells (Ingerson-Mahar et al. 2010, Liu 2010, Noree et al. 2010). It was questioned whether these fibers are enzymatically active, if they serve as cytoskeletal elements, or if they fulfill other, as yet undetermined, regulatory or structural roles. However, far from being a unique occurrence, these fibers are just the latest in a growing assortment of intracellular bodies formed by metabolic enzymes.

In fact, observations of such bodies have increased markedly in recent years. The technological capacity to perform large-scale microscopy screens of protein localization has made possible cell-biological studies that focus on particular cellular conditions. Such screens have been further abetted by the availability of extensive libraries of bacterial, yeast, and mammalian cells expressing proteins fused to reporters. In parallel, biologists and bioengineers have begun constructing novel multi-enzyme complexes to alter or improve metabolic capacity. Thus, the cellular principles underlying the assembly of large protein bodies and their contribution to the dynamic regulation of metabolism are ripe for exploration and exploitation.

This Chapter serves to survey classic and recent examples of intracellular protein bodies, focusing on those formed by metabolic enzymes. We then discuss the current understanding of when and why such bodies form and what their formation

implies for the functionality—and dysfunctionality—of the enzymes that comprise them. Finally, we speculate on more general models for how such complex quaternary structures form in the first place and why they may be inevitable.

A significant portion of the work in this Chapter is indebted to Jeremy O'Connell, who experimentally investigated foci in yeast and extensively reviewed the historic literature on the panoply of intracellular protein assemblies. My primary contributions to this body of work are reviewing and interpreting these assemblies, especially in light of additional discoveries uncovered in more recent literature. I also experimentally investigated such assemblies formed by purine biosynthesis enzymes in human cells.

2.3 Types of intracellular bodies

There are a wide variety of intracellular bodies that can be classified roughly on the basis of their composition and structures. Bacteria, for the most part (although exceptions exist), do not have membranous compartments, and thus their intracellular bodies tend to be almost exclusively proteinaceous and serve as subcellular compartments with specialized interiors optimized for their relevant biological roles (bacterial micro-compartments)(Bobik, Lehman, and Yeates 2015). Originally called polyhedral bodies when discovered in electron micrographs of the cyanobacterium *Phormidium uncinatum* in 1956 (Drews and Niklowitz 1956), bacterial microcompartments are icosahedral-type polyhedrons of ~90–400 nm in diameter, composed of up to several thousand protein subunits of 10-20 different

types totaling up to more than a gigadalton in mass, and delimited by a protein shell (**Figure 2-1**) (Bobik, Lehman, and Yeates 2015).

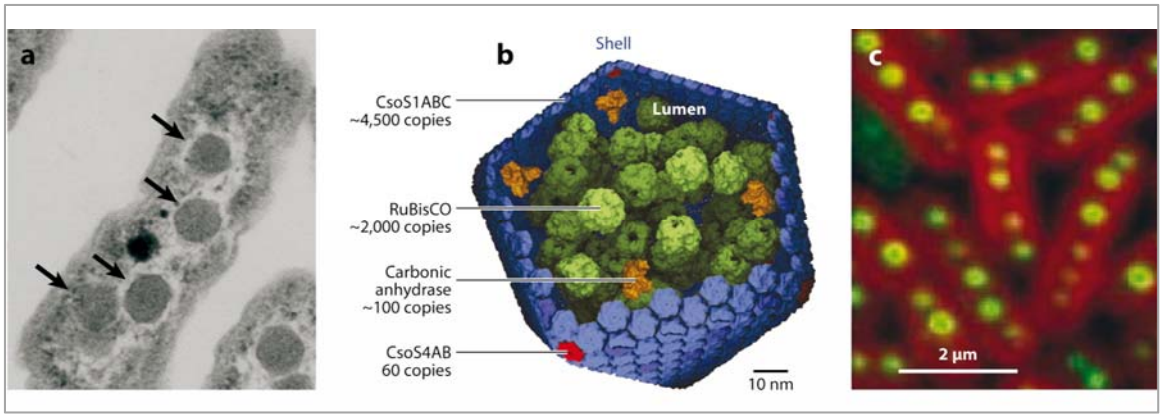

Figure 2-1. Bacterial microcompartments as exemplified by carboxysomes.

Bacterial microcompartments as exemplified by carboxysomes. (**a**) Transmission electron micrographs of *Halothiobacillus neapolitanus* cells showing carboxysomes (arrows) as polyhedral, protein-dense bodies. Adapted from Yeates et al. (2007). (**b**) The major shell protein (CsoS1ABC in *H. neapolitanus*) is a hexagonal subunit that oligomerizes into massive sheets that are bent into joined facets at the vertices by a second, pentagonal protein (CsoS4AB) to complete the shell. The interior of the shell is packed with RuBisCO and carbonic anhydrase to maximize CO_2 capture. Adapted from Bonacci et al. (2012). (**c**) Fluorescence microscopy of carboxysomes (green) shows that their *in vivo* location within cyanobacteria is regulated such that they are centrally aligned and evenly spaced roughly 0.5 µm apart. Adapted from Savage et al. (2010).

In eukaryotes, the formation of membranous compartments is more the norm, and beyond mitochondria and chloroplasts a profusion of specialized intracellular metabolic compartments has been discovered in recent years. These compartments bear a particular relevance to studies of self-assembly of metabolic enzymes, as they frequently exhibit a high level of enzyme self-organization. It can be argued that the crystalline or quasi-crystalline organization of at least some metabolic enzymes

within intracellular membranous compartments is a direct result of their increased concentrations within these microbodies.

Finally, beyond the microcompartments and microbodies discussed above, individual metabolic enzymes have been observed to form intracellular fibers and foci. Textbook cases for which the fiber is known to be the enzymatically active form include acetyl-CoA carboxylase and β-glucosidase. Many additional examples of fiber-forming metabolic enzymes have been identified recently, but their functionality is not yet established.

2.4 Bacterial microcompartments

The archetypal bacterial microcompartment, the carboxysome, earned its name and association with carbon fixation in 1973, when the CO_2-fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) co-purified with polyhedral bodies from the aerobic sulfur bacterium *Halothiobacillus neapolitanus* (then called *Thiobacillus neapolitanus*)(Shively et al. 1973). Initial evidence such as increase in bodies during nutrient depletion and loss of bodies upon refeeding/reactivation supported the idea that carboxysomes were storage bodies rather than the major sites of carbon fixation (Shively 1974, Bock and Heinrich 1971). However, studies over the next 30 years showed that RuBisCO in carboxysomes are active and that carboxysomes also contain carbonic anhydrases to convert bicarbonate into CO_2 (Cannon, English, and Shively 1991, So et al. 2004).

Crystal structures of carboxysome shell proteins model the wall as a single layer of interlocking hexagonal subunits with a pore in the center of each. The sheets of hexagonal subunits are joined at the vertexes by pentagonal subunits to complete the icosahedrons (Iancu et al. 2007, Kerfeld et al. 2005). The shell seems to provide selective permeability to metabolites while blocking the diffusion of even smaller gas molecules by an unknown mechanism (Dou et al. 2008). The co-localization of enzymes and substrates within the diffusion barrier of the carboxysome shell greatly increases the efficiency of carbon fixation, and carboxysomes are the main sites of CO_2 capture in cyanobacteria.

Searches for sequence homology to the major shell proteins have revealed that at least 189 bacterial species contain shell protein orthologs in gene clusters that potentially produce microcompartments (Yeates et al. 2008). Among these, microcompartments with functions other than CO_2 capture have begun to be characterized. In *Escherichia coli* and some *Salmonella enterica* serovars, orthologs of the carboxysome shell proteins were found in operons containing genes essential for growth on ethanolamine (Stojiljkovic, Baumler, and Heffron 1995). Acetaldehyde is a toxic and volatile intermediate in this pathway and might otherwise diffuse away rapidly in the absence of the shell proteins. Thus, a portion of the ethanolamine degradation pathway occurs within an ethanolamine-utilizing microcompartment (Brinsmade, Paldon, and Escalante-Semerena 2005). A similar 1–2 propanediol-utilizing (PDU) microcompartment was found in *S. enterica* serovar t*yphimurium* (Chen, Andersson, and Roth 1994). The PDU microcompartment may shield the cell from the toxic intermediate propionaldehyde (Havemann, Sampson, and Bobik 2002,

Havemann and Bobik 2003). The discovery of an N-terminal targeting sequence for loading proteins into PDU microcompartments and the transgenic expression of functional carboxysomes containing fusion proteins in *E. coli* mark the first steps toward rationally engineering bacterial microcompartments (Bonacci et al. 2012, Fan et al. 2010).

2.5 Aggregates with microbodies

It is interesting that protein microcompartments are found primarily in bacteria rather than in eukaryotes, in the same way that it is interesting that bacteriophages primarily have protein coats, whereas their eukaryotic counterparts have membranous coats. It may be that as the complexity of a system scales, the opportunities for inadvertent protein aggregation also scale (a point we touch on later), and that the less precise organization of lipids relative to proteins is a hedge against such aggregates. Conversely, as suggested above, enzyme compartmentalization may provide opportunities for increasing concentrations to the point at which ordered aggregation is possible. Whatever the reason, eukaryotes frequently compartmentalize metabolic enzymes within membranes. We highlight three cases of ordered aggregates within peroxisomes as examples typical of bodies seen in other membrane-bound compartments such as the mitochondria (Polianskyte et al. 2009) and the chloroplast (Engelbrecht and Esau 1963, Price, Martinez, and Warmke 1966).

Peroxisomes are eukaryotic microbodies bound by a single lipid membrane. These structures, typically 100 nm to 1 µm across, compartmentalize enzymes and substrates at high concentrations for particular metabolic pathways to improve the rates of catalysis or to shield the rest of the cell from the potential damaging effects of reactive intermediates (**Figure 2-2**).

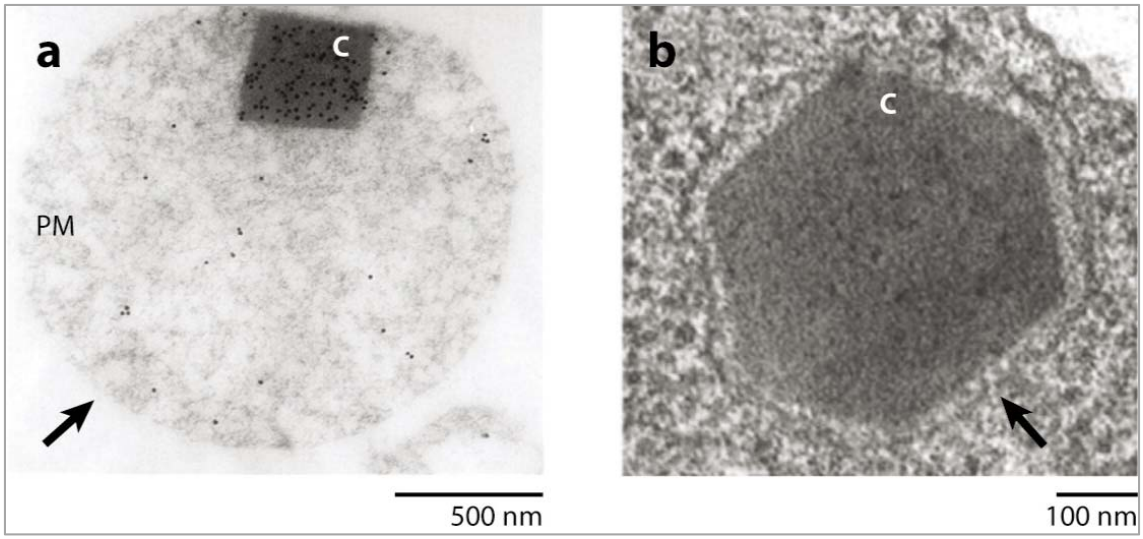

Figure 2-2. Examples of microbodies visualized by TEM.

Microbodies visualized by thin-section transmission electron microscopy include: (**a**) Peroxisomes (arrow) within sunflower cotyledon mesophyll cells. Crystalline inclusion bodies are formed from catalase, as shown by the immunogold nanoparticle localization (black dots). Adapted from Tenberge and Eising (1995). (**b**) *Aspergillus nidulans* showing a Woronin body filled with a HEX-1 protein crystal. Adapted from Yuan et al. (2003). Abbreviations: C, crystalline inclusion bodies; PM, peroxisomal matrix.

Most peroxisomes possess one or more enzymes for purine catabolism or salvage, usually xanthine oxidase and urate oxidase. These proteins are often clustered, even to the extent of forming amorphous and crystalline inclusions. Urate

oxidase is a cuproprotein that normally forms homotetrameric rings, but in some mammalian peroxisomes these tetramers stack into fibers, which combine into crystalline cores (Angermuller et al. 1987). These crystalline cores are a common feature of many types of peroxisomes, though their effects on the enzymatic activity are unknown.

In plants, peroxisomes commonly specialize in β-oxidation and contain high concentrations of oxidases and enzymes to degrade hydrogen peroxide produced by the oxidases. Electron microscopy of plant peroxisomes reveals crystalline and amorphous inclusions likely composed of catalase, a homotetrameric enzyme that degrades hydrogen peroxide (Heinze et al. 2000). Biochemical studies comparing crystalline catalase with diffuse catalase found that the crystalline catalase had up to tenfold-less specific activity but greater stability under UV, pH, and temperature stresses (Eising et al. 1998). Although there is a clear loss of function owing to aggregation, the gain of structural stability in a highly oxidizing environment could be an example of adaptive change in catalytic potential mediated by forming larger aggregates.

Specialized types of peroxisomes, called Woronin bodies, are found in filamentous fungi and staunch the flow of cytoplasm from hyphal wounds (Jedd and Chua 2000). The major and essential component of Woronin bodies is an aggregate of HEX protein oligomers, which form the eponymous hexameric crystalline cores. HEX proteins are related most closely in structure and sequence to eIF-5a, though the residues responsible for inter-subunit contacts are different, and the precise

enzymatic function is unknown (Yuan et al. 2003). Mutation of residues at the oligomerization interface abolishes both the wound-healing function and the classic polymerization phenotype, which shows this to be an example of a functional aggregate (Yuan et al. 2003).

In mammals, the processed form of the β-lactamase-like protein polymerizes into ordered filaments hundreds of micrometers in length in the intermembrane space of mitochondria (Polianskyte et al. 2009). The role of these fibers within mitochondria and the factors that drove the evolution of fiber formation from the non-fiber forming bacterial ortholog are unclear.

2.6 Fibers and foci

Although micro-compartments and membranous organelles are complex and highly structured organizing centers of metabolism, metabolic enzymes also self-assemble into a wide array of simpler intracellular bodies, among them fibers and foci. Many such bodies have been discovered; some are clearly functional, whereas the functions of others have yet to be established. The probability of forming functional aggregates arguably scales with the concentrations of the enzymes involved, and thus, perhaps unsurprisingly, many of these fibers and foci seem to be enzymes that support key metabolic processes such as carbon utilization, nitrogen fixation, and nucleotide biosynthesis.

2.6.1 Fibers of metabolic proteins: carbon utilization

2.6.1.1 Acetyl-CoA carboxylase

Self-assembling fibers of a single enzyme are perhaps the simplest metabolic bodies observed in cells. The textbook example of a functional fiber is acetyl-CoA carboxylase (Gregolin et al. 1966, Ahmad et al. 1978); it was the first and remains the best characterized of all enzymatic fibers. Most mammals encode acetyl-CoA carboxylase as a single multi-domain protein, nominally a homodimer, but one that can assemble into polymers with >50-fold greater activity (Beaty and Lane 1983a, Meredith and Lane 1978). In mammals, there are two isoforms of the enzyme, one localized to the cytoplasm and the other to the mitochondria. Both isoforms carry out the first and rate-limiting step of fatty acid biosynthesis by carboxylating acetyl-CoA to malonyl-CoA. Recently, acetyl-CoA carboxylase filaments have been revealed in yeast (Acc1) under light microscopy, in which Acc1p-GFP assembles into long filaments under prolonged starvation (Shen et al. 2016, Suresh et al. 2015). In bacteria, the beta-carboxyltransferase subunit of acetyl-CoA carboxylase fused to mCherry was observed to assemble into foci (Werner et al. 2009).

Although polymerization is dependent on enzyme concentration, many other mechanisms shift the enzymes between inactive monomers and the active polymer form (Beaty and Lane 1983b). Phosphorylation or de-phosphorylation will decrease or increase acetyl-CoA carboxylase polymerization, respectively; allosteric binding of citrate induces polymerization, whereas excess product (malonyl-CoA) triggers de-polymerization. Binding of MIG12, an effector protein, to acetyl-CoA carboxylase

also induces its polymerization and further increases fatty acid synthesis and triglyceride accumulation (Kim et al. 2010). Interestingly, MIG12 is also involved in stabilizing another well-known polymer: microtubules (Berti et al. 2004).

2.6.1.2 β-glucosidase

Gunning (1965) first reported fibers of β-glucosidase in electron micrographs of oat plastids in 1965. This enzyme is nominally a homohexamer that hydrolyzes β1→4 glucose bonds; it also cleaves avenacosides as an antifungal defense. Structural studies 40 years later revealed rings of hexamers stacked into fibers up to 2 μm in length (Kim et al. 2005), and biochemical analyses provided evidence that the fibers might be the active form of the enzyme: longer fibers were more active in hydrolyzing avenacosides and more resistant to inhibitors than were shorter fibers (Kim et al. 2005).

2.6.2 Fibers of metabolic proteins: nitrogen utilization

2.6.2.1 Glutamine synthetase

Several enzymes at the core of nitrogen metabolism also form fibers. Studies of *E. coli* glutamine synthase (Class I), which catalyzes the ATP-driven addition of ammonia to glutamate, found that zinc progressively induces the dodecameric rings of the enzyme complex to aggregate into fibers (Miller, Shelton, and Stadtman 1974). Mammals have an eye-specific glutamine synthetase (Class I-like) that is catalytically inactive but essential for proper eye development (Harding et al. 2008).

Like crystallins, it is one of many chance adaptations of structured enzyme aggregates to serve as eye proteins. In yeast, glutamine synthetase (Class II) is a decamer that forms foci (Narayanaswamy, Levy, et al. 2009, O'Connell et al. 2014) and filaments (Petrovska et al. 2014) both *in vivo* and *in vitro* under various conditions such as starvation or low pH. Glutamine synthetase filaments constituted *in vitro* with cobaltous ions can form highly ordered helical cables (Frey, Eisenberg, and Eiserling 1975) and wheat sheaf bundles (Valentine, Shapiro, and Stadtman 1968). The formation of these assemblies *in vivo* were shown to have implications on the protein's enzymatic activity and the ability of yeast cells survive and recover from severe starvation (Petrovska et al. 2014), further discussed in **Section 2.7**.

2.6.2.2 Glutamate dehydrogenase and glutamate synthetase

When cells are starved of glutamine, they degrade glutamate to α-ketoglutarate and ammonia using the NAD-dependent glutamate dehydrogenase in mitochondria. In mammals, this enzyme normally forms homohexamers but at high concentrations associates into either highly ordered filaments or helical fiber bundles (Josephs and Borisy 1972). In quiescent yeast, glutamate dehydrogenase (Gdh2p) assembles into long filaments under light microscopy (Shen et al. 2016). It is unclear what effect polymerization has on enzyme activity (Zeiri and Reisler 1978); however, glutamate dehydrogenase has a host of allosteric regulators, including zinc, ADP, and GTP, and those that activate function also promote polymerization and decrease thermal aggregation (Cioni and Strambini 1989, Eisenberg and Reisler 1971, Frieden

1958, Olson and Anfinsen 1952, Sabbaghian, Ebrahim-Habibi, and Nemat-Gorgani 2009).

The reverse reaction—glutamate synthesis—can be carried out either by highly related enzymes using NADPH as a cofactor or by glutamate synthetases, which catalyze a transamination from glutamine to α-ketoglutarate, forming two molecules of glutamate. In starved yeast, the green fluorescent protein (GFP)-tagged glutamate synthetase forms fibers (Noree et al. 2010); similar to glutamate dehydrogenase, the enzyme forms homohexamers, and the fibers observed by fluorescence microscopy may represent fibrillar bundles of the homohexamers.

2.6.3 Fibers of metabolic proteins: nucleotide biosynthesis

2.6.3.1 Cytidine triphosphate (CTP) synthase

Multiple groups have reported that CTP synthase forms filamentous, foci, and ring structures *in vivo* in fly (Chen et al. 2011, Liu 2010, Noree et al. 2010), bacteria (Ingerson-Mahar et al. 2010), budding yeast (Noree et al. 2010), fission yeast (Zhang, Hulme, and Liu 2014), rat (Noree et al. 2010), and human cells (Chen et al. 2011, Carcamo et al. 2011) (**Figure 2-3**). In the crescent-shaped bacterium *Caulobacter crescentus*, CTP synthase forms a single rod lining the inner curvature of the cell. In *Drosophila*, CTP synthase filaments, termed cytoophidia (meaning "cell snake"), exist in two forms, termed microcytoophidia and macrocytoophidia, the latter being considerably thicker and longer than the former. In yeast, CTP synthase assembles into filaments as cells enter quiescence upon

culture growth to stationary phase. In fact, yeast cells express two CTP synthase isozymes, Ura7p and Ura8p, which co-localize within the same filaments but not with other known filament-forming proteins (Noree et al. 2010).

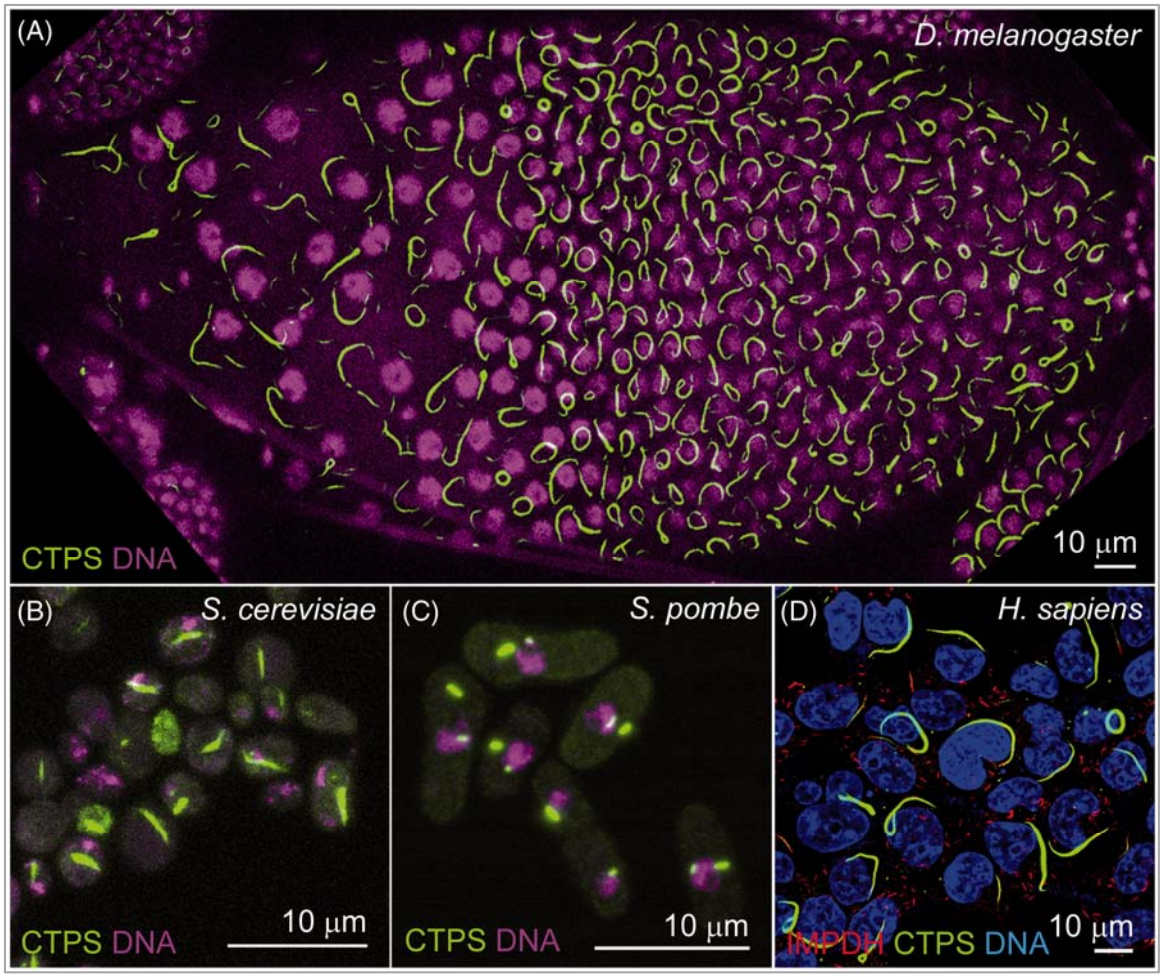

Figure 2-3. CTP synthase fibers within various organisms.

CTP synthase labeled with GFP forms filaments in fruit flies (Liu 2010) (**A**), budding yeast (Noree et al. 2010) (**B**), fission yeast (Zhang, Hulme, and Liu 2014) (**C**), and human HEK293T cells (Carcamo et al. 2011, Chen et al. 2011) (**D**). Nuclei are labeled by DNA dyes (magenta in **A**–**C**; blue in **D**). Scale bars, 10 μm. The fibers exhibit a diverse range of lengths. Originally from Aughey and Liu (2016).

Initial data suggest that these fibers may be formed by the self-association of CTP synthase and do not require energy or special machinery (such as chaperones or scaffolds) for formation. Filaments are seen to form upon heterologous expression of *C. crescentus* CTP synthase in *E. coli*, and native *E. coli* CTP synthase forms filaments both *in vivo* and *in vitro* (Ingerson-Mahar et al. 2010). That said, other studies now implicate IMPDH as additional components in these fibers within higher eukaryotes (Carcamo et al. 2011, Chang et al. 2015), although this observation has since been contested (Probst et al. 2013, Carcamo et al. 2014).

Although the filaments show distinct intracellular locations, these apparently vary in different organisms. In *C. crescentus*, CTP synthase filaments co-localize to crescentin along the inner curvature of the cell and apparently help regulate the crescent cell shape (Ingerson-Mahar et al. 2010). In rat neurons, CTP synthase filaments occur in axons but not dendrites (Noree et al. 2010). Fly microcytoophidia associate with Golgi complexes and in some cases with microtubules (Liu 2010). In contrast, human and yeast CTP synthase filaments were not observed to co-localize with microtubules (Carcamo et al. 2011, Noree et al. 2010), and human CTP synthase filaments also were not observed to co-localize to Golgi complexes or centrosomes and were not enriched in actin or vimentin (Carcamo et al. 2011).

Compounds that modulate CTP synthase protein function modulate fiber formation. For example, CTP synthase inhibitors such as acivicin and the glutamine analogs 6-diazo-5-oxo-L-norleucine (DON) and azaserine (Carcamo et al. 2011, Chen et al. 2011) have dramatic effects on CTP synthase filaments. Interestingly, CTP

synthase inhibitors produce different effects in different organisms: DON treatment disrupts CTP synthase filament formation in *C. crescentus* and also disrupts filaments of heterologously expressed *C. crescentus* CTP synthase in *Schizosaccharomyces pombe* or *E. coli* (Ingerson-Mahar et al. 2010). In contrast, DON and azaserine treatments promote filament formation in fly cells, and DON and acivicin treatments induce filament formation in human cells (Carcamo et al. 2011, Chen et al. 2011). Notably, DON binding induces tertiary (Levitzki, Stallcup, and Koshland 1971) and quaternary (Robertson 1995) structural changes of CTP synthase in *E. coli*. As in the glutamine synthetase case mentioned above, such allosteric changes also alter the conformational state of neighboring subunits coordinately (Levitzki, Stallcup, and Koshland 1971), which is consistent with the hypothesis that differences in inter-subunit amino acid contacts might be exposed or hidden by conformational changes. This would help to explain not only the evolution of fiber formation but also how fiber formation is functionally related to regulatory logic.

Filament formation also can vary broadly according to systemic conditions, including cell types, stages, and growth conditions. Ingerson-Mahar et al. (2010) reported that mCherry-CTP synthase filaments were generally shorter in newly formed stalked cells of *C. crescentus* and then elongated with the progression of the cell cycle. Additionally, subcellular localization was also dependent on cell cycle. Carcamo et al. (2011) observed the opposite in human cells, in which expression of filaments occurred in all phases of the cell cycle in HEp-2 cells, and Chen et al. (2011) found a similar ubiquity in HeLa cells. However, in other

human cells, the filaments varied by cell type and culture conditions. For example, undifferentiated, uninduced human embryonic stem cells contained CTP synthase rings and rods but lost them when stimulated to differentiate by retinoic acid (Carcamo et al. 2011). Similarly, filaments of the yeast CTP synthase Ura7p are induced strongly by growth to saturation or glucose depletion but disappear upon addition of fresh medium. Thus, CTP synthase fibers currently represent something of a quandary: Although widely observed, they show mixed regulatory logic across organisms, and their enzymatic functionality is further explored in **Sections 2.7.1.2** and **2.8.3**.

2.6.3.2 Inosine 5'-monophosphate dehydrogenase, types I and II

A study in human cells that was later contested demonstrated that cytoplasmic filaments consisting of human CTP synthase also contained inosine 5′-monophosphate dehydrogenase (IMPDH) (Probst et al. 2013, Carcamo et al. 2014, Carcamo et al. 2011). IMPDH catalyzes the NAD-dependent oxidation of inosine monophosphate (IMP) to xanthosine monophosphate, the first and rate-limiting step for the synthesis of guanosine nucleotides. The predominant isotype, IMPDH type II, is highly expressed in neoplastic and differentiating cells, which makes it an inviting target for anti-proliferation drugs such as the immunosuppressive, noncompetitive inhibitor mycophenolic acid (MPA) (Nagai et al. 1991, Woolfolk and Stadtman 1967). MPA induced IMPDH2 rings and filaments in cultured human cells, reduced the enzyme's specific activity in cell lysates, and induced purified IMPDH2 homotetramers of ~15 nm in diameter to form large, disordered aggregates *in vitro*

(Ji et al. 2006). Independent studies have observed that IMPDH2 fibers have a regular diameter of about 8.5 nm and consist of regularly shaped particles with a length of approximately 11 nm (Juda et al. 2014). Both MPA-induced rings and fibers *in vivo* and aggregation *in vitro* could be dispersed by the addition of GTP at physiological concentrations, with GTP addition restoring the activity of MPA-inhibited IMPDH2 in cell lysates (Ji et al. 2006). These results are consistent with a hypothesis that MPA stabilizes inter-tetramer interactions and shifts the equilibrium between active IMPDH2 tetramers and inactive fibers to favor the inactive form.

IMPDH type I (IMPDH1), which shares 84% amino acid identity and virtually indistinguishable catalytic activity with IMPDH2, was shown to form intracellular fibers with higher propensity than IMPDH2 (Gunter et al. 2008). Purified human IMPDH1 in the presence of an allosteric effector (Mg-ATP) form fibers by a linear stacking of two conformations of IMPDH1 octamers, in which the eight Mg-ATP-binding domains (CBS modules) point out perpendicular to the fiber axis (Labesse et al. 2013). The CBS module harbors the location of the D226N point mutation in human IMPDH1 responsible for retinal degeneration and IMPDH1-based retinitis pigmentosa. Interestingly, the D226N variant has a high propensity to spontaneously form intracellular fibers (Thomas et al. 2012). *In vitro*, the purified D226N variant show intertwined fibers with strong interactions between adjacent fibers (Labesse et al. 2013). The aggregation of such fibers in the autosomal dominant mutant D226N could explain the onset of such IMPDH1-based diseases (Labesse et al. 2013).

2.6.3.3 Purine biosynthesis and purinosomes

The enzymes of nucleotide metabolism appear to be particularly prone to form intracellular bodies (Shen et al. 2016). In particular, the existence of a multi-enzyme complex consisting of all the members of the *de novo* purine biosynthesis pathway, termed the purinosome, has been postulated for some time on the basis of the accumulation of a variety of evidence from many experiments; this topic is explored in-depth in the next Chapter. The *de novo* purine biosynthesis pathway encompasses a ten-step enzymatic reaction converting phosphoribosyl pyrophosphate (PRPP) to inosine monophosphate (IMP). In higher eukaryotes such as mammals, *de novo* purine biosynthesis is carried out by phosphoribosyl pyrophosphate amidotransferase (PPAT), the trifunctional phosphoribosylglycinamide formyltransferase/phosphoribosylglycinamide synthetase/phosphoribosylaminoimidazole synthetase (GART), phosphoribosylformyl-glycinamidine synthase (FGAMS), the bifunctional phosphoribosylaminoimidazole carboxylase/phosphoribosylaminoimidazole succinocarboxamide synthetase (PAICS), adenylosuccinate lyase (ADSL), and the bifunctional 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase (ATIC). The *de novo* purine biosynthesis pathway is upregulated when exogenous hypoxanthine is unavailable (Becker and Kim 1987, Yamaoka et al. 2001).

In higher eukaryotes, these ten enzymes have fused into six polypeptide chains, three of which possess multiple active sites, as listed above. Early hints that the purinosome might form included co-immunoprecipitation of trifunctional GART

and bifunctional PAICS from chicken liver (Caperelli et al. 1980), along with serine hydroxymethyl transferase and the trifunctional methylene tetrahydrofolate dehydrogenase (MTHFD1). These latter two enzymes form a cycle for the production of the labile 10-formyl tetrahydrofolate coenzymes required for steps 3 and 9 of purine biosynthesis. Cross-linking studies *in vitro* found a physical interaction between GART and MTHFD1, and their interaction increased GART activity (Smith et al. 1980). Fluorescence microscopy of transiently expressed, fluorescently tagged proteins showed co-localization between tetrahydrofolate synthase and FGAMS or GART (Field, Anderson, and Stover 2011).

In 2008, the Benkovic group reported the first direct observation of the purinosome in cells (An et al. 2008). Fluorescent protein constructs of the six mammalian purine biosynthesis enzymes formed intracellular punctate foci when transiently overexpressed in HeLa cells in purine-depleted culture medium. FGAMS-OFP was shown to co-localize with the five other enzymes of the pathway, which suggested the assembly of *de novo* purine biosynthesis enzymes into purinosomes. FGAMS-GFP foci could be dissolved by exchanging purine-depleted for purine-rich medium, although the addition hypoxanthine, a purine derivative, to purine-depleted medium did not dissolve the foci. Intracellular foci of PPAT, GART, ADSL, and ATIC were also detected via immunofluorescence confocal microscopy of endogenous enzymes in many human cell types, including primary keratinocytes. The bodies formed and dispersed when culturing cells in purine-depleted and purine-rich medium, respectively (Baresova et al. 2012).

Inhibitor studies have yielded some insights into the mechanism of purine enzyme foci formation, which will be expounded upon in the next Chapter. The study of these foci is a case study in how the methods used to query intracellular bodies potentially influence outcomes. These foci have been observed largely by transient overexpression with fluorescent protein tags. In contrast, previous stably expressed purine biosynthesis enzymes did not form foci (Gooljarsingh et al. 2001). Overexpression of proteins can result in deleterious disruption of homeostasis leading to aggregation (Tartaglia et al. 2007). As well, transient transfection is well-known to induce stress and cell death. The transiently overexpressed purine biosynthetic enzymes formed foci to extents that correspond to each enzyme's predicted aggregation propensity (Zhao et al. 2013). Foci formation could be induced by general stress agents, and foci were marked by ubiquitin and heat-shock chaperones, which suggested that they may represent aggregated protein clusters (Zhao et al. 2013).

The formation of functional purinosomes is not necessarily inconsistent with the observation of aggregated purine biosynthetic enzymes. Indeed, both may be true, depending on the cell state and method of observation. Foci formed by purine enzymes appear to be quite heterogeneous, as the penetrance of fluorescent body formation varied broadly for each individual enzyme, ranging from 5% to 77% of the cells (An et al. 2008). A similar broad distribution was observed for CK2 inhibitor-mediated effects (ranging from 15% to 95% penetrance) (An, Kyoung, et al. 2010). Regulatory influences can also be interpreted in terms of either functional or nonfunctional aggregation (or both): although microtubules and CK2 could regulate

purine enzyme foci, they have also been implicated in general protein aggregation (Muchowski and Wacker 2005, Watabe and Nakaki 2011, 2012). Nocodazole inhibition of purine enzyme foci formation (An, Deng, et al. 2010) is consistent with functional assembly requiring microtubules, but nocodazole treatment also inhibits inclusion body formation of aggregated huntingtin (Kaminosono et al. 2008, Muchowski et al. 2002), suggesting general protein aggregation mechanisms. Future studies will clearly be required to address the key questions of functionality of the purine enzyme foci, ideally leading to purification and *in vitro* characterization, and to assess the relationships, if any, between the intracellular foci formed by the recombinant purine biosynthetic enzymes and the endogenous forms of those proteins.

2.6.4 Large-scale screens in various systems reveal many additional intracellular foci and fibers

The frequency at which intracellular bodies have been found during the study of metabolism indicates either that metabolic enzymes are prone to aggregation or, perhaps, that all enzymes are prone to aggregation and, to date, biochemists just happen to have largely studied metabolic enzymes. In fact, large-scale microscopy screens of protein localization dynamics have revealed tens to hundreds of additional enzymes that form intracellular bodies (**Figure 2-4**). Using cell microarrays, Narayanaswamy, Levy, et al. (2009) surveyed large-scale trends in yeast (*Saccharomyces cerevisiae*) protein reorganization using a genome-wide GFP fusion library (Huh et al. 2003). When ~800 yeast strains expressing normally cytosolic

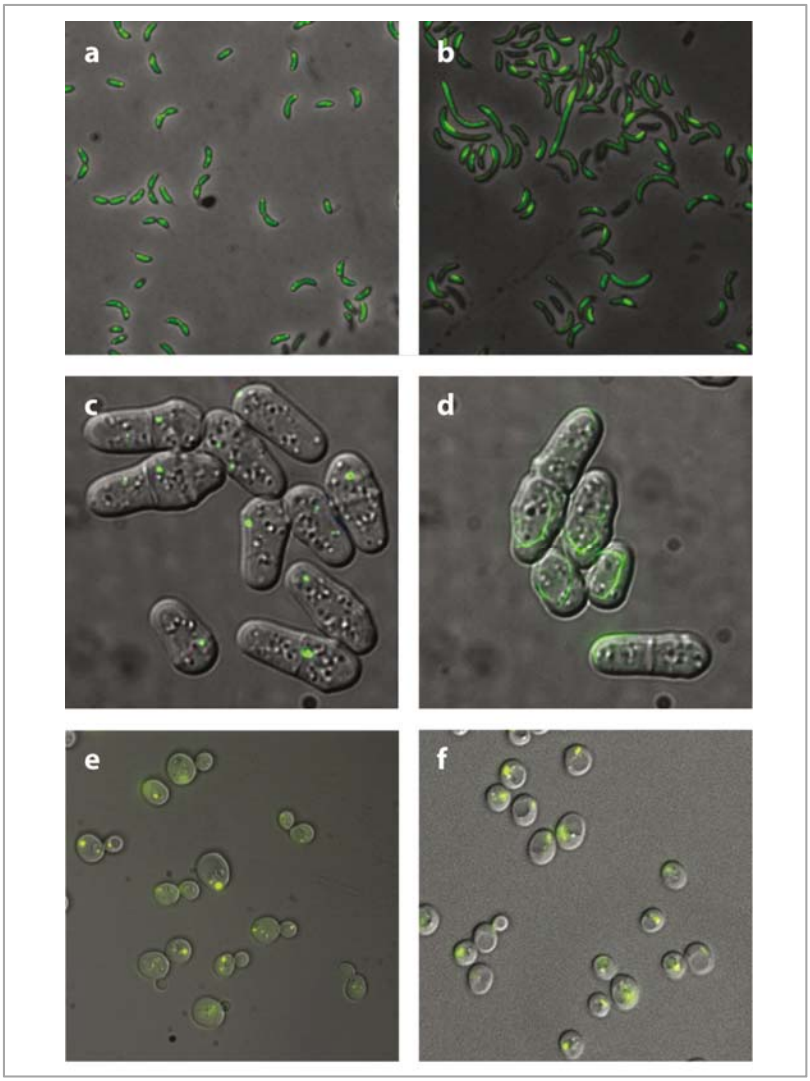

Figure 2-4. Hundreds of foci- and fiber-forming proteins discovered in various organisms.

Hundreds of foci- and fiber-forming proteins have been discovered in systematic protein localization screens; most of these intracellular bodies are still largely uncharacterized. (**a,b**) Representative foci (acetyl-coA carboxylase carboxyl transferase β-subunit)- and fiber (UDP-N-acetylmuramate-alanine ligase)-forming mCherry fluorescent protein fusion proteins, respectively, from *Caulobacter crescentus.* Adapted from Werner et al. (2009). (**c,d**) Representative foci (Ade4p)- and fiber (Pil1p)-forming yellow fluorescent protein–fusion proteins, respectively, from *Schizosaccharomyces pombe*. Adapted from Matsuyama et al. (2006). (**e,f**) Representative foci (Gln1p)- and fiber (Asn2p)-forming GFP-fusion proteins, respectively, from *Saccharomyces cerevisiae*. Adapted from Narayanaswamy, Levy, et al. (2009).

GFP-tagged proteins from their native locus in the genome were grown to stationary phase, 180 proteins involved in intermediary metabolism and stress response were observed to form cytoplasmic punctate foci. The formation of many of these protein macrostructures was confirmed also by immunofluorescence and mass spectrometry of untagged proteins, with 33 proteins confirmed by both microscopy and mass spectrometry (Narayanaswamy, Levy, et al. 2009).

A second screen of a portion of the *S. cerevisiae* GFP library by Noree et al. (2010) found 29 foci- and 9 filament-forming proteins. Of these, three distinct fibers were formed by metabolic proteins. One was formed by CTP synthases Ura7p and Ura8p. The other two fibers, composed of Psa1p and Glt1p, were novel observations. Psa1 is a GDP-mannose pyrophosphorylase essential for building the glycoproteins of the cell wall and is highly conserved across eukaryotes. The other, Glt1, is an NAD-dependent glutamate synthase, which, along with Gln1, forms one of the core ammonia incorporation pathways. CTP synthase fibers have been further extensively characterized and suggested to comprise inactive forms of the enzyme (Noree et al. 2014), further discussed in **Section 2.7**.

A third screen of the entire *S. cerevisiae* GFP library (which represents ~75% of the *S. cerevisiae* genome) by Shen et al. (2016) report 23 filament-forming proteins, confirming previous findings and declaring new filament-forming proteins, including asparagine synthetases Asn1p and Asn2p, glycogen debranching enzyme Gdb1p, glutamate dehydrogenase Gdh2p, phosphofructokinases Pfk1p and Pfk2p, and thioredoxin peroxidase Tsa1p. Most of the identified filament-forming proteins

are metabolic enzymes, a point we touch on later in **Section 2.8.1**. Culture conditions apparently affected the occurrence and length of the metabolic filaments. Additionally, Asn1p/Asn2p filaments were also found in the nucleus.

As well, several systematic microscopy screens of protein localization have been performed in the fission yeast *S. pombe* (Ding et al. 2000, Hayashi et al. 2009, Matsuyama et al. 2006, Sawin and Nurse 1996). Although these screens were not searching specifically for new structures, they still describe many proteins appearing as cytoplasmic dots (e.g., CTP synthase Ura7p and Ura8, and phosphoribosylpyrophosphate amidotransferase Ade4p, as were seen also in *S. cerevisiae* (Narayanaswamy, Levy, et al. 2009)) or in fibrous morphologies (e.g., Pil1p, an essential cell wall peptidoglycan synthetase).

A localization study of ~300 cytoplasmic proteins labeled with mCherry in the asymmetric bacterium *Caulobacter crescentus* identified many proteins that showed non-diffuse localization (Werner et al. 2009). Of these, they described 29 proteins' cellular distributions as discrete foci, 129 proteins as patchy/spotty, and 3 as filament-forming. Two of the filament-forming proteins are CTP synthase and an associated structural protein, CreS. The third fiber is formed by UDP-N-acetylmuramate-alanine ligase. Similar to Psa1, it is an essential enzyme for peptidoglycan synthesis.

These large-scale screens clearly reveal a remarkably extensive assortment of intracellular bodies forming across diverse environmental conditions. It would appear

that at least some of these bodies are metabolically inducible and form reversibly, which strongly suggests functionality. For example, the yeast purine biosynthetic enzyme Ade4p-GFP formed foci in the absence of adenine, and cycling between punctate and diffuse phenotypes could be controlled by adenine removal and supplementation, respectively. Similarly, yeast glutamine synthetase (Gln1p-GFP) foci cycled reversibly in the absence and presence of glucose (Narayanaswamy, Levy, et al. 2009). Finally, there is evidence for the yeast translation initiation regulatory complexes eIF2 and eIF2B polymerizing into filaments during log phase growth (Noree et al. 2010, Campbell, Hoyle, and Ashe 2005). As this is precisely when translation rates are the highest in yeast, it suggests the fibers may be functional and regulated. However, as with the purine metabolism enzyme foci, caution must be used in interpreting whether the bodies form for functional roles or whether regulated changes in concentration or fluorescent protein tagging inadvertently led to intracellular aggregates.

All of these large-scale localization studies use fusion proteins, and the properties of the tag can affect the solubility and interactions of the tagged protein. One recent study found that a commonly accepted proteinaceous body formed by Clp proteases was entirely dependent on certain fluorescent tags (Landgraf et al. 2012). In the absence of a tag or expressed with a GFP evolved for monomeric expression in *E. coli*, Clp proteases did not form the bodies. Similarly re-tagging five additional proteins that were reported to form bright foci in multiple fluorescent protein libraries resulted in greatly reduced or entirely eliminated foci (Landgraf et al. 2012). The authors proposed that the dimerization of fluorescent protein tags (a

known property of many fluorescent proteins) cause homo-oligomeric complexes to assemble into an extended network to produce an intracellular body (Landgraf et al. 2012). If this is a major cause of intracellular bodies, it predicts the set of foci forming proteins should be strongly enriched for oligomers. Thus, structures discovered by large-scale fluorescence localization screens need to be confirmed by orthogonal, preferably tag-free, methods to verify their biological relevance. Further evidence of tag-induced mis-localization is discussed in later chapters.

2.7 Three potential roles for intracellular bodies

The panoply of intracellular protein bodies also raises interesting questions regarding their evolution and maintenance within cells. We conclude by offering additional speculations on several particular aspects of how and why such bodies might evolve.

Case studies such as CTP synthase and the purine biosynthesis enzymes raise many interesting questions regarding the functionality of intracellular bodies. It is thus useful to consider why metabolic enzymes might assemble into such large intracellular assemblies. In general, metabolic enzymes are notable for often forming large complex quaternary structures (e.g., pyruvate dehydrogenase). These massive intracellular assemblies may provide functional advantages to the cell, such as catalytic efficiency or improved regulation. Alternatively, these structures might be depots for the storage of functional proteins or the disposal of dysfunctional ones. Although numerous metabolic enzymes clearly form functional and well-

characterized meta-assemblies, it is not yet clear for many of the most recently discovered intracellular bodies whether they represent functional structures, storage bodies, or aggregates. Distinguishing these roles remains one of the major challenges for understanding these structures.

2.7.1 Catalytic efficiency and improved regulation

2.7.1.1 Substrate channeling

Classically, enzymes have been thought to organize into multi-subunit assemblies to improve their functionality. Quaternary structures enable the channeling of substrates between active sites on individual subunits, thereby protecting labile intermediates from side reactions in the cell or protecting the cell from toxic reaction intermediates. For example, the first intermediate substrate in *de novo* purine biosynthesis, 5-phosphoribosylamine, has a cellular half-life of 38 seconds, and channeling is essential for its subsequent coupling to glycine (Schendel et al. 1988). Also, the addition of CO_2 to aminoimidazole ribonucleotide to form carboxyaminoimidazole ribonucleotide (CAIR) through two enzymatic steps with N^5-CAIR as an unstable intermediate substrate may have driven the fusion of the enzymes responsible for these two steps in eukaryotes (Meyer et al. 1999). Finally, the N^{10}-formyl-tetrahydrofolate coenzyme used in steps 3 and 9 is moderately labile, with a half-life of 30 minutes (Smith et al. 1980). Such observations underpin the search for the purinosome, which would in principle localize the *de novo* purine biosynthesis enzymes and their coenzymes within sufficient proximity to prevent the diffusion of unstable intermediates or substrates in a fashion similar to how the

peroxisome shields the cytosol from peroxide radicals generated by oxidases in fatty acid and purine catabolism. A related phenomenon is that of ethanolamine-utilizing micro-compartments, which localizes the production and degradation of toxic aldehydes within a protein shell.

Channeling substrates between active sites in a quaternary structure or within a compartment can also improve metabolic efficiency greatly, even when no side reactions are in play. For instance, carboxysomes prevent the diffusive loss of CO_2 during carbon fixation and thereby allow unicellular organisms to achieve C4 plantlike efficiencies. In an even grander organization of metabolic machinery, the cellulosome anchors enzymes involved in cellulose production via elaborate, interlocking multidomain protein scaffolds on the cell surface (Beguin and Lemaire 1996).

2.7.1.2 Allostery and cooperativity

Finally, quaternary structure formation allows regulation by cooperative interactions and allosteric effectors. There are, of course, many such known examples of cooperativity in enzyme oligomers. Among these, the dodecameric glutamine synthetase from *E. coli* is one of the best understood. In addition to two covalent modification enzyme systems, glutamine synthetase has eight direct allosteric inhibitors that (individually) partially and (together) cooperatively inhibit activity (Eisenberg et al. 2000). The allosteric inhibitors bind active sites that are positioned at the interface between the subunits of the two hexameric rings, and in this way

binding is transduced into structural changes that can be transmitted between enzymes in both rings. This mode of regulation effectively integrates information about the metabolic state of the cell with overall enzyme activity (Woolfolk and Stadtman 1967), and such changes could in principle modulate the formation of glutamine synthetase foci and fibers *in vivo*. In yeast, glutamine synthetase (Gln1p) is composed of two similarly arranged pentameric rings that interact to form a face-to-face homo-decamer. Moreover, the crystal structure of yeast Gln1p revealed a back-to-back association between two homo-decamers, a mechanism from which Gln1p filaments may form (He et al. 2009). Indeed, Petrovska et al. (2014) uncovered mutations at the decamer-decamer interface that either abrogated the ability of Gln1p to conditionally form filaments or constitutively form filaments regardless of the growth conditions. However, Gln1p mutants that constitutively formed filaments had strongly reduced synthetase activity (Petrovska et al. 2014), suggesting that while filament formation may regulate Gln1p's enzymatic activity, it is likely not by increasing catalytic efficiency. How the filament formation induces enzymatic inactivation is unclear, although reasonable hypotheses are that the filament-assembled Gln1p is sterically restricted from conformational changes required for enzymatic activity, or that the filaments block substrate access to Gln1p's catalytic site.

Due to our understanding of multiple well-established mechanisms that regulate its activity, CTP synthase serves as a candidate model for addressing whether regulation of enzyme activity is coupled to its fiber formation. Using multiple substrate-binding-site and phosphorylation site mutations, Noree et al.

(2014) found that allosteric activation and a phosphorylation site on the enzyme's glutamine amidotransferase domain, along with sites of substrate binding and end-product inhibition within the amidoligase (synthase) domain, modulated filament length. In particular, mutations in CTP synthase's "L11 loop", a region covering the GTP-induced activation site, is known to alter allosteric regulation of the enzyme. Mutations in this loop that inhibit access of the active site to positive allosteric regulators such as GTP result in Ura7p filaments longer than those observed of wildtype Ura7p (Noree et al. 2014). Conversely, shorter filaments were observed with a loop mutation that encourages GTP interaction (Noree et al. 2014). Interesting, both types of mutations increased the frequency of filament formation, and the reasons for this are unclear (Noree et al. 2014). However, these observations suggest that regulation of the enzyme's activity is tightly coupled to the control of filament formation and/or filament length (Noree et al. 2014). Similar to Gln1p filaments, CTP synthase filaments are also suggested to comprise of inactive enzymes (Barry et al. 2014, Aughey et al. 2014, Noree et al. 2014). While filament formation may be coupled to its enzymatic activity, it does not seem to be for enhancing catalytic efficiency.

2.7.2 Storage depots

It has also been hypothesized that proteins may assemble into macromolecular depots, in which unused components can be held transiently in localized bodies before its eventual re-activation and release back into the cellular pool. The advantage of such depots is that proteins need not be resynthesized but are instead retained for

potential future use, especially in conditions in which rapid re-deployment may be required. For example, in the stationary (quiescent) phase, yeast cells have a remarkable ability to weather extreme stress conditions but can rapidly re-enter the cell cycle (Gray et al. 2004). Many changes accompany the transition into stationary phase; for example, exhibit decreased metabolic rates and increased size and density and in general cease proliferating. The cell walls increase in thickness to provide osmo- and thermo-tolerance, and the cells accumulate intracellular carbohydrates, including glycogen and trehalose, which may serve to help protect the cells against a variety of stresses (Gray et al. 2004). Although rates of transcription and translation are decreased dramatically in quiescent cells as compared with exponentially growing cells (Choder 1991, Fuge, Braun, and Werner-Washburne 1994), the quiescent cells are able to rapidly restart growth when nutrients become available. Quiescent cells have been shown to maintain available pools of important cellular components in forms that can be mobilized quickly, including cytoplasmic processing bodies containing mRNAs that can be translated upon restarting growth (Brengues, Teixeira, and Parker 2005) and actin bodies—localized accumulations of actin that can reassemble into actin fibers and patches as necessary when cellular growth restarts (Sagot et al. 2006). Overall, quiescent cells appear to be rich with dynamic depots that are important for re-entry into the mitotic cell cycle. Such a trend is consistent with the tendency for many cellular proteins to be organized–both spatially and functionally–in a manner consistent with the needs of the cell.

In particular, the aforementioned formation of Gln1p filaments in yeast were shown to be essential for re-entry into the cell cycle after recovery from severe

starvation (0.1 M phosphate buffer). Gln1p filaments were shown to dissolve rapidly when yeast cells re-enter the cell cycle, and the disassembled enzyme complexes were presumably functional, arguing that Gln1p filaments serve as crucial storage depots during starvation (Petrovska et al. 2014).

2.7.3 Aggregation of dysfunctional and/or misfolded proteins

In contrast to the above examples of intracellular bodies with active functional roles or storage of functional potential, a third major category of intracellular bodies is now well established: those composed of aggregated and possibly dysfunctional unfolded proteins. Such bodies often are the product of active cellular processes for collecting, sequestering, and disposing of the aggregates (Tyedmers, Mogk, and Bukau 2010) but also can form when high levels of expression trigger self-association and aggregation (Wickner 1994). Aggregated proteins are commonly sequestered, often via active transport along the cytoskeleton, to specific cellular sites such as the aggresome (Johnston, Ward, and Kopito 1998), IPOD, and JUNQ (Kaganovich, Kopito, and Frydman 2008).

The best-characterized examples of protein aggregates and intracellular aggregation bodies include amyloid fibers and inclusion bodies. Amyloid fibers, in particular, are linked to a range of human diseases, including Alzheimer's, Parkinson's, and Huntington's (Ross and Poirier 2004), but they also occur broadly across proteins and organisms (Halfmann et al. 2012, Goldschmidt et al. 2010). The toxicity of protein aggregates is generally attributed not only to the depletion of

functional machinery (Stefani and Dobson 2003) but also to the creation of pores inside cell membranes by small oligomers (Ahmed et al. 2010). At the early stages of aggregation, small oligomers of amyloid aggregates have structural similarities to pore-forming bacterial toxins and eukaryotic pore proteins (Hirakura and Kagan 2001), and their insertion into membranes leads to ion loss and cell death (Kourie et al. 2002). Both yeast and human prions constitute a continuous spectrum of aggregation with multiple morphologies (Edskes et al. 2009, Legname et al. 2006).

In contrast to bodies formed by aggregates of dysfunctional proteins, some bodies may be formed by aggregates of functional proteins. For instance, recent evidence suggests RNA granules are formed by aggregation of partially unfolded RNA-binding proteins whose low complexity sequences form hydrogels of amyloid aggregates. However, the proteins retain their RNA-binding capacity, and unlike other amyloid aggregates these show highly dynamic assembly and disassembly rates and can incorporate heterogeneous sequences (Kato et al. 2012). These RNA granules still serve various biological functions despite being formed by amyloid aggregates, which are historically thought to be pathological and often associated with cell death in neurodegenerative diseases.

2.8 Speculations and conclusions

Similar to the large-scale screens of protein localization, microscopy screens have also identified numerous new yeast prions (Alberti et al. 2009). Whether assembling bodies for functional reasons, perhaps in response to metabolic cues, or

simply aggregating pathologically, a much larger set of proteins than is broadly appreciated may assemble into bodies or aggregate *in vivo*, perhaps whenever their abundances exceed tolerated limits (Tartaglia et al. 2007). The breadth of these phenomena raises the interesting possibility that all ordered proteins may exhibit some level of self-aggregation or self-assembly.

2.8.1 Are metabolic enzymes intrinsically more likely to self-assemble?

Because metabolic enzymes often exhibit complex quaternary structures and inter-subunit allostery for the purpose of regulation, such quaternary structures might be intrinsically more susceptible to forming intracellular foci and fibers simply as a result of a symmetrical arrangement of enzymes, which would replicate any favorable inter-subunit interactions around the structure's axes of symmetry (**Figure 2-5**). For example, stacked *E. coli* glutamine synthase dodecamers (dimers of hexamers) present six identical interfaces, one between each pair of the six repeated monomers around the dodecamer-dodecamer interface. Any favorable interaction at one such interface is therefore copied six times around the rings, making—via avidity—for a robust overall interface. Such a mechanism of fiber formation is therefore intrinsically more likely for proteins with complex quaternary structures typical of those found among metabolic enzymes. **Figure 2-6** shows a gallery of metabolic enzyme quaternary structures and fibers thought to be formed by such stacking mechanisms.

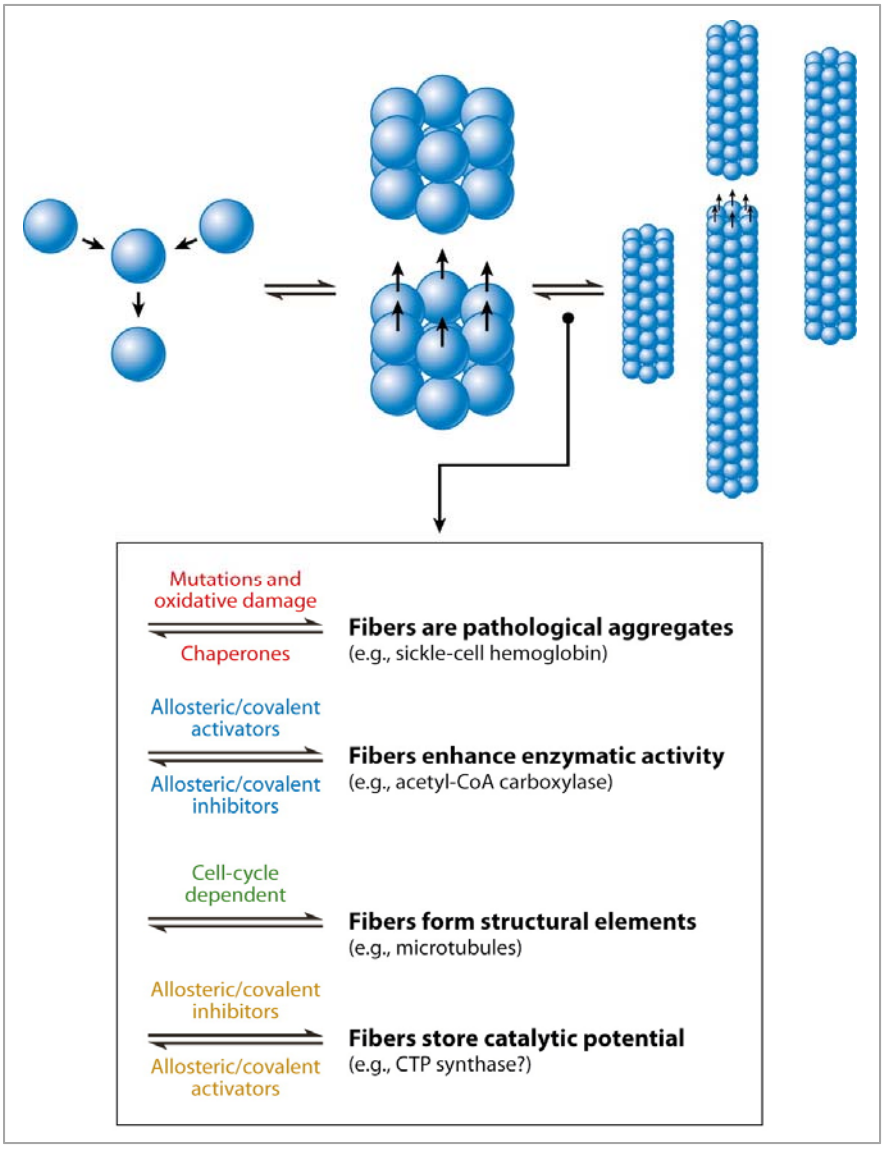

Figure 2-5. Inter-subunit interactions that favor ordered structure assembly

Proteins that assemble into symmetric quaternary structures should in principle have a higher propensity to form fibers, because effectors, whether allosteric, covalent, or mutational, that enhance binding between the oligomeric faces may be multiplied around the axis of symmetry, leading to enhanced fiber stability. The types of effectors and their contribution to fiber stability or destabilization can inform about an enzymatic fiber's role within cells.

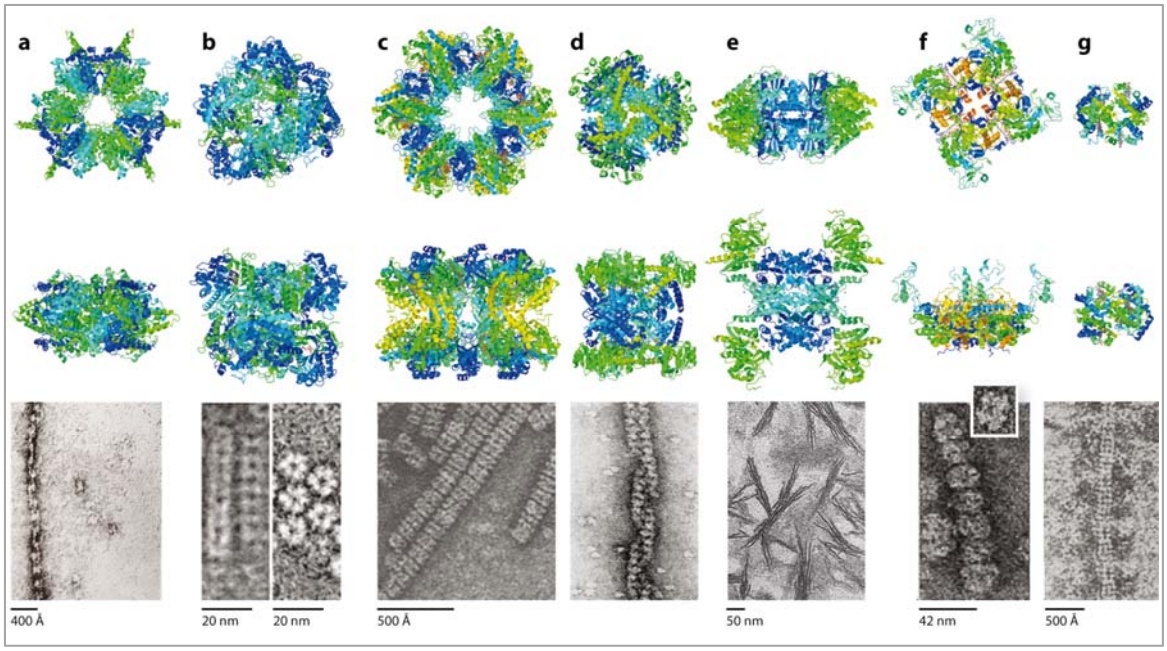

Figure 2-6. A sampling of metabolic enzymes that self-assemble into fibers.

The quaternary structure of each enzyme is illustrated schematically (*top row*), following 90° rotation (*middle row*), and imaged in fiber form by electron microscopy (*bottom row*). (**a**) Acetyl-CoA carboxylase: crystal structure of *Streptomyces coelicolor* acetyl-CoA carboxylase β-subunit, PDB ID: 1XO6 (Diacovich et al. 2004) and electron micrograph of rat liver acetyl-CoA carboxylase (Lehninger, Nelson, and Cox 2005). (**b**) β-glucosidase: crystal structure of wheat β-glucosidase, PDB ID: 2DGA (Sue et al. 2006), and electron micrograph of oat β-glucosidase (Kim et al. 2005). (**c**) Glutamine synthetase: crystal structure of *Escherichia coli* glutamine synthetase, PDB ID: 1FY (HS and D 2001), and electron micrograph of *E. coli* glutamine synthetase (Frey, Eisenberg, and Eiserling 1975). (**d**) Glutamate dehydrogenase: crystal structure of *Clostridium symbiosum* glutamate dehydrogenase, PDB ID: 1BGV (Stillman et al. 1993), and electron micrograph of cow liver glutamate dehydrogenase (Josephs and Borisy 1972). Scale not provided. (**e**) CTP synthase: human CTP synthase 2, PDB ID: 3IHL (M. Moche et al., unpublished data), and electron micrograph of *Drosophila* CTP synthase (Liu 2010). (**f**) Inosine monophosphate dehydrogenase: crystal structure of human type II inosine monophosphate dehydrogenase, PDB ID: 1NF7 (D. Risal, M.D. Strickler, B.M. Goldstein, unpublished data), and electron micrograph of human type II inosine monophosphate dehydrogenase (Ingerson-Mahar et al. 2010). (**g**) Human sickle-cell mutant hemoglobin: crystal structure, PDB ID: 2HBS (Harrington, Adachi, and Royer 1997), and electron micrograph human sickle-cell mutant hemoglobin (Ohtsuki et al. 1977). All crystal structure images created with Jmol.

Given that metabolic enzymes often are also highly expressed and prone to allosteric regulation of alternate conformations, they would seem to be likely cellular candidates for self-assembly into higher-order structures. The resulting fibers might provide regulatory functionality—e.g., fine-tuning of enzymatic output by integrating allosteric interactions across the polymeric interfaces. Such appears to be the case for glutamine synthase, at least on the scale of binding of two homopentameric or hexameric rings (Eisenberg et al. 2000). Such fibers or assemblies might also provide stability or rigidity, especially with regards to the formation of metabolic enzymes into more extensive structures such as crystals, which we might expect to serve as self-chaperoning structures, decreasing the likelihood for component proteins to unfold and form aggregates. Indeed, lens crystallins are derived frequently from metabolic enzymes (Piatigorsky 1993). Similarly, the crystalline cores of peroxisomes as well as mitochondrial glutamate dehydrogenase fibers are cases in which homo-oligomeric enzymes are highly concentrated in stressful environments.

2.8.2 Protein aggregation as an evolutionary compromise

As we have discussed, the formation of intracellular protein aggregates is widespread, and such aggregates can be functional. This seems to be a phenomenon that is not highly selected against and that actually may be beneficial. Although, phylogenetic analyses of multiple proteins suggest that a negative correlation exists between the rate of sequence change and the level of expression (Drummond and Wilke 2008, 2009), which has been interpreted to mean that mistranslation of a highly expressed protein is more likely to lead to dysfunctional aggregation than the

mistranslation of scarcer proteins. Thus, despite the widespread existence of protein aggregates, the entire proteome is constantly under selection to avoid aggregation. The resolution of this conundrum is seemingly that aggregates are largely unavoidable and that functional organization around aggregates is an evolutionary compromise.

This begs the question of why aggregates are unavoidable, given that evolutionary optimization often grinds genotypes and phenotypes to a very fine degree. The answer must lie in the realm of physical principles that not even evolution can refine, and as described below, sickle-cell hemoglobin remains one of the most illustrative examples. A single mutation that occurred independently several times within a relatively short span of evolutionary time (Wainscoat et al. 1983) leads to polymer formation, which confers the beneficial consequence of partial protection against malaria for carriers of this mutation. We would argue that such mutations are largely unavoidable, especially in oligomeric proteins that by definition have the opportunity to form multiple, geometrically repeated contact points.

Because oligomers often are allosterically regulated, and thus by definition assume multiple conformations, there may be unique opportunities for the formation of new mutational contact points. In many cases, new aggregates will be deleterious, as with the previously mentioned proteins associated with Alzheimer's and other prion-based diseases, and evolution will eventually constrain the sequence of the protein in a concentration-dependent manner. In some cases, the aggregates will be neutral, and quaternary structures will form that have the opportunity to eventually

benefit the cell—possibly through the adoption of a new regulatory functionality, as in the case of acetyl-CoA carboxylase, and likely possible for the cases of IMPDH and CTP synthase.

2.8.3 A case study in ambiguity: are CTP synthase fibers cytoskeletal elements or bacterial sickle-cell disease?

The evolutionary conservation of CTP synthase filaments suggests that filament assembly of CTP synthase may provide a biologically useful purpose for cells, such as enhancing enzymatic regulation or even forming new cytoskeletal elements. Experimental evidence now suggests that these filaments consist of inactive enzymes and may serve a regulatory role (Noree et al. 2014, Aughey et al. 2014, Barry et al. 2014). Further functional evidence for CTP synthase filaments includes the observation that they help determine the *C. crescentus* cell curvature. However, for other model organisms in which CTP synthase filaments are found do not exhibit curved cell structures, their functional roles are limited to potential enzymatic activity regulation. Notably, filament formation and enzymatic roles are at least partially separable; for example, Ingerson-Mahar et al. (2010) separated the filament-forming properties of CTP synthase from one of its two enzymatic activities by showing that catalytically inactive synthetase domain mutants retained the ability to form filaments in *C. crescentus*. Similarly, mutation of the active-site cysteine of the Ura7p glutamine amidotransferase domain known to eliminate enzyme activity had no effect on its filament formation (Noree et al. 2014).

The induction of cell straightening by CTP synthase mutants in *C. crescentus* is an interesting comparison to that of sickled red blood cells harboring mutant hemoglobin (**Figure 2-7**). Sickle cell hemoglobin possess a mutation that exposes a hydrophobic amino acid, causing preferential interaction with other hemoglobin molecules rather than association with the cellular environment. This preferential interaction of the hemoglobin molecules with each other drives its polymerization into fibers, which over time grow in size. The long crystallized aggregates are responsible for the rigidity and deformation of sickled red blood cells (Cerami and Peterson 1975). Under electron microscopy, the fibers of CTP synthase lining the inner curve of the cell look remarkably similar to those of sickle cell hemoglobin.

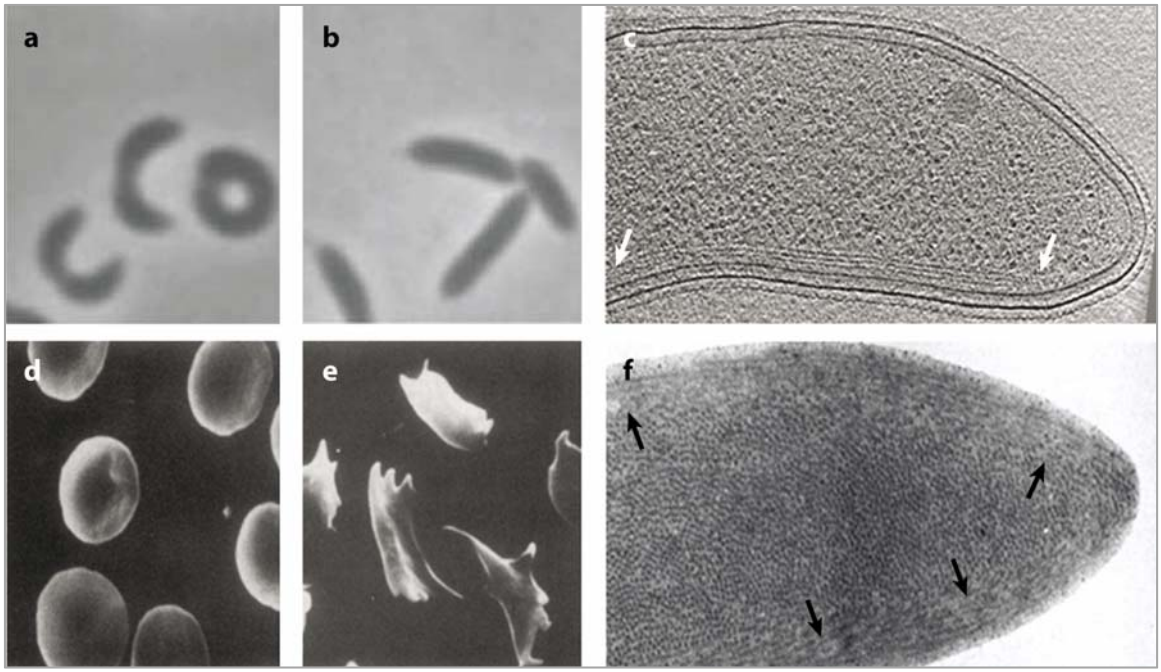

Figure 2-7. Analogous effects on cell morphology by CTP synthase and sickle-cell hemoglobin.

An illustration of the analogous effects on cell morphology by cytidine triphosphate (CTP) synthase and sickle-cell mutant hemoglobin (HbS). (**a**) Bright field images of *Caulobacter crescentus* cells depleted of CTP synthase show severe bending, some to the point of circularization. (**b**) Cells overexpressing CTP synthase are straightened markedly. (**c**) Transmission electron micrograph of a *C. crescentus* cell with CTP synthase fibers (*arrows*) along the cell wall, altering cell morphology. Panels *a–c* adapted from Ingerson-Mahar et al. (2010). (**d**) Analogous images of red blood cells showing their normal, round morphology when oxygenated, and (**e**) their highly straightened morphology when the deoxygenated HbS forms into fibers. Adapted from Kaul et al. (1983). (**f**) Transmission electron micrograph showing HbS fibers (*arrows*) along the cell border, altering cell morphology. Adapted from Dobler and Bertles (1968).

2.8.4 High prevalence of GAT domains in assembled structures

Glutamine amidotransferases (GATs) catalyze the hydrolysis of glutamine to glutamate and ammonia (glutaminase reaction), and the transfer of ammonia from a glutamine molecule to an acceptor substrate (synthase reaction), creating a new

carbon-nitrogen group on the substrate. GATs are grouped into Class-I (triad) and Class-II/N-terminal nucleophile superfamilies. Class-I GAT domains are defined by a conserved catalytic triad consisting of cysteine, histidine, and glutamate, and include CTP synthase. Its glutaminase domain commonly adopts an α−β−α structure fold. Class-II GAT domains contain an active site cysteine present at the N-terminal extremity of the enzyme, whose structure mainly consists of a sandwich of anti-parallel β sheets surrounded by two layers of α helices (α−β−β−α). Class-II GATs include previously mentioned enzymes phosphoribosylpyrophosphate amidotransferase (Ade4) and glutamate synthase (Glt1).

We noticed a high prevalence of either GAT domain superfamilies within the subset of cytoplasmic proteins that formed discrete structures upon culturing into late stationary phase. 12 of the 15 tested GAT domain-containing proteins were able to form structures (**Figure 2-8**, hypergeometric test cumulative $p<3E-6$). Similarly, Shen et al. (2016) also found that many filament-forming proteins shared either GAT domain superfamilies. This domain may serve as an attractive starting point for understanding a potential unifying molecular mechanism of fiber or foci formation. However, the rationale for this observation currently remains elusive.

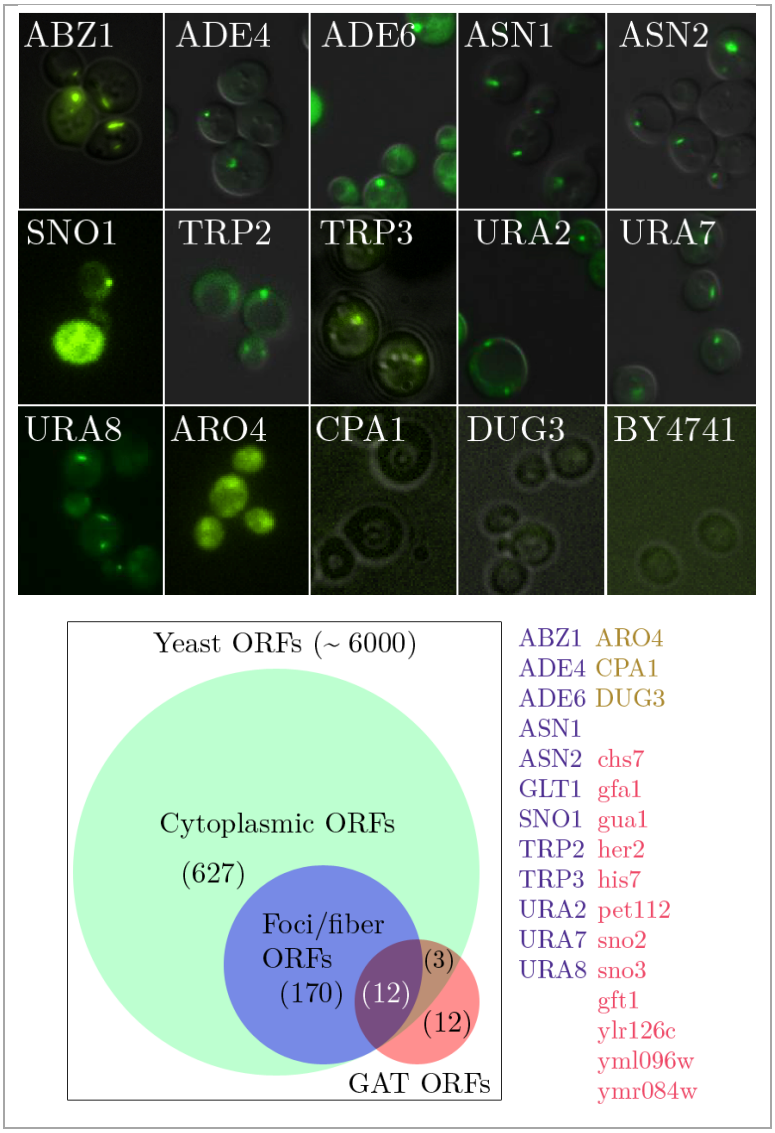

Figure 2-8. Proteins with GAT domain preferentially form structures

Top: Yeast GFP-tagged proteins containing glutamine amidotranferase (GAT) domains form fibers and foci after 5 days in late stationary phase. A fiber-former that is not shown here is GLT1 (Noree et al. 2010). 12 out of the 15 GAT domain-containing cytoplasmic ORFs show ability to form foci and/or fibers. Of the 3 that did not form foci, 2 were not expressed at levels in which subcellular localization was evident (CPA1 and DUG3) compared to the background strain without GFP (BY4741) and 1 is subjectively ambiguous (ARO4) as to whether structures were forming. **Bottom**: GAT ORF enrichment within foci/fiber former's hypergeometric cumulative p<3E-6. Genes that contain GAT domains are listed. Blue gene list contains positive foci/fiber formers, and brown gene list contains ambiguous or negative for foci/fiber formers. Red gene list was not tested due to lack of availability in our GFP library or to the protein subcellular localization not scored as cytoplasmic.

Chapter 3: Purine biosynthesis, purinosomes, and punctate foci formed by purine biosynthesis enzymes

3.1 Abstract

It has been hypothesized that components of enzymatic pathways might organize into intracellular assemblies to improve their catalytic efficiency or lead to coordinate regulation. Accordingly, *de novo* purine biosynthesis enzymes may form a multi-enzyme complex termed the purinosome in the absence of purines. The search for direct *in vivo* evidence of the purinosome culminated in the discovery of intracellular punctate bodies formed in part by purine biosynthesis enzymes. We investigated the mechanism by which human purine biosynthetic enzymes might be organized, especially under differing cellular conditions. Our discoveries support a model where these bodies behave in a manner unlike that expected for functional purinosomes and instead have numerous features one might expect of simple protein aggregates or stress-induced bodies. We demonstrate that the intracellular bodies formed by transiently transfected, fluorescently tagged human purine biosynthesis proteins are best explained as protein aggregation.

Parts of Chapter 3 have been previously published in:

(1) Zhao, A., M. Tsechansky, J. Swaminathan, L. Cook, A. D. Ellington, and E. M. Marcotte. 2013. PLoS One 8 (2):e56203. doi: 10.1371/journal.pone.0056203.

(2) Zhao, A., M. Tsechansky, A. D. Ellington, and E. M. Marcotte. 2014. Molecular BioSystems 10 (3):369-374. doi: 10.1039/C3MB70397E.

(3) PLoS One Reader's Comments, http://www.plosone.org/article/comments/info%3Adoi%2F10.1371%2Fjournal.pone.0056203.

3.2 Introduction

The cytoplasm of living cells is exquisitely organized and dynamic. As previously mentioned, enzymes have previously been found to organize into intracellular assemblies that may improve their catalytic efficiency or lead to coordinate regulation. For example, the trypanosome glycolytic pathway is organized into a glycosome. Similarly, many such cellular bodies occur naturally, functioning in degradation or storage (e.g., P-bodies (Brengues, Teixeira, and Parker 2005) or actin bodies (Sagot et al. 2006)). While such functional bodies continue to be found, it is also the case that overexpression of proteins in cells can lead to aggregation, (Johnston, Ward, and Kopito 1998, Alberti et al. 2009), such as the formation of inclusion bodies. Proteins have the inherent capability to misfold, escape cellular clearance mechanisms, and form intractable aggregates which are frequently associated with neurological diseases.

As I mentioned in the previous chapter, while measuring the localization of green fluorescent protein (GFP)-tagged proteins with the yeast cell chip, we identified a surprisingly large number of intracellular punctate foci or other supramolecular structures (referred here throughout and in other literature as foci, puncta, puncti, punctate foci, bodies, clusters) that accumulated during various environmental growth conditions including nutrient starvation (Narayanaswamy, Levy, et al. 2009). This led us to further speculate whether such bodies were

The work in Chapter 3 was supported by grants from the United States National Institute of Health, National Science Foundation, and the Welch Foundation (F1515), and a Packard Foundation Fellowship to Edward M. Marcotte.

representative of endogenous, functional assemblies, or accidental or pathological aggregates.

Among >100 proteins forming such bodies, we observed that the *de novo* purine biosynthetic enzyme phosphoribosyl pyrophosphate amidotransferase (encoded by the yeast gene ADE4, homolog of the human enzyme PPAT) reversibly formed intracellular bodies in the presence and absence of adenine (Narayanaswamy, Levy, et al. 2009). We initially thought that such bodies might be depots for functional enzymes, and the observation of intracellular bodies associated both with the yeast Ade4 and the human PPAT enzyme suggested the possibility of a functional purine-biosynthetic intracellular body conserved between yeast and humans. In particular, Benkovic and co-workers have identified a cellular body (composed in part of the PPAT enzyme) that they called the purinosome (An et al. 2008), forming in human cell culture in the absence of purines, and whose assembly has been shown to be assisted by microtubules and perturbed by casein kinase II inhibitors (An, Kyoung, et al. 2010, An, Deng, et al. 2010) and which may be under GPCR control (Verrier et al. 2011).

However, the possibility remained that the manipulation of the gene (via fusion to GFP) or its expression (via transfection and/or starvation) had led to aggregate formation. Therefore, we sought to establish whether the observed punctate bodies were formed by transiently expressing the recombinant enzymes in cultured cells. We characterized punctate formation in greater depth and devised

explicit tests to determine whether the bodies may have arisen due to non-native protein expression, stress, or aggregation.

My contributions to the work described in this Chapter include designing and conducting experiments and interpreting data with input from Mark Tsechansky. Lindsey Cook conducted experiments under my guidance, and Jagannath Swaminathan constructed a number of the expression clones.

3.2.1 Purine biosynthesis pathway enzymes

Purines are ubiquitous and essential components of DNA and RNA, and their derivatives participate in numerous biological processes. Adenine and guanine nucleotides are derived from the compound inosine monophosphate (IMP), which is synthesized *de novo* from phosphoribosyl pyrophosphate (PRPP) through a highly conserved multi-step *de novo* purine biosynthesis pathway. In higher eukaryotes (such as humans), the pathway consists of six enzymes catalyzing ten sequential reactions converting PRPP to IMP (**Table 3-1**). *De novo* purine biosynthesis activity is up-regulated when the cellular demand for purines exceeds that supplied by the purine salvage pathway which utilizes purines taken up from the cellular environment, consisting of a single-step conversion of hypoxanthine to IMP catalyzed by hypoxanthine phosphoribosyltransferase (HPRT). Conversely, *de novo* biosynthesis is down-regulated when exogenous purine, *i.e.* hypoxanthine, is available (Becker and Kim 1987, Yamaoka et al. 2001).

Step(s)	Gene	Description
1	PPAT	Phosphoribosyl pyrophosphate amidotransferase
2,3,5	GART	Trifunctional phosphoribosylglycinamide formyltransferase/ phosphoribosylglycinamide synthetase/phosphoribosylaminoimidazole synthetase
4	FGAMS	Phosphoribosylformylglycinamidine synthase
6,7	PAICS	Bifunctional phosphoribosylaminoimidazole carboxylase/phosphoribosylaminoimidazole succinocarboxamide synthetase
8	ADSL	Adenylosuccinate lyase
9,10	ATIC	Bifunctional 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase

Table 3-1. Six human enzymes catalyze the ten-step conversion of phosphoribosyl pyrophosphate to inosine monophosphate.

3.2.2 Circumstantial evidence for the theoretical purinosome

The existence of a multienzyme complex consisting of all the members of the *de novo* purine biosynthesis pathway, termed the "purinosome", has been postulated for some time on the basis of the accumulation of a variety of circumstantial evidence. Some metabolic pathway enzymes are known to organize into multi-enzyme complexes for reasons of catalytic efficiency, metabolite channeling, and other advantages of compartmentalization. Though in the last several decades there has been growing research on multi-enzyme complexes with beneficial roles in the cell, there has been little evidence of anticipated protein-protein interactions that would enable substrate channeling and regulation of the metabolic flux through the *de novo* purine biosynthesis pathway.

The hypothesis that these purine biosynthesis enzymes organize into a multi-enzyme complex has long been attractive based at least in part on the chemical instability of 5-phosphoribosylamine, the first intermediate substrate in the pathway, which suggests an essential direct transfer between PPAT and GART (Schendel et al. 1988). Furthermore, the consolidation of several individual enzymatic functions onto single bifunctional or trifunctional polypeptide chains has been observed in many organisms (Henikoff et al. 1986, Aimi et al. 1990), which suggests that stable physical interactions between these enzymes may exist even in organisms which do not consolidate these enzymes on a single polypeptide chain (Marcotte et al. 1999). Perhaps the most interesting of these is the trifunctional protein GART, which catalyzes the second, third, and fifth enzymatic steps in the pathway. The joining of the non-sequential steps into a single trifunctional enzyme in humans also suggests that this polypeptide may be further, non-covalently juxtaposed with the enzyme for step 4.

3.2.3 Early experimental difficulties in demonstrating the theoretical purinosome

Driven by compelling circumstantial evidence for the purinosome, early experiments aimed at isolating the intact multi-enzyme complex found few interacting pairs. Affinity purification of mammalian GART resulted in near homogeneous protein (Caperelli 1985), although a cruder purification of avian GART resulted in co-purification of PAICS and folate coenzymes (Caperelli et al. 1980). Pairs of purine biosynthesis pathway members (including PPAT and GART, and

ATIC and GART) could be enriched by partial co-fractionation under some conditions, yet no biophysical support was found for a larger multi-enzyme complex or a fully intact purinosome (McCairns et al. 1983, Rowe et al. 1978). Follow-up attempts failed to detect the anticipated protein-protein interaction between PPAT and GART previously suggested by kinetic studies (Rudolph and Stubbe 1995). While biochemical experiments to isolate an intact purinosome were largely unsuccessful, trials with detection in living cells also failed: transfected recombinant GART was not found to be localized to any cellular architecture that might serve as a structural scaffold to assemble a multi-enzyme complex (Gooljarsingh et al. 2001). Despite experimental difficulties to demonstrate the purinosome, a complex representing the purinosome may still physically exist, and it may form only transiently or in a condition-specific manner.

3.2.4 Discovery of transiently expressed purine biosynthetic enzyme bodies

In 2008, the Benkovic group reported the first direct observation of purine enzyme foci in human cells, suggesting that these foci were purinosomes (An et al. 2008). Fluorescent protein constructs of the six mammalian purine biosynthesis enzymes formed intracellular punctate bodies when transiently expressed in HeLa cells in purine-depleted culture medium. FGAMS-OFP foci was shown to co-localize with foci formed by the five other enzymes of the pathway, supporting that these foci contained all direct participants of the *de novo* purine biosynthesis pathway and thus they encompassed purinosomes. These so-called "purinosome clusters"

apparently formed in purine-depleted medium and disassembled in purine-rich medium, a potentially strong indication of function.

Follow-up studies perturbing these foci have yielded some insights into the mechanism of purine enzyme foci formation. Addition of the microtubule-disrupting agent nocodazole reduced foci formation and decreased overall cellular purine synthesis (An, Deng, et al. 2010), which suggests that the formation of these foci may require cytoskeletally directed active transport. Casein kinase 2 (CK2) inhibitors 4,5,6,7-tetrabromo-1Hbenzimidazole and 2-dimethylamino-4,5,6,7-tetrabromo-1H-benzimidazole induced foci formation by FGAMS-GFP, GART-GFP, or PPAT-GFP, whereas the subsequent addition of a different CK2 inhibitor, 4,5,6,7- tetrabromobenzotriazole, appeared to reverse or suppress foci formation. In addition, although CK2 inhibitors affected FGAMS-GFP, GART-GFP, and PPAT-GFP foci formation, similar effects were not seen on PAICS-GFP, ADSL-GFP, and GFP-ATIC unless each was co-transfected with one of the former three proteins tagged with orange fluorescent protein (An, Kyoung, et al. 2010). These results suggest a possible, but as yet unclear, role for CK2 in purinosome regulation. One possibility is that phosphorylation by CK2 disperses assemblies, as seen for acetyl-CoA carboxylase (Meggio and Pinna 2003).

3.2.5 However, purine biosynthesis enzyme bodies do not have to be purinosomes

As purinosomes are structures proposed to enhance purine biosynthesis, the foci should exhibit abundant metabolic flux. It has been shown that flux through the *de novo* purine biosynthesis pathway is suppressed when foci assembly is disrupted by nocodazole. While this correlation supports a role for microtubules in establishing purine enzyme foci in live cells (An, Deng, et al. 2010), these observations may also arise from an alternative explanation, namely that given that nocodazole is a cell cycle arresting agent, the combination of nutrient-poor medium and arrested cell cycle might also be expected to greatly impede metabolic flux. Such experiments illustrate the intrinsic difficulties in distinguishing flux contributed by the bodies and flux contributed by the free, un-localized pool of the same enzymes. Experiments have so far monitored the complete cellular complement of purine biosynthetic enzymes, not explicitly flux through the bodies (An, Deng, et al. 2010, Zhao et al. 2015); to date, no experiments have demonstrated that the foci themselves provide any metabolic flux. Thus, further kinetic experiments are required before purine enzyme foci meet the standard of proof associated with other well-characterized metabolic enhancing structures like acetyl-CoA carboxylase polymers (Beaty and Lane 1983b, a, Meredith and Lane 1978, Clarke and Clarke 1982) or quaternary structures verified for substrate channeling, such as for tryptophan synthase (Pan, Woehl, and Dunn 1997).

3.3 Characteristics of purine enzyme foci

Earlier in this Chapter, I introduced the compelling reasons behind the idea of a multi-enzyme complex consisting of all the members of the *de novo* purine biosynthesis pathway, termed the "purinosome." Although it has never been directly observed nor fully supported on the basis of the accumulation of a variety of biochemical evidence from many experiments, it has been established that the transient expression of recombinant tagged purine biosynthetic enzymes form punctate foci visible by conventional microscopy, referred to here as "purine biosynthetic enzyme foci", "purine enzyme foci", "punctate foci", "foci", "puncta", "bodies" and in other literature as "clusters", "purinosomes", or "purinosome bodies." I addressed some of the key findings supporting the interpretation of the foci formed by purine enzymes encompassing actual purinosomes, yet I also urged caution for this interpretation due to insufficient critical evaluation of other explanations of the same data. Overexpression of proteins above their native levels and/or protein-tagging has been shown to result in their aggregation (Cromwell, Hilario, and Jacobson 2006, Klein and Dhurjati 1995, Mayer and Buchner 2004, Yokota, Kamijo, and Oda 2000, Zhang et al. 2004, Landgraf et al. 2012), and there is no *a priori* reason to believe that purine biosynthetic enzymes are somehow different from enzymes as a whole. Thus, the possibility that purine enzyme foci are no different than aggregated protein bodies should be explored.

In the following sections, I will examine the likely possibility that foci formed by purine biosynthesis enzymes are protein aggregates by addressing many characteristics of such bodies, including their morphological similarity to known

protein aggregates, their non-specific induction patterns, and their association with other proteins known to associate with protein aggregates.

3.3.1 Purine enzyme foci formation in the context of purine dependency

While there is supporting evidence for the hypothesis that purine enzyme foci form in a purine-dependent and reversible manner (An et al. 2008), it is possible that these effects are not solely due to the presence or absence of purines. For example, while it has been reported that foci form when cells are cultured in purine-depleted medium, the specific growth medium employed for these experiments actually significantly differed from the purine-rich control medium in ways that went beyond the mere presence or absence of purines. The purine-depleted medium was generated by serum dialysis with a 25,000 Da pore size, which would have removed a variety of compounds other than purines (which are only ~100–500 Da); additionally, the base media were different between the two conditions (MEM for "purine-rich" and RPMI for "purine-depleted"), and the serum supplementation was doubled for the "purine-rich" medium (An et al. 2008). These non-conservative changes clearly alters many additional components beyond purines and confounds the argument of purine-dependency in those observations. Notably, altering only purine levels did not affect foci formation: purine enzyme foci were unaffected after specifically adding an exogenous purine source (hypoxanthine) back to "purine-depleted" medium (An et al. 2008), or after specifically adding a purine antagonist (azaserine) (An et al. 2008). These contrary observations suggest that the bodies may not form in response to purine levels, but are instead dependent on other factors such as protein

aggregation.

We studied the mechanism by which transfected genes encoding human purine biosynthetic enzymes might be organized into cellular aggregates, especially under different nutrient conditions. HeLa and HEK293 cells transiently expressing GFP- or RFP-tagged purine biosynthetic enzymes were cultured in purine-poor media. Media that was identical to that in An et al. (2008) was reproduced, but to ensure that any observed outcomes could be attributed more specifically to purine depletion, I formulated more optimized media in which our base medium used in both conditions was DMEM supplemented with 10% FBS, in which purine depletion was carried out with a more stringent 1,000 Da pore size dialysis. Punctate foci were observed, resembling those seen by An et al. (2008). While variable penetrance of different purine biosynthetic enzymes was confirmed**,** we were surprised to find that this penetrance was independent of the presence or absence of purines in the growth medium (**Figure 3-1**).

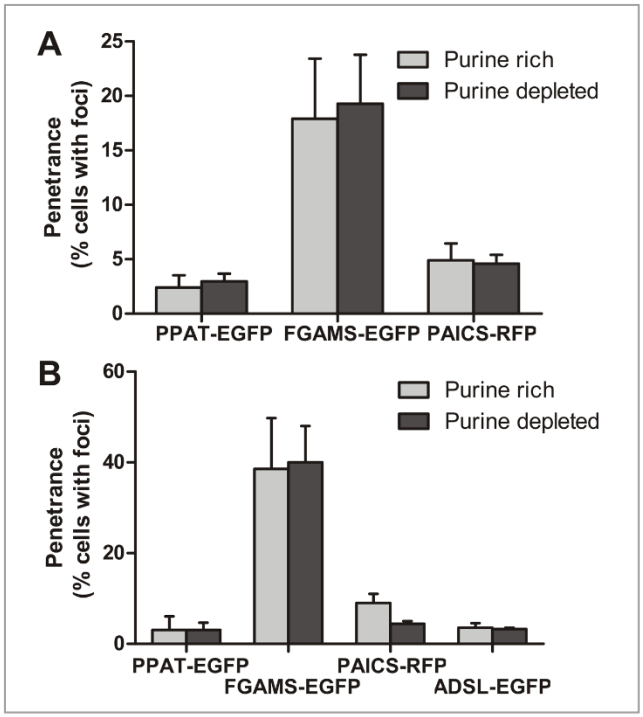

Figure 3-1. The formation of intracellular foci by transfected purine biosynthetic enzymes is not strongly influenced by purine availability

(**A**) The frequency with which intracellular bodies appeared across the population of transfected cells varied for individual *de novo* purine biosynthetic enzymes, and did not depend upon the purine content of the cell growth medium, shown here for HEK293 cells cultured using the optimized medium formulations as described in the Methods. Unless otherwise specified, for HEK293 cells assays in this and later figures within this Chapter, 200 ng of DNA were transfected for each construct. Bars in all experiments represent the average and +/- 1 s.d. from at least 3 replicates, counting n = 769, 742, 755, 665, 1225, and 1016 cells, respectively. See Methods for abbreviations of protein names. (**B**) This trend was similar for HeLa cells cultured in purine-rich medium versus purine-depleted medium, formulated as described previously (An et al. 2008). Unless otherwise specified, for HeLa cell assays in this and later figures, 1.6 µg of DNA were transfected for each construct. Bars in all experiments represent the average and +/- 1 s.d. from at least 3 replicates, here counting n = 578, 821, 662, 832, 1519, 732, 674, and 791 total cells, respectively.

3.3.2 Heterogeneity of foci morphologies and varying prevalence

Published purine enzyme bodies have varied widely in their morphologies, ranging from pinpoint foci to oil droplet-like, and methods have not been developed yet either to classify these bodies or to distinguish them from bodies of any different nature (i.e. aggregates) based on morphology (**Figure 3-2**). In my own studies, I have observed a dynamic spectrum of morphologies, as well as both increases and decreases in the numbers of foci per cell even over the course of unperturbed growth. For my experiments, I did not apply any selective counting of purine enzyme foci; all cells exhibiting visible foci were chosen for analysis.

There is also large heterogeneity observed in the penetrance of foci formation. The penetrance of body formation following transient transfection varies widely across enzymes, with individual enzyme foci formation rates ranging from 5% to 77% of the transfected cells (Zhao et al. 2013, An et al. 2008). This heterogeneity, persisting even in cell populations treated with pharmacophores that promote the formation of purine enzyme foci (15 to 95% penetrance of assembly) (An, Kyoung, et al. 2010), is inconsistent with the hypothesis that each foci contains all members of the pathway, and at a minimum, such foci are unlikely to represent complete purinosomes.

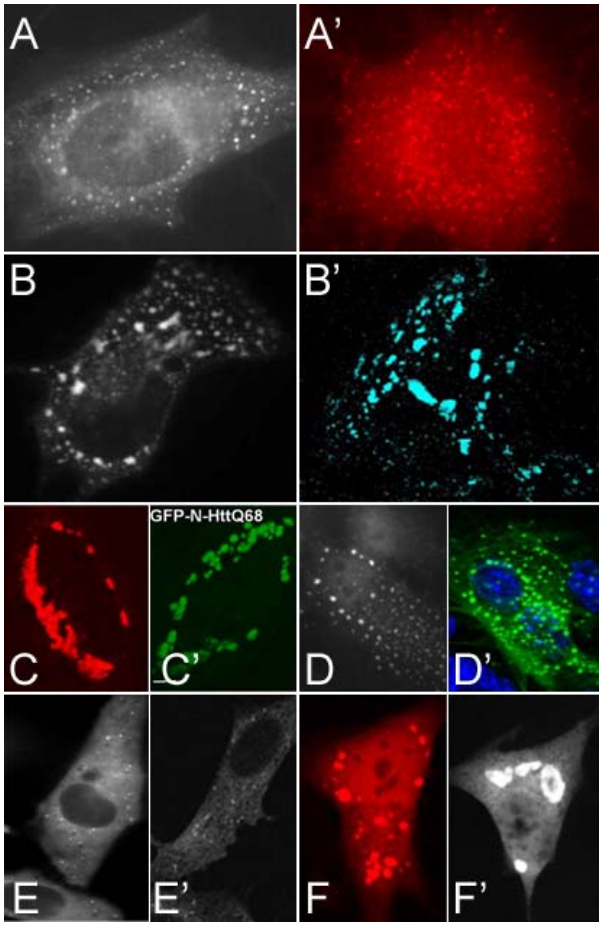

Figure 3-2. Purine enzyme foci display a spectrum of morphologies similar to those of canonical protein aggregates.

(**A**–**F**) Purine enzyme foci in HeLa cells. (**A'**–**F'**) Various disease-associated protein aggregates in indicated cell lines. These representative (but not exhaustive) morphologies can be roughly classified into categories: pinpoint foci of (**A**) TrifGART-GFP (An et al. 2008) and (**A'**) α-synuclein in oligodendroglial cells (Riedel et al. 2010), Droplet-like clusters of (**B**) FGAMS-GFP (An et al. 2008) and (**B'**) mutant glial fibrillary acidic protein (GFAP)-CFP in astrocytes (Mignot et al. 2007). Bulk clumps of (**C**) FGAMS-OFP (Field, Anderson, and Stover 2011) and (**C'**) GFP-tagged huntingtin fragment in HeLa cells (Bjorkoy et al. 2005). Spherical puncta of (**D**) FGAMS-OFP (An, Kyoung, et al. 2010) and (**D'**) GFAP-GFP in astrocytes (Mignot et al. 2007). Sparse pinpoint foci formed by (**E**) FGAMS-GFP (An, Deng, et al. 2010) and (**E'**) a fragment of p62 protein, a common component of disease-associated protein aggregates (Kuusisto, Salminen, and Alafuzoff 2001, 2002, Zatloukal et al. 2002, Nagaoka et al. 2004) tagged to GFP in NIH3T3 fibroblasts (although the authors classify this subtler morphology as diffuse (Bjorkoy et al. 2005)). Large oil-droplet-like bodies of (**F**) hTrifGART-OFP in HeLa cells (Field, Anderson, and Stover 2011) and (**F'**) α-synuclein-EGFP into H4 neuroglioma cells (McLean, Kawamata, and Hyman 2001).

3.3.3 Low rates of co-localization among purine biosynthetic enzymes

A key finding in support of the foci encompassing the functional multi-enzyme purinosome complex is the co-localization of FGAMS-OFP with each of the five other enzymes in the pathway (An et al. 2008). To reproduce these results, I revisited the possible co-localization of pairs of fusion purine biosynthetic enzymes by co-transfecting FGAMS-EGFP with PPAT-RFP, PAICS-RFP, or ADSL-RFP, and also tested PPAT-EGFP with PAICS-RFP or with ADSL-RFP. Only ~1% on average of co-transfected cells showed co-localization of different purine biosynthetic enzymes to the same puncta (**Figure 3-3, Table 3-2, Table 3-3**). Increasing the total dosage of DNA transfected did not alter the rate of co-localization, nor did varying growth in purine-rich or purine-depleted media. More generally, co-transfection tended to suppress puncta formation, or individual enzymes aggregated separately (**Table 3-2**, **Table 3-3**).

We also attempted to co-localize endogenous purine biosynthetic enzymes with protein puncta by use of immunofluorescence with antibodies against endogenous PPAT, PAICS, and GART. However, we did not find commercial antibodies satisfactory for immunofluorescence imaging, as all showed speckling in immunofluorescent experiments, regardless of purine availability (for an example, see **Figure A-1** in **Appendix A**), and no co-localization with transfected recombinant proteins, arguing that the speckles are immunofluorescence artifacts. Similarly, An et al. (2008) also reported puncta, or clustering, with immunofluorescence regardless of cell growth conditions. Baresova et al. (2012) observed punctate structures using

immunofluorescent against purine biosynthesis enzymes, although it is difficult to determine whether those structures are immunofluorescent artifacts or representative of the actual targeted proteins.

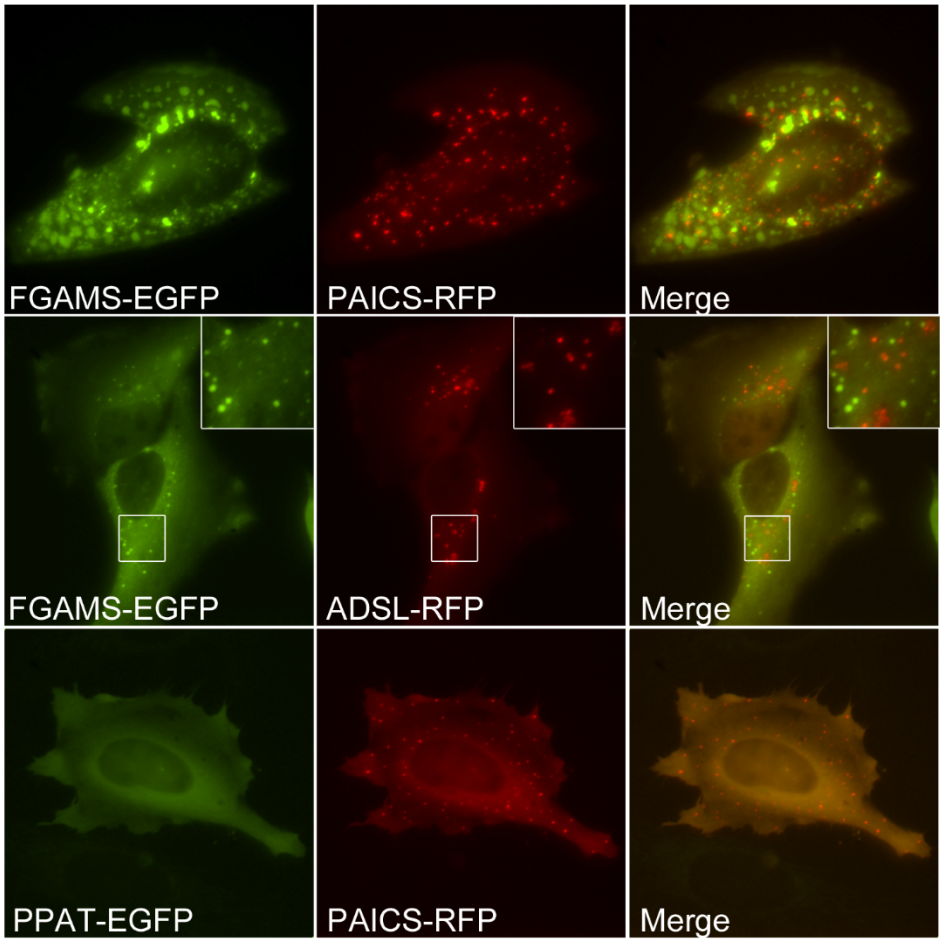

Figure 3-3. Purine biosynthesis enzymes commonly do not co-localize to the same foci.

The top row shows an example of partial but minimal co-localization of FGAMS-EGFP and PAICS-RFP foci. The middle row shows an example of non-co-localizing FGAMS-EGFP and ADSL-RFP foci. The bottom row shows an example of a more typical case, non-co-localization due to the formation of foci by only one protein in doubly-transfected cells, as shown here for PPAT-EGFP and PAICS-RFP.

PPAT-EGFP and PAICS-RFP						
Experiment	DNA (ug)	Both diffuse	P1 foci P2 diffuse	P1 diffuse P2 foci	Both foci NOT co-loc.	Co-localizing foci
6/3/2011	0.8/0.8	343 (99%)	1 (0.3%)	2 (0.6%)	0 (0%)	0 (0%)
6/6/2011	0.8/0.8	592 (99%)	2 (0.3%)	1 (0.2%)	1 (0.2%)	2 (0.3%)
6/20/2011	0.8/0.8	325 (98%)	6 (1.8%)	1 (0.3%)	0 (0%)	1 (0.3%)
6/3/2011	1.6/1.6	425 (100%)	1 (0.2%)	0 (0%)	0 (0%)	1 (0.2%)
6/6/2011	1.6/1.6	686 (99%)	3 (0.4%)	5 (0.7%)	1 (0.1%)	0 (0%)
6/20/2011	1.6/1.6	291 (96%)	5 (1.7%)	3 (1.0%)	2 (0.7%)	1 (0.3%)
PPAT-EGFP and ADSL-RFP						
Experiment	DNA (ug)	Both diffuse	P1 foci P2 diffuse	P1 diffuse P2 foci	Both foci NOT co-loc.	Co-localizing foci
6/3/2011	0.8/0.8	371 (97%)	1 (0.3%)	6 (1.6%)	0 (0%)	4 (1%)
6/6/2011	0.8/0.8	770 (97%)	1 (0.1%)	15 (1.9%)	1 (0.1%)	4 (0.5%)
6/20/2011	0.8/0.8	324 (96%)	4 (1.2%)	7 (2.1%)	1 (0.30)	1 (0.3%)
6/3/2011	1.6/1.6	382 (96%)	0 (0%)	15 (3.8%)	0 (0%)	1 (0.3%)
6/6/2011	1.6/1.6	808 (97%)	1 (0.1%)	14 (1.7%)	2 (0.2%)	7 (0.8%)
6/20/2011	1.6/1.6	360 (95%)	1 (0.3%)	11 (2.9%)	1 (0.26)	5 (1.3%)
FGAMS-EGFP and PPAT-RFP						
Experiment	DNA (ug)	Both diffuse	P1 foci P2 diffuse	P1 diffuse P2 foci	Both foci NOT co-loc.	Co-localizing foci
6/3/2011	0.8/0.8	215 (97%)	0 (0%)	2 (0.9%)	0 (0%)	5 (2.3%)
6/6/2011	0.8/0.8	611 (96%)	3 (0.5%)	10 (1.6%)	1 (0.2%)	13 (2.0%)
6/20/2011	0.8/0.8	386 (93%)	5 (1.2%)	4 (1.0%)	0 (0%)	18 (4.4%)
6/3/2011	1.6/1.6	241 (97%)	0 (0%)	2 (0.8%)	0 (0%)	5 (2.0%)
6/6/2011	1.6/1.6	574 (97%)	0 (0%)	2 (0.3%)	0 (0%)	13 (2.2%)
6/20/2011	1.6/1.6	354 (95%)	5 (1.3%)	6 (1.6%)	0 (0%)	8 (2.1%)

Table 3-2. Minimal co-localization of purine biosynthetic enzyme pairs into foci in "purine-depleted" medium.

Differential expression localization quantified for different purine biosynthetic enzyme pairs, reporting counts and percentages of successfully co-transected HeLa cells as a function of protein localization. Cells were grown in RPMI + 5% 25kDa dialyzed FBS.

FGAMS-EGFP and ADSL-RFP						
Experiment	DNA (ug)	Both diffuse	P1 foci P2 diffuse	P1 diffuse P2 foci	Both foci NOT co-loc.	Co-localizing foci
5/18/2011	1.6/1.6	174 (93%)	6 (3.2%)	5 (2.7%)	0 (0%)	2 (1.1%)
5/19/2011	0.8/0.8	108 (95%)	3 (2.6%)	0 (0%)	0 (0%)	3 (2.6%)
5/20/2011	0.8/0.8	124 (89%)	8 (5.8%)	4 (2.9%)	1 (0.7%)	2 (1.4%)

FGAMS-EGFP and PAICS-RFP						
Experiment	DNA (ug)	Both diffuse	P1 foci P2 diffuse	P1 diffuse P2 foci	Both foci NOT co-loc.	Co-localizing foci
5/18/2011	1.6/1.6	137 (94%)	7 (4.8%)	0 (0%)	0 (0%)	1 (0.7%)
5/19/2011	0.8/0.8	148 (98%)	1 (0.7%)	2 (1.3%)	0 (0%)	0 (0%)
5/20/2011	0.8/0.8	148 (94%)	5 (3.2%)	4 (2.5%)	0 (0%)	0 (0%)

Table 3-3. Minimal co-localization of purine biosynthetic enzyme pairs into foci in "purine-rich" medium.

Differential expression localization quantified for different purine biosynthetic enzyme pairs, reporting counts and percentages of successfully co-transected HeLa cells as a function of protein localization. Cells were grown in MEM + 10% FBS.

3.3.4 Time-lapse dynamics of foci formation and dissolution

Unperturbed time course of live cells transfected with purine biosynthetic enzyme constructs show variable dynamic range. Foci were observed to spontaneously appear and disappear within 1-2 hours (**Figure 3-4**). Replacing the culture medium seemed to induce foci formation, however this induction was apparently regardless of whether "purine-rich" or "purine-depleted" medium was newly introduced. This observation challenges the previous report that the cycling of foci formation and dissolution was dependent on the purine content in the media in the newly introduced media, and instead suggests that a gentle process of replacing media might impose short-lived mechanical or osmotic stress on the cells which results in the temporary aggregation of proteins.

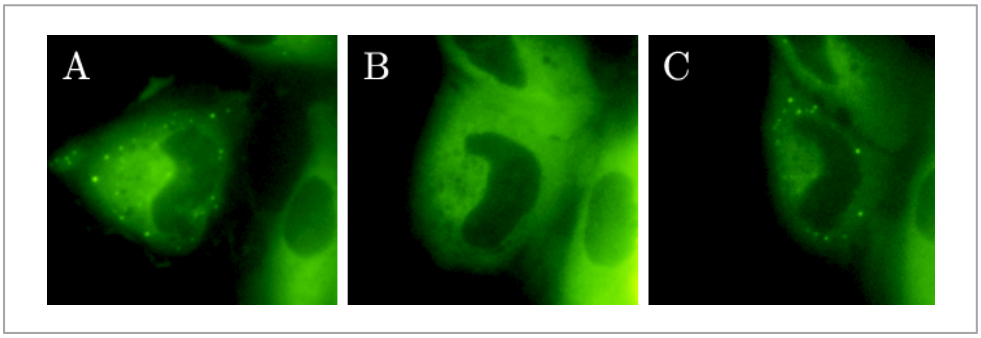

Figure 3-4: FGAMS-EGFP foci form after purine-independent medium replacement and disappear with adaption to the medium.

HeLa cells grown in "purine-rich" medium with FGAMS-EGFP foci may dissolve foci after normal culture condition incubation without perturbation (2 hour incubation, **A** to **B**). However, upon media replacement with identical "purine-rich" media induces foci to reappear within 30 minutes post-media replacement (**C**). A larger field is provided in **Figure A-2 (Appendix A)**.

3.3.5 Foci formation rate correlates with transfected DNA quantity

For single enzyme transfections, we observed a positive correlation between foci formation rate and amount of transfected DNA (**Figure 3-5**). This observation is consistent with the known phenomenon that higher protein production result in higher probability of protein aggregation (Swartz 2001).

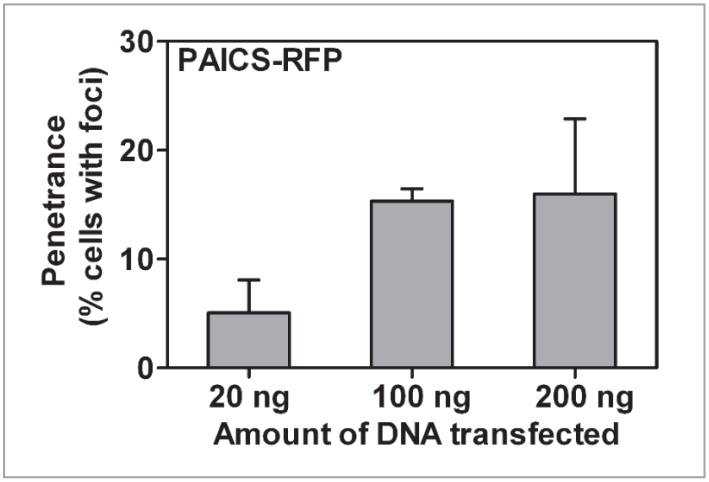

Figure 3-5. Formation of PAICS-RFP bodies in HEK293T cells scaled with DNA transfected.

Among successfully transfected cells, the fraction of the cell population exhibiting PAICS-RFP puncta correlated strongly with the quantity of plasmid DNA transfected. Bars indicate average +/- 1 s. d. across at least 3 replicates, n = 498, 627, and 591 cells, respectively.

3.4 Mounting evidence for protein aggregation

In the previous section, I speculated that the formation of the punctate foci by purine biosynthesis enzymes might be due to factors other than purine starvation.

Previously, the formation of protein puncta in yeast (Narayanaswamy, Levy, et al. 2009) was correlated with protein aggregation potentials measured by TANGO scores (Fernandez-Escamilla et al. 2004), and many metabolic proteins that formed punctate foci seemed to be insoluble aggregates (O'Connell et al. 2014). This suggested the possibility that the recombinant human proteins might be aggregating. Indeed, the fraction of cells in which individual purine biosynthetic enzymes formed detectable punctate bodies correlated well with the predicted aggregation potentials of the enzymes **(Figure 3-6)**. As well, the bodies exhibit overlapping co-localization with a recombinant fusion protein known to mark aggregates and aggresomes (**Figure 3-7**).

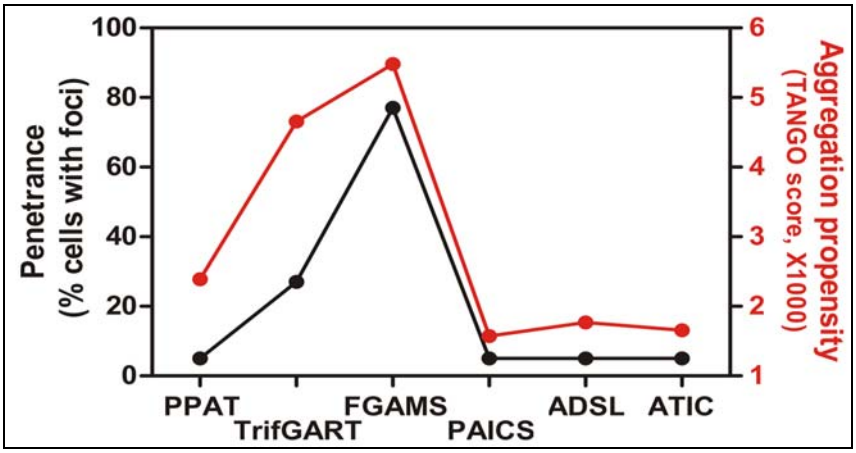

Figure 3-6. Correlation between purine biosynthesis enzyme foci formation rates and their respective TANGO score

The frequency with which transfected cells exhibited punctate foci (shown here for values reported in An et al. (2008)) correlated strongly with each protein's predicted aggregation potential (Fernandez-Escamilla et al. 2004). For comparison, the unrelated enzyme GAPDH has a TANGO score of 879.

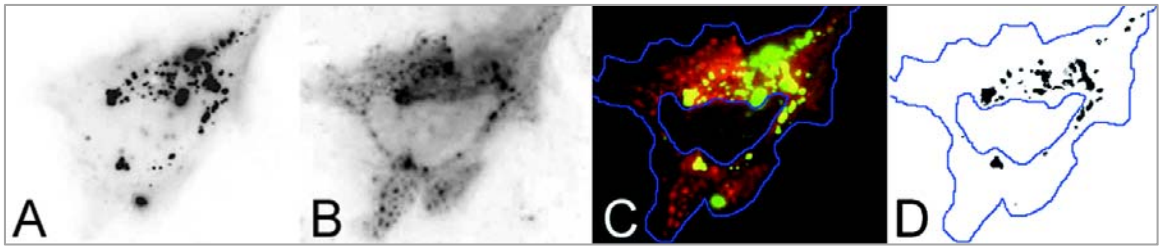

Figure 3-7. Bodies partially overlap with chimera protein aggregate.

Partial co-localization of (**A**) gp-250 (also known as GFP-250) with (**B**) FGAMS-OFP, co-transfected in the same cell (French et al. 2013). The overlay in (**C**) shows the merge of panel A in green and panel B in red (Pearson's coefficient of 0.4 and overlap coefficient of 0.41) with regions of co-localized expression extracted and plotted in (**D**) for clarity. Notably, overexpression of the GFP-chimera, gp-250, has been reported to form insoluble aggregates that are delivered to aggresomes (Garcia-Mata et al. 1999, Garcia-Mata, Gao, and Sztul 2002). Adapted from French et al. (2013).

If the proximal cause of the formation of punctate foci was aggregation, then it seemed likely that: (a) cells may experience differential survivability due to inability to cope with such structures, (b) such structures could be induced by cellular stress, and (c) such structures would contain proteins normally associated with aggregates, such as chaperones and ubiquitin.

3.4.1 Fate of cells with and without purine biosynthetic enzyme foci

During time-lapse microscopy experiments of transfected cells, I observed variation both in punctate body dynamics as well as in cell survival (**Figure 3-8**, **Figure 3-9**). Microscopy analysis of individual transfected cells over the course of 1-2 hours revealed that transfected cells without punctate bodies exhibited significantly higher survival rates than cells marked by punctate bodies (**Figure 3-9**). This difference in survival rate was largely independent of purine availability,

and persisted even when the cell growth medium was exchanged from purine-depleted to purine-rich. The observation that cells marked by punctate bodies died at greater rates than transfected cells lacking such aggregates suggested that these cells likely experienced greater levels of stress, although lacking in an established mechanism and known molecular signals, it is not clear from these data if the punctate bodies were a cause or consequence of that stress.

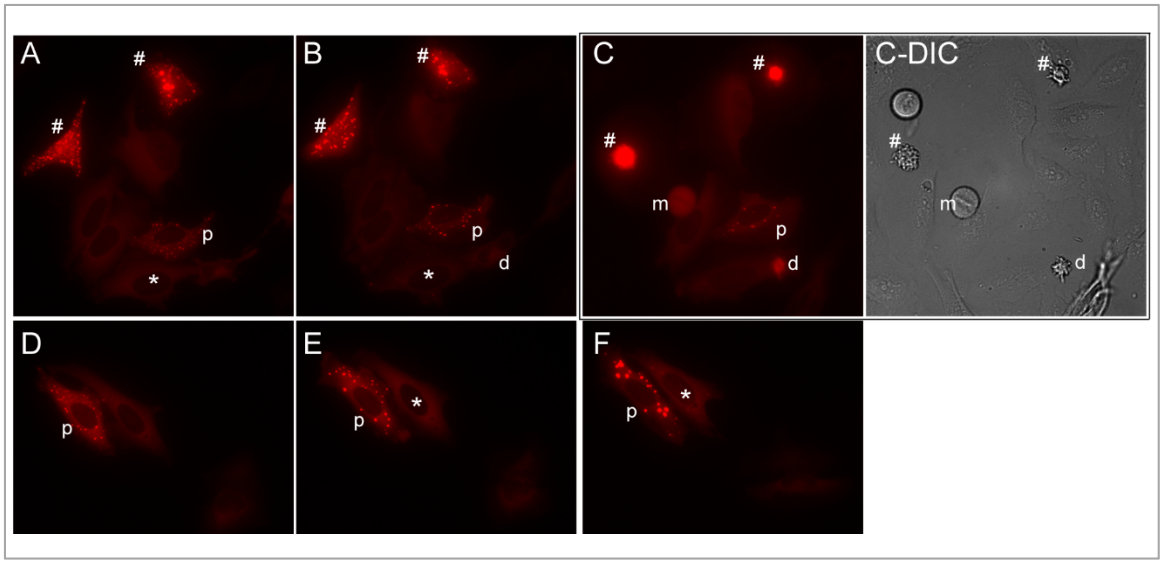

Figure 3-8. Time course imaging reveals that punctate foci are dynamic.

Shown here for HeLa cells transfected with PAICS-RFP grown in purine-depleted medium in two replicate 4 hour time series. Panels (**A**) and (**D**) show the time zero condition; panels (**B**) and (**E**) show the same cells as in (**A**) and (**D**), respectively, following two hours of growth in the same medium; panels (**C**) and (**F**) show the same cells following two hours of growth after exchanging the growth medium to purine-rich. Cells marked by * display formation of punctate foci over the time series, while cells marked by **p** display variable dynamics of punctate foci. The # sign marks cells with punctate bodies that die over the course of the series; the cell marked by **d** dies in the absence of punctate bodies. Cell death was determined by marked cell shrinkage and membrane blebbing, detected by differential interference contrast (**DIC**) microscopy, as in panel (**C-DIC**), accompanied by markedly increased cellular fluorescence, easily distinguishable from flat healthy cells and mitotic cells (one is marked by **m** in panels (**C**) and (**C-DIC**)). Notably, punctate bodies are detectable in both purine-poor and rich media, with some forming even after the shift into purine-rich medium, as for the cell marked * in (**E-F**).

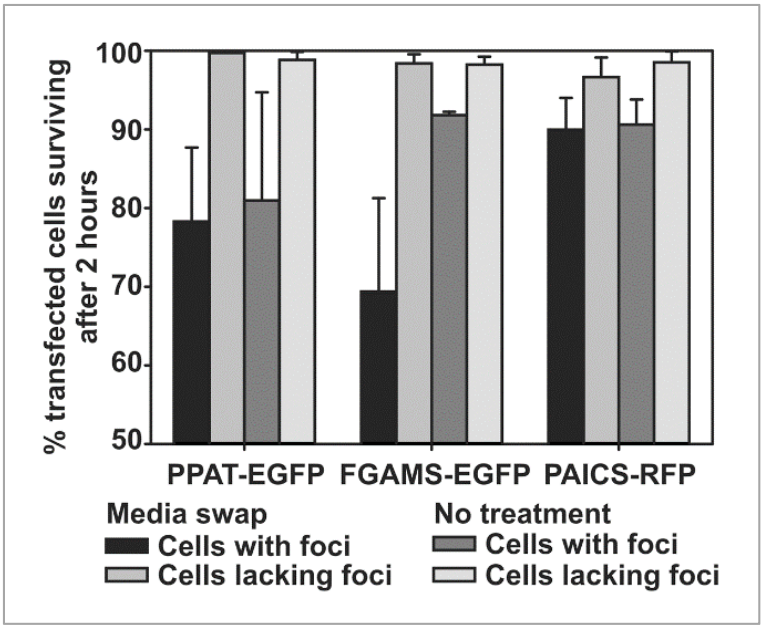

Figure 3-9. Transfected cells exhibiting purine enzyme foci die at a greater rate than transfected cells lacking foci.

Shown here are HeLa cells two hours after either no treatment or after exchanging the growth medium from purine-depleted to purine-rich. All comparisons between cells with and without intracellular foci are statistically significant (p-values ranging from 10^{-4} to $<10^{-16}$). Cell death was measured as marked cell shrinkage and membrane blebbing accompanied by markedly increased cellular fluorescence during time lapse fluorescence microscopy (e.g., as for the example cells in **Figure 3-8**).

3.4.2 Oxidative induction of purine biosynthetic enzyme foci formation

Oxidative stress may induce aberrant disulfide bond formation that can result in improper protein folding, and has also been found to impair proteasome function (Goswami et al. 2006). Application of oxidative stimuli has been shown to promote the aggregation of various disease-associated proteins (Hashimoto et al. 1999, Goswami et al. 2006). Similarly, I found that hydrogen peroxide can induce the

formation of purine biosynthesis enzyme bodies irrespective of the presence or absence of purines in the medium (**Figure 3-10**). Likewise, menadione-induced oxidative stress also led to an increase of foci formation. Pyruvate is known to protect against oxidative stress through direct scavenging of reactive oxygen species, and has previous been shown to prevent oxidation-induced protein aggregation (Aouffen et al. 2004). Cells grown in the presence of pyruvate have lower penetrance of purine enzyme foci formation (**Figure 3-11**), suggesting that oxidative stress at least partially underlie the basal formation of purine enzyme foci.

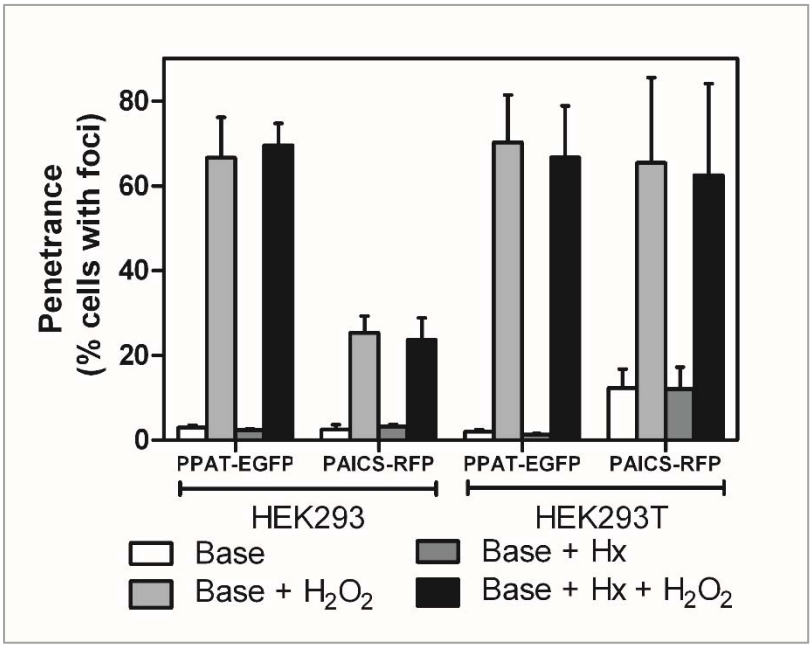

Figure 3-10. Hydrogen peroxide strongly induced purine biosynthetic enzyme foci regardless of hypoxanthine.

Base medium is DMEM supplemented with 10% FBS. As indicated, medium was also supplemented with 1 mM hydrogen peroxide (H_2O_2) and/or 35 µM hypoxanthine (Hx) as described in Methods. For HEK293 cells transfected with, n = 4419, 2652, 3088, 3182 cells per bar. For HEK293 cells transfected with PAICS-RFP, n = 2970, 1944, 1880, 1760. For HEK293T cells transfected with PPAT-EGFP, n = 4537, 2267, 2411, 2947. For HEK293T cells transfected with PAICS-RFP, n = 4612, 3660, 4211, 3760. Bars indicate average +/- 1 s. d. across at least 3 replicates.

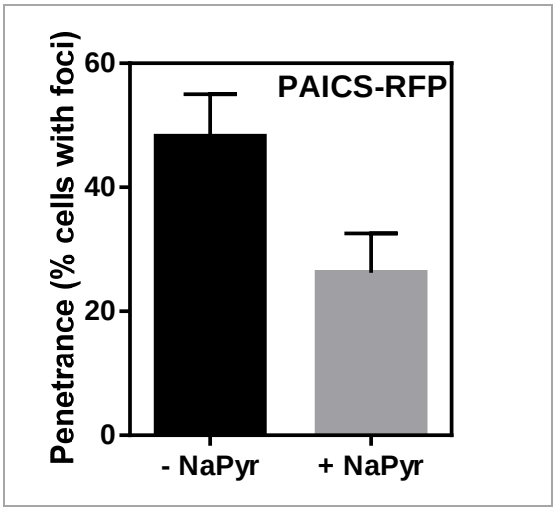

Figure 3-11. Pyruvate reduces the formation of PAICS-RFP bodies

HEK293T cells cultured in DMEM without sodium pyruvate (NaPyr) show nearly double the rate of PAICS-RFP foci formation compared to cells cultured in DMEM with 110mg/L sodium pyruvate. n = 314 and 785 cells. –NaPyr bar represents average +/- 1 s.d. across 2 replicates, +NaPyr bar represents average +/- 1 s.d. across 4 replicates.

3.4.3 Abundance of proteostasis chaperones in purine biosynthetic enzyme foci

The punctate foci appeared to be universally associated with the chaperone HSP70, as measured by co-transfection with a plasmid expressing fluorescently-tagged HSP70 (**Figure 3-12A-I**). Similarly, I observed the chaperone HSP90 to co-localize to punctate foci (**Figure 3-12J-L**). I verified via immunofluorescence that this was not an overexpression artifact, showing that endogenous HSP70 localized with the foci (**Figure 3-13A-C**). As heat shock proteins are highly multifunctional and interact with many proteins in the cell, I further tested whether purine enzyme foci might be aggregates by assaying for their association with ubiquitin. Using immunofluorescence, I observed the co-localization of endogenous ubiquitin to the

foci (**Figure 3-13D-F**). In contrast, control experiments employing only secondary antibody or primary antibody targeting the unrelated enzymes GAPDH or glutamine synthetase showed no such co-localization with the foci (**Figure 3-13G-O**).

Beyond HSP70 and HSP90 co-localization to purine enzyme foci which have been independently reproduced (French et al. 2013), a multitude of additional other proteins not previously implicated in purine biosynthesis have been localized to purine enzyme foci, including Bag2 (binds to ATPase domain of Hsp70 and inhibit its chaperone activity), Bag5 (inhibits ubiquitination of target proteins), Stip1/Hop (adaptor protein that binds both HSP70 and HSP90), p23 (a lesser-understood HSP90 cochaperone), DnaJ-C7 (Hsp40; crucial partners for HSP70 cochaperones), and DnaJ-A1 (Hsp40) (French et al. 2013, Zhao et al. 2013). Many of these proteins are commonly associated with deleterious protein aggregates and global regulation of protein quality control (Broadley and Hartl 2009, Ren et al. 2009, Chai et al. 1999, Uryu et al. 2006, Jana et al. 2000, Schmidt et al. 2002, McLean et al. 2002, Kalia et al. 2011, Shinder et al. 2001, Muchowski and Wacker 2005, Dai et al. 2005, Kalia et al. 2004, Freeman, Toft, and Morimoto 1996). The relatively high rate of co-localization between an individual purine biosynthesis enzyme and HSP70 compared to that among pairs of purine biosynthesis enzymes gives rise to the hypothesis that the purine enzyme foci are non-specific aggregates rather than the long sought-after functional metabolic purinosome complex (**Table 3-4**). If the foci are aggregates, they do not seem to be sorted for lysosomal degradation (**Figure A-3** in **Appendix A**).

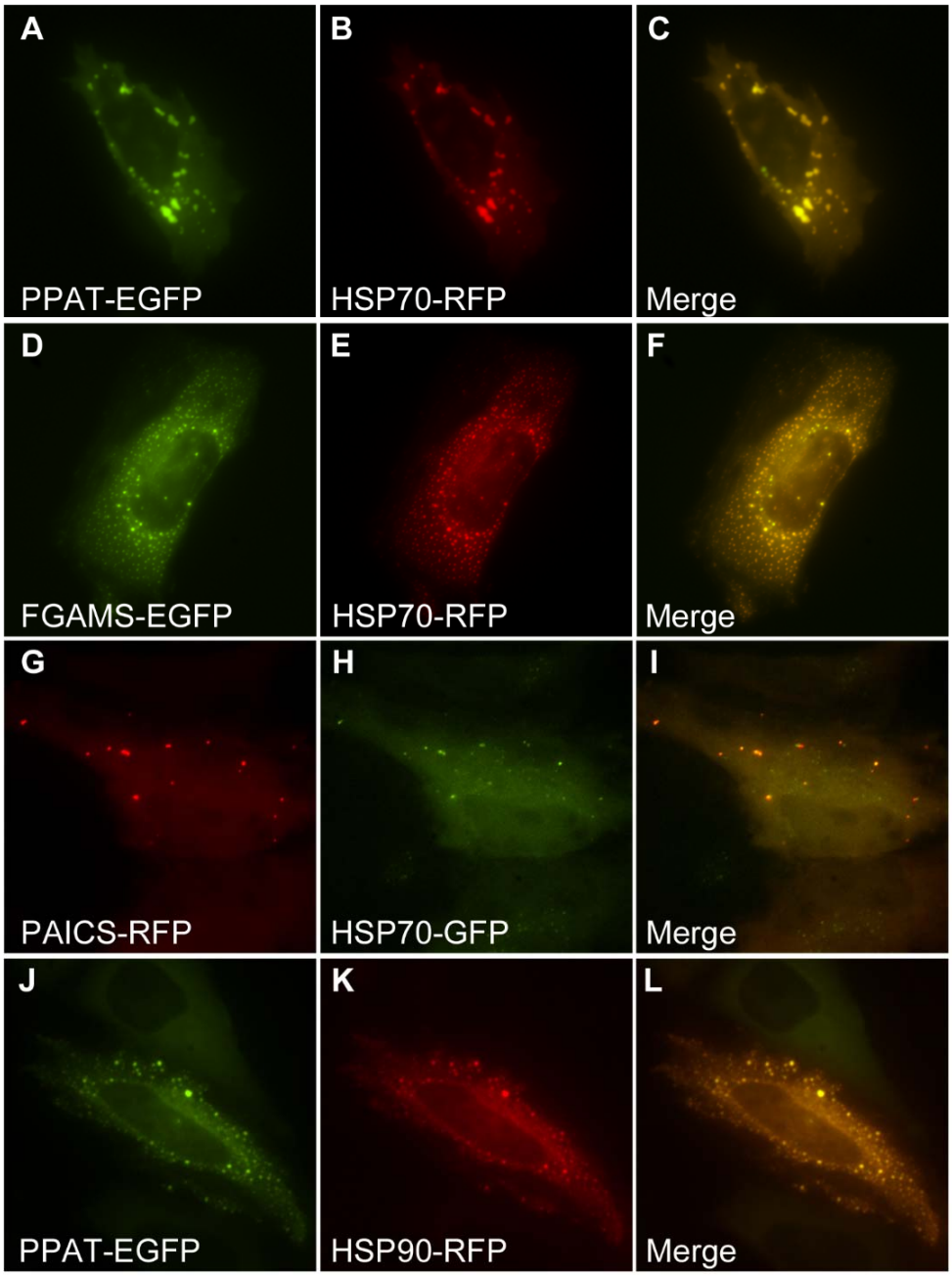

Figure 3-12. Co-expressed HSP70 and HSP90 chaperones marked purine biosynthesis enzyme bodies.

Intracellular bodies formed by (**A**) PPAT-EGFP, (**D**) FGAMS-EGFP and (**G**) PAICS-RFP co-localized with co-transfected (**B**,**E**) HSP70-RFP or (**H**) HSP70-GFP, shown here in HeLa cells. (**C**, **F** and **I**) show merged images. (**J**-**L**) Intracellular bodies were also often observed to co-localize with co-transfected HSP90, shown here with PPAT-EGFP.

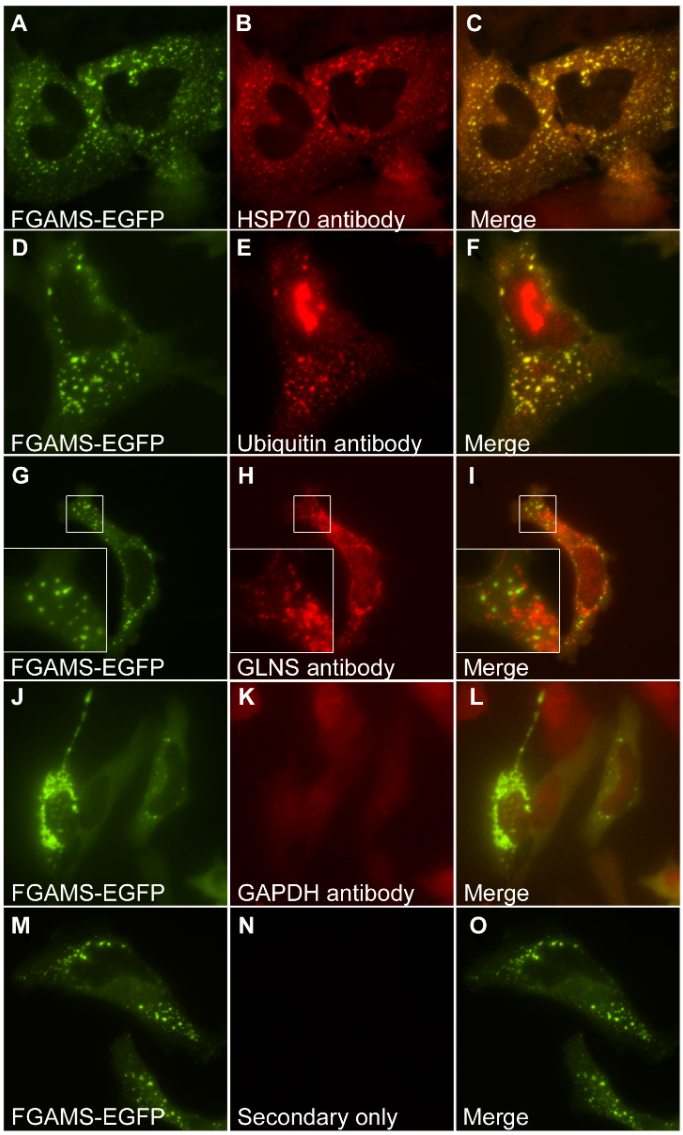

Figure 3-13. Endogenous markers of aggregated proteins associated with intracellular foci of transfected purine biosynthetic enzymes.

(**A-C**) Endogenous HSP70 and (**D-F**) ubiquitin co-localized with bodies formed in cells transfected with FGAMS-EGFP, as assayed using immunofluorescence. Immunofluorescence against endogenous glutamine synthetase (**G-I**) or GAPDH (**J-L**), which are not markers for protein aggregation, confirmed that these proteins did not co-localize with the bodies. (**M-O**) Additional control experiments employing only the secondary antibodies (tested for both secondary antibodies and shown here for Alexa Fluor 594-conjugated goat anti-rabbit) exhibited no positive signal and, with the experiments in panels (**G-L**), ruled out the possibility of non-specific antibody-mediated localization to the bodies.

FGAMS-EGFP and HSP70-RFP						
Experiment	DNA (ug)	Both diffuse	P1 foci P2 diffuse	P1 diffuse P2 foci	Both foci NOT co-localized	Co-localizing foci
6/30/2011	0.8/0.8	385 (94%)	0 (0%)	2 (0.5%)	0 (0%)	21 (5.1%)
6/30/2011	1.6/1.6	349 (90%)	2 (0.5%)	7 (1.8%)	0 (0%)	28 (7.3%)

PPAT-EGFP and HSP70-RFP						
Experiment	DNA (ug)	Both diffuse	P1 foci P2 diffuse	P1 diffuse P2 foci	Both foci NOT co-localized	Co-localizing foci
6/30/2011	0.8/0.8	393 (96%)	0 (0%)	12 (2.9%)	0 (0%)	4 (1%)

FGAMS-EGFP and HSP70 (Native, IF)						
Experiment	DNA (ug)	Both diffuse	P1 foci P2 diffuse	P1 diffuse P2 foci	Both foci NOT co-localized	Co-localizing foci
7/6/2011	1.6 FGAMS	200 (61%)	7 (2.1%)	33 (10%)	1 (0.3%)	85 (26%)
7/6/2011	1.6 FGAMS	223 (69%)	4 (1.2%)	14 (4.4%)	0 (0%)	80 (25%)
7/6/2011	1.6 FGAMS	199 (62%)	7 (2.2%)	24 (7.5%)	0 (0%)	89 (28%)

Table 3-4. Purine biosynthesis foci are enriched for HSP70.

Co-transfection of FGAMS and HSP70 as well as staining endogenous HSP70 in FGAMS-EGFP transfected HeLa cells show high rates of co-localizing puncta, reporting counts and percentages of successfully co-transected (or for the immunofluorescence labelling (IF), transfected) HeLa cells as a function of protein localization. Cells were grown in RPMI + 5% 25kDa dialyzed FBS.

3.4.4 Proteostasis machinery perturbation affects foci formation

Association with HSP70 and HSP90 did not appear to be accidental; short-term treatment of cells with the HSP90 chaperone inhibitor geldanamycin increased accumulation of the punctate foci (**Figure 3-14A**). However, longer term treatment with low doses of geldanamycin is known to induce HSP70 chaperone production via hormesis (McLean et al. 2004), and accordingly longer term treatment with geldanamycin not only inhibited the formation of these foci, but also prevented their basal formation (**Figure 3-14C**) as well as their induced formation by hydrogen peroxide (**Figure 3-14B**).

The association of purine biosynthetic enzyme foci with ubiquitin as measured by immunofluorescence also suggested the possible involvement of the ubiquitin-proteasome proteolytic pathway. Consistent with this involvement, I observed strong induction of foci formation by treatment with the proteasome inhibitor MG-132 (**Figure 3-15**). The induction of foci by blocking protein degradation suggests that they form in cases of excess protein buildup or compromised protein homeostasis, conditions which can lead to intracellular aggregation of proteins.

MG-132 is known to induce restructuring of the distributed vimentin network into a perinuclear cage surrounding the aggresome, a large inclusion body composed of aggregated ubiquinated proteins (Johnston, Ward, and Kopito 1998) (**Figure 3-16A**). Should the punctate foci represent large protein aggregates, they would likely be marked for degradation, consistent with their co-localization with ubiquitin.

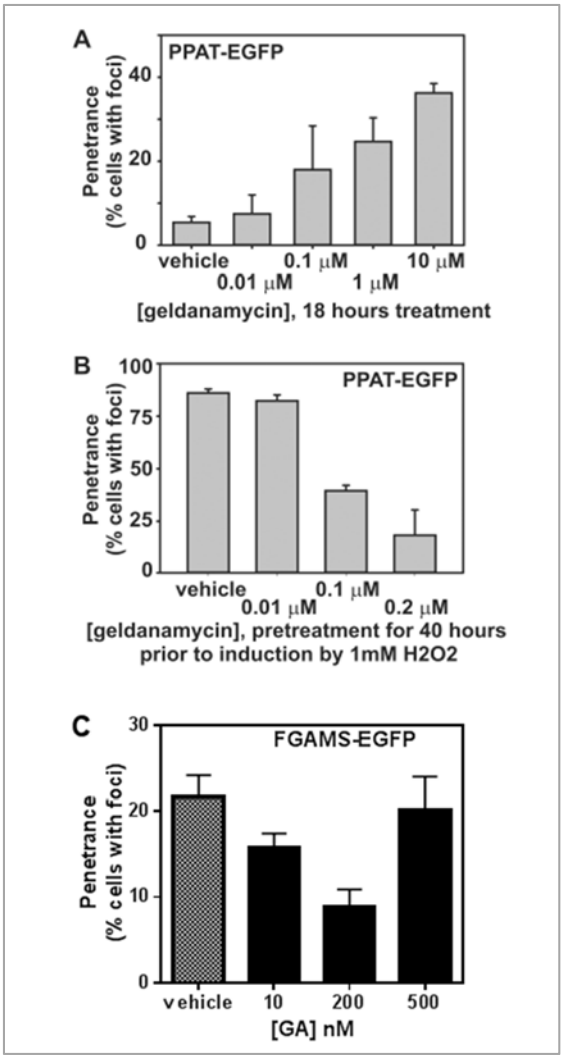

Figure 3-14. Chaperone activity modulated the formation of intracellular foci comprised of purine biosynthetic enzymes

(**A**) Short-term treatment with the HSP90 chaperone inhibitor geldanamycin induced puncta formation in a concentration-dependent manner, shown here for PPAT-EGFP in HEK293T cells. Bars indicate average +/- 1 s. d. across 3 replicates, n = 680, 351, 565, 601, and 616 total cells, respectively. (**B**) Low-dosage geldanamycin pre-treatment—known to induce HSP70 activity (McLean et al. 2004)—suppressed oxidatively-induced puncta formation, shown here for PPAT-EGFP in HEK293T cells. Bars indicate average +/- 1 s. d. across 3 replicates, n = 555, 500, 608, and 601 total cells, respectively. Geldanamycin pre-treatment yielded similar foci-inhibitory results tested in the absence of hydrogen peroxide for FGAMS-EGFP in HEK293T cells (**C**), although a high concentration of geldanamycin (GA) abolished the foci-inhibitory affect; bars indicate average +/- 1 s.d. across 3 replicates, n = 907, 1254, 1222, 620 cells.

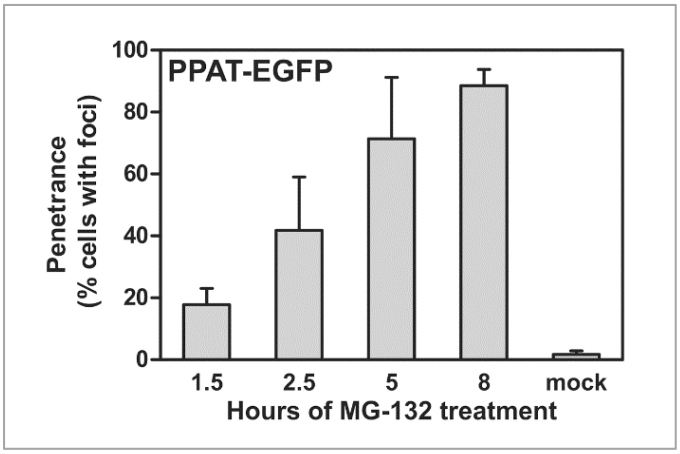

Figure 3-15. Proteasome inhibition induce PPAT-EGFP foci in a time-dependent manner.

HeLa cells treated with 20 µM MG-132 for 1.5, 2.5, 5, and 8 hours showed increasing cells populations with foci. Carrier-only control (DMSO) for 8 hours showed only minimal penetrance, consistent with no treatment (see **Figure 3-1**). Bars indicate average +/- 1 s. d. across 3 replicates, n = 625, 793, 1024, 466, and 397 cells, respectively.

Foci induced with MG-132 were usually accompanied by a single dominant perinuclear purine enzyme body surrounded by vimentin (**Figure 3-16B**), and smaller bodies were occasionally juxtaposed to vimentin filaments (**Figure 3-16C**). Thus, collectively, these data indicate that purine enzyme foci associated strongly with known markers of aggregation, and that their formation is modulated in a manner consistent with aggregation (Howard et al. 2007, Sampathu et al. 2003, Kim et al. 2002, Proctor and Lorimer 2011).

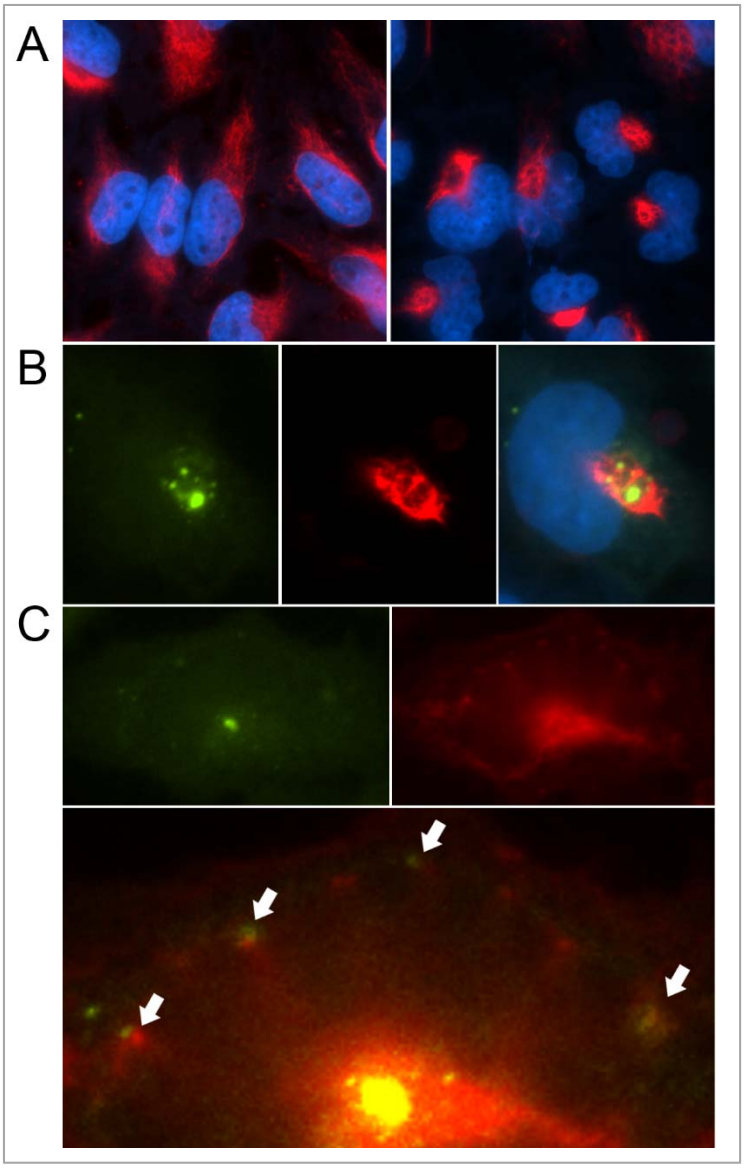

Figure 3-16. Proteasome inhibition induces aggresome-localization of purine enzyme foci.

(A) The intermediate filament vimentin (left, untreated) collapses into a perinuclear ring surrounding the aggresome upon treatment with MG-132 (right, 20 µM for 8 hours). (**B**) MG-132-induced PPAT-EGFP foci are predominated by a single large body at the site of the aggresome, encased by vimentin. (**C**) Smaller foci (left) distributed across the cytoplasm are frequently associated with vimentin filaments (right) extending from the aggresome; examples are marked by arrows (overlay, bottom) (20 µM MG-132 for 5 hours). For all panels, red: anti-vimentin antibody; green: PPAT-EGFP; blue: nucleus stained with Hoechst.

3.5 Undirected investigation of protein interactions suggests sequential enzyme pairing

The repertoire of current studies on purine biosynthesis enzyme interactions rely heavily on fluorescence microscopy of transfected recombinant tagged constructs, which are known to influence expression behavior of foci (Landgraf et al. 2012). To access the existence of the putative purinosome as a true native multi-enzyme complex, we used undirected, unbiased method orthogonal to fluorescence microscopy including (1) a high-scale proteomic-level mass spectroscopy profiling of the proteome to search for native protein complexes, and (2) mass spectroscopy coupled with immunoprecipitation (IPMS) of FGAMS-EGFP to identify associated endogenous proteins. The high-scale protein complex search described in **Section 3.5.1** was done in collaboration with Pierre Havugimana and Cuihong Wan from Andrew Emili's lab at the University of Toronto, who carried out all the fractionation experiments, and with Blake Borgeson, who optimized the computation pipeline for data processing and map assembly. The IPMS experiments described in **Section 3.5.2** were done in collaboration with Daniel Boutz, who collected and analyzed the data from the experiments that I designed and conducted. We find no strong evidence for a native purinosome and no consistent associations of endogenous proteins to FGAMS-EGFP.

We devised a method of identifying conserved protein complexes in an untargeted manner by combining scaled-up biochemical fractionation with mass spectroscopy profiling (Havugimana et al. 2012). To assemble candidate protein complexes, we employed multiple separation techniques to resolve proteins followed

by computational filtering and clustering methods. Using this methodology, we not only recapitulated many previously reported complexes (**Figure 3-17**) but also inferred hundreds of new putative complexes, including a partial purinosome (**Figure 3-18**).

3.5.1 Endogenous enzymes may associate in pairs but not in full complex

Purine enzyme foci average 560 nm (ranging 200-900 nm) (French et al. 2013). One might expect that such large structures, even if partially disassociated, might still retain protein-protein interaction with other foci constituents. Some of the largest known protein complexes such as the human proteasome and 80S ribosome which are no larger than 50 nm (Jones et al. 2014, da Fonseca and Morris 2008) are robustly recapitulated in our biochemical fractionations (**Figure 3-18A**). Five of the six enzymes of the *de novo* purine biosynthesis pathway were predicted as a complex via our mapping pipeline (**Figure 3-18B**). However, the elution profiles of the pathway proteins indicated a pattern of protein "pairing", or the possibility of multiple sub-complexes, as opposed to a single complex consisting of at least all six enzymes (**Figure 3-19**). The pattern of mutually exclusive pairwise associations at each sequential step is consistent with the idea that the purine biosynthesis enzymes associate to exchange unstable intermediates and may facilitate metabolic channeling, and is inconsistent with the idea that the pathway coalesces into a single stable macromolecular complex.

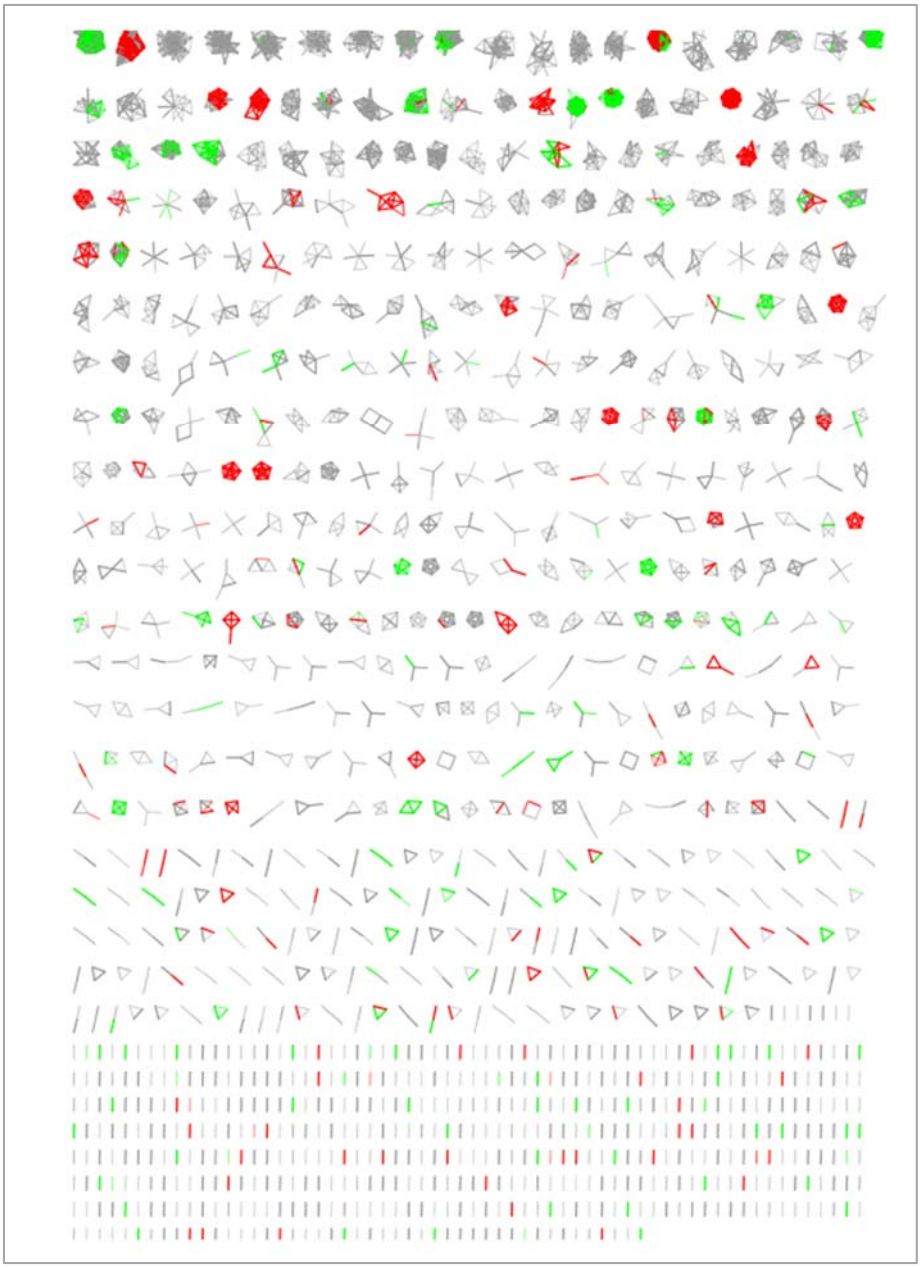

Figure 3-17. All inferred conserved animal complexes from large scale fractionations.

Each cluster represents a conserved animal complex inferred from co-elution data, with nodes presenting proteins connected by weighted edges presenting strength of association. Curated human interactions from CORUM (Ruepp et al. 2010) are shown colored: red associations were given as training set, and green associations were not given as training set (thus served as validation of our prediction set). Grey edges and associations represent novel predicted interactions and new complexes.

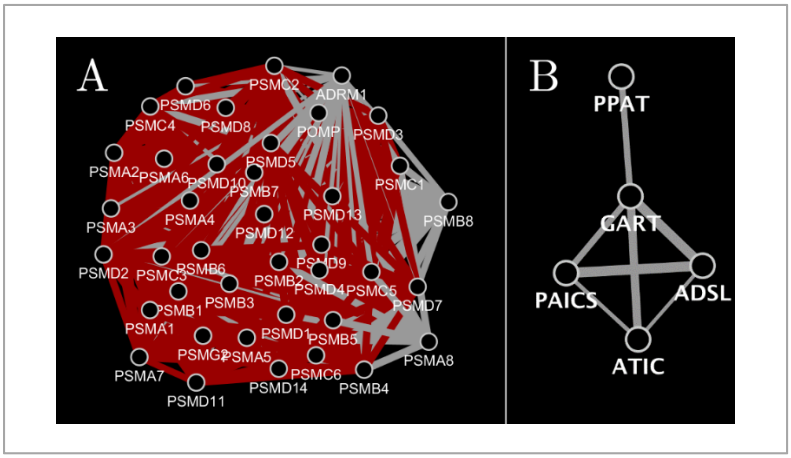

Figure 3-18. Recapitulated known proteasome complex and inferred novel purine biosynthesis complex.

(**A**) Complex discovery method recapitulated the proteasome, with proteins shown as circular nodes connected by weighted edges presenting the calculated probability of association. Some known interactions from CORUM associations (red edges) were used to train the clustering algorithm. (**B**) 5 of the 6 human purine biosynthetic pathway enzymes clustered into a single module representing a "novel" putative complex.

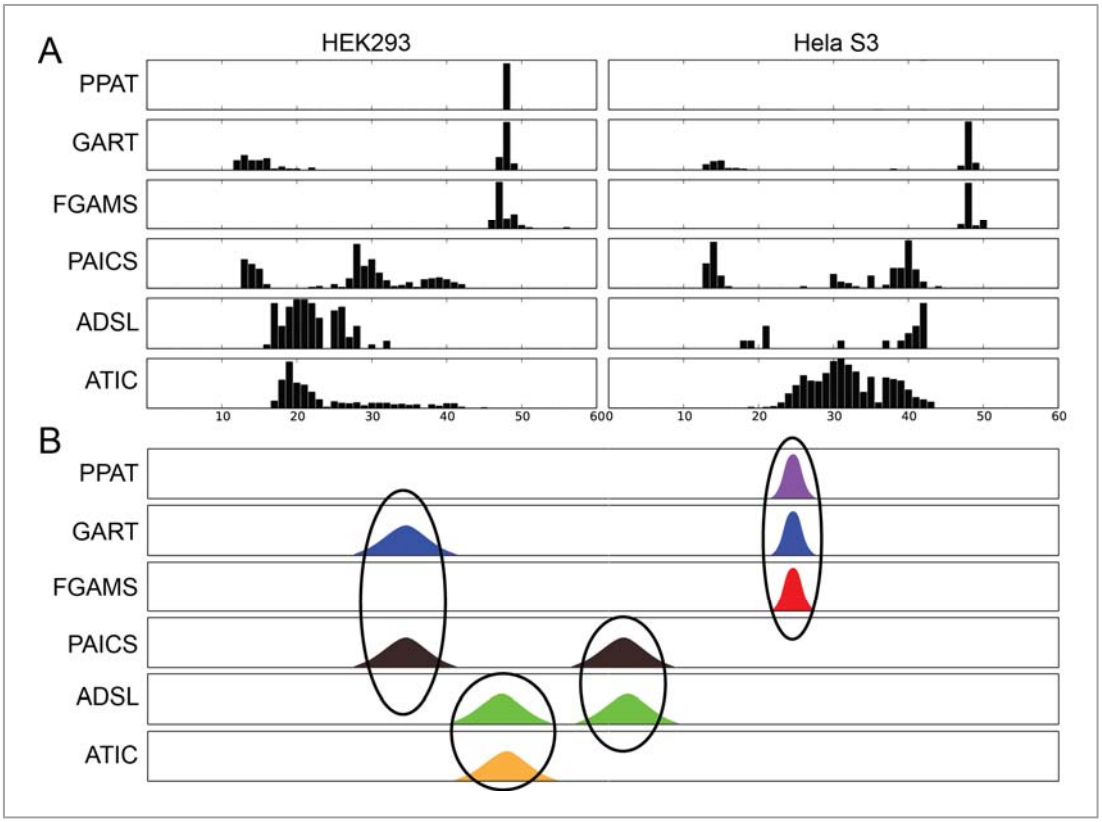

Figure 3-19. Elution profile for purine biosynthesis pathway enzymes.

(**A**) Actual elution profiles for dual heparin-ionic exchange fractionation of partial cellular extracts from HEK293 and HeLa S3 cells. The six human purine biosynthesis pathway enzymes appear in different fractions plotted along the x-axis. Each enzyme's normalized mass spectral counts are plotted along the y-axis. Peak height is relative to the maximum count observed for each particular individual protein. (**B**) Cartoon representation of the elution profiles highlighting mutually exclusive putative interactions at each step of the sequential pathway. The chemical reaction steps are carried out by the enzymes in the following order: PPAT → GART → FGAMS → GART → PAICS → ADSL → ATIC. Methods used are those described in Wan et al. (2015).

3.5.2 No evidence of complex by IPMS of FGAMS-EGFP

Immunoprecipitation of FGAMS-EGFP subject to mass spectroscopy identification of co-precipitated proteins (IPMS) revealed a lack of high-confidence

interacting proteins (**Figure 3-20**) which would have formed the purinosome complex. The notion that FGAMS-EGFP foci are homogenous is consistent with the fact that protein aggregates are highly homogenous (Rinas and Bailey 1992). A large fraction (>60%) of the immunoprecipitated FGAMS-EGFP should be derived from the diffuse expression state, suggesting that even the diffuse protein does not strongly associate with any interacting partner proteins. Consistent with this observation, independent IPMS of transfected FGAMS-MYC also failed to identify strong associations with other enzymes, including the purine biosynthetic pathway enzymes (French et al. 2013).

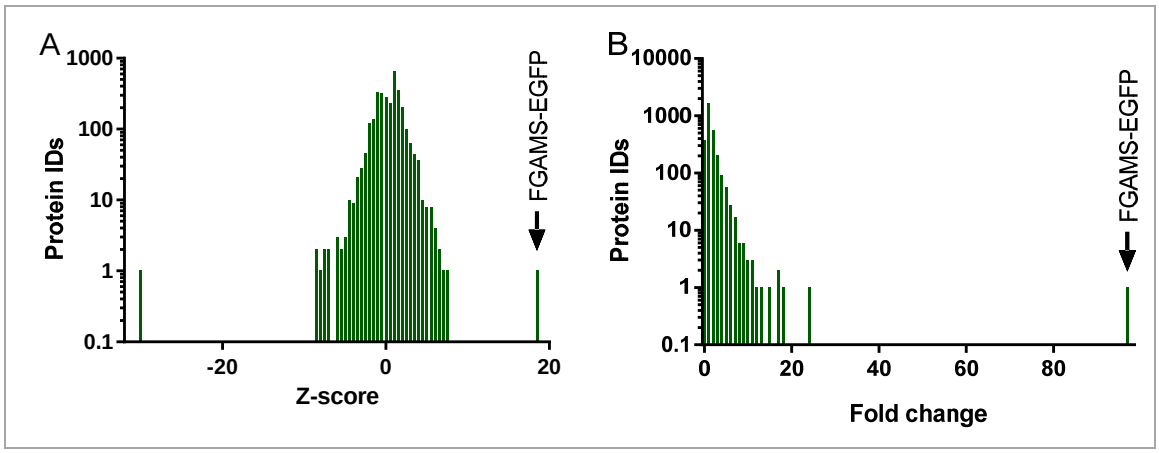

Figure 3-20. IPMS of FGAMS-EGFP show no strongly enriched interacting proteins.

Immunoprecipitated FGAMS-EGFP from HeLa cells grown in RPMI + 10% dialyzed FBS was subject to mass spectroscopy identification. Immunoprecipitated GFP from cells grown in identical condition was used as control. (**A**) Combined Z-scores with most significant protein IDs being: 18.35 FGAMS-EGFP and 7.15 transketolase, the latter presumably a contaminant. The significant Z-score for antibody (-30.23) resulted from detection of abundant antibody used in the IP. (**B**) Fold-change enrichments suggest FGAMS-EGFP does not strongly associate with other proteins. Protein IDs with highest fold-change are 97.49 FGAMS-EGFP and 23.7 dynein, the latter presumably a contaminant. All values were calculated from PSMs (see Methods) from 3 experiments.

3.6 Discussion and Remarks

Multi-enzyme complexes often engage in various forms of substrate channeling, in which sequential pathway enzymes "hand off" intermediate metabolic products amongst each other rather than release them into bulk solution. The advantages of such complexes include improved efficiency and optimized usage of short-lived intermediates.

One such possible complex involves the enzymes for *de novo* purine biosynthesis. Termed the "purinosome," the complex's existence has long been suggested by circumstantial evidence. The search for direct *in vivo* evidence of the purinosome culminated in the discovery of intracellular punctate foci formed in part by purine biosynthesis enzymes. An et al. (2008) argued, based in part upon the now-debated evidence for purine dependency, that these were functional complexes. However, the new discoveries described in this dissertation support a model where these foci behave in a manner unlike that expected for functional purinosomes and instead have numerous features one might expect of simple protein aggregates or stress bodies. Specifically, the transiently overexpressed purine biosynthetic enzymes form intracellular bodies at frequencies that correspond to each enzyme's predicted aggregation propensity. The resultant bodies are marked by ubiquitin and heat-shock chaperones, and they were inducible using a number of nonspecific cellular insults.

In this discussion section, I will address issues with current literature that models purine enzyme foci as metabolically active associations (purinosomes), and demonstrate that the same experiments do not disentangle the bodies from possibly

being general protein aggregation. Then I will address the issues that emerge with characterizing the bodies with evidence obtained with a single method.

3.6.1 Conciliating seemingly contradictory observations

Reports of seemingly contradicting observations of purine biosynthetic enzyme foci by different groups or in differently designed experiments calls a need to reconcile the apparent contradictions.

3.6.1.1 "Purine-dependency" experiments were not purine-specific experiments

Functionality of cellular bodies is strongly suggested when they are metabolically inducible and form reversibly. Based in part upon evidence for purine dependent formation and dissolution of purine enzyme foci, they were argued to be functional complexes. However, as stated above, the medium used to observe purinosomes by An *et al.* lacked bulk nutrients up to 50X larger than purines, and was supplemented with only half of the serum concentration compared to that for purine-rich medium (An et al. 2008). Unsurprisingly, a switch to this nutrient-depleted medium may be stressful for some cells (e.g., it was reported that HTB-125 cells did not survive in this "purine-depleted" medium (An et al. 2008)). It is therefore possible that this stress contributed to the visible aggregation of recombinant purine enzyme constructs.

Although HeLa cells seem to fare better than HTB-125 cells in "purine-depleted" medium (An et al. 2008), claims that "purine-depleted" medium promotes assembly of puncta while "purine-rich" medium promotes disassembly are not supported by more quantitative measurements of puncta penetrance across HeLa cells. In experiments replicated exactly or replicated with optimized media formulations, I did not see differences in puncta penetrance between different growth conditions. While the foci formation rate for HeLa cells grown in "purine-rich" medium was not reported in An et al., their HeLa cells grown in "purine-depleted" medium showed foci rates similar to those I report. In particular, puncta penetrance in four out of the six pathway enzymes is agreeably very low (~5-8%), and the opposite of might be predicted if they were functional entities, given that the *de novo* purine biosynthesis pathway is also active and detectable in "purine-rich" mammalian cell culture (Yamaoka et al. 1997, Yamaoka et al. 2001).

Interestingly, in my own work and in other independent labs, no differences in foci rates between cells grown in the two media are observed (personal communication with Dr. Michael Rosen, UTSW; personal communication with Dr. Donald Anderson, UTSW), although one group saw increased foci rates only when the "purine-depleted" medium was further filtered with charcoal (Field, Anderson, and Stover 2011), a method that powerfully and non-specifically depletes organic compounds. These discrepancies may have risen from unintentional differences in cell line sub-culturing practices or imaging techniques used by different groups. For example, immediately prior to live imaging, one group washed the cells for several

times in variously supplemented buffered saline solutions (An et al. 2008) for which both reasons and implications are unclear.

3.6.1.2 "Purine-dependency" claim is not supported by purine-specific experiments.

While there is no convincing published evidence that levels of purines in the medium alone affect foci formation, it has been shown that levels of purines in the medium do *not* affect the foci formation. Specifically, the observation that purine enzyme foci remained robust after specifically adding an exogenous purine source, hypoxanthine, back to "purine-depleted" medium (An et al. 2008) strongly argues that foci formation is not controlled solely by purine levels in the growth medium. In this regard, azaserine (a purine antagonist) also has no effect on the formation of purine biosynthesis enzyme foci (An et al. 2008). The lack of purine-specific dependencies is inconsistent with the hypothesis that the inducible foci represent true purinosomes or functional purine biosynthesis enzymes.

3.6.1.3 Purine-independent foci cycling

Furthermore, An et al. (2008) showed cycling of foci assembly and disassembly by exchanging the "purine-rich" for "purine-depleted" media and vice-versa; however, no control was performed. **Figure 3-4** showed cycling could also be achieved by simply removing and replacing with the same medium, arguing that the

induction and disappearance of bodies may be due to the shock and adaptation to the brief mechanical or oxidative stress associated with media swap.

3.6.1.4 Heat shock chaperone perturbation

The association of heat shock chaperones (particularly of HSP70) with purine enzyme foci has been independently established (Zhao et al. 2013, French et al. 2013), although there are opposing views on the nature of this association. As shown earlier in this manuscript in support of purine enzyme foci as aggregated proteins, HSP70 induction by geldanamcyin reduces foci prevalence, even in the presence of foci induction by peroxide. This observation supports the bodies as aggregating proteins which may be chaperoned into their solubilized with the aid of HSP70. Contrarily, HSP70 induction by geranylgeranylacetone induces purine enzyme foci (French et al. 2013), arguing that the foci are functional multi-enzyme complexes that form with the aid of HSP70. However, at the concentration used to induce foci, geranylgeranylacetone has also been shown to also induce the unfolded protein response and the accumulation of unfolded/misfolded proteins within the ER (Endo et al. 2007). The complexities of cell culture and imaging make it difficult to ascertain which precise cellular response is observed, thus the apparent induction of purine enzyme foci may have instead resulted from this ER stress.

3.6.1.5 Effects of other pharmacophores

It is argued that the pharmaceutical induction of foci supports their functionality. However, the aforementioned effects of CK2 inhibitors on foci formation are fully consistent with an alternative interpretation that the foci are intracellular aggregates. CK2 has been shown to have crucial roles in clearing general protein aggregation and has also been implicated as an anti-stress factor (Watabe and Nakaki 2011, 2012). Importantly, CK2 has many protein substrates, and its inhibition influences a wide array of cellular activity (Bibby and Litchfield 2005, Litchfield 2003). For example, CK2 inhibition is known to induce apoptosis (Slaton et al. 2004, Ahmad et al. 2008, Trembley et al. 2009). The chemical inhibitors of CK2 are also rather nonspecific (Pagano et al. 2008) and are themselves associated with increased apoptosis and increased production of reaction oxygen species (Schneider et al. 2009). Thus, possible signal transduction via CK2 does not strongly constrain the mechanism underlying purine enzyme foci formation.

The claim that the formation of foci by purine biosynthesis enzymes can be controlled by the addition of Gαi agonists or antagonists (Verrier et al. 2011) may again be more simply explained by aggregation, rather than functional complex formation. Indeed, downstream Gαi targets (e.g., the PI3K/Akt pathway) include regulators of various stress-related cellular responses, such as cell survival and protein synthesis, which will also influence protein homeostasis and aggregation. Notably, Gαi signaling has been implicated in autophagy (a mechanism known to clear aggregated proteins) in human cell lines (Petiot et al. 1999, Ogier-Denis et al. 1996),

and a molecule which sustains activated Gα has been reported to inhibit the formation of autophagic vacuoles (Kadowaki et al. 1994).

These data are fully consistent with a simpler hypothesis that cell stress, which may be induced by a variety of factors, might be the trigger for non-specific foci formation.

3.6.1.6 Foci constituents disagree with biochemically-identified purine enzyme associations

The constituents of purine enzyme foci as measured by microscopy are in many ways in disagreement with biochemically-captured purine enzyme protein–protein associations. As previously shown through partial co-purification, two folate metabolism enzymes (serine hydroxymethyltransferase 1 and methylenetetrahydrofolate dehydrogenase 1) associate with the folate-utilizing purine biosynthesis enzymes GART and ATIC (Caperelli et al. 1980, Smith et al. 1980). GART's catalytic activity actually requires interaction with methylenetetrahydrofolate dehydrogenase 1 or its analog (Smith et al. 1980, Smith et al. 1981). However, these two folate metabolism enzymes were found to be excluded from the fluorescent purine enzyme bodies (An et al. 2008, French et al. 2013).

3.6.2 Reliance on transient transfection of recombinant proteins and insufficient examination of endogenous proteins

With one exception (Baresova et al. 2012), all literature evidence for purine enzyme foci has thus far relied on transiently expressed recombinant fluorescent protein fusion constructs. This may be a consequence of reported difficulties and possible artifacts surrounding immunofluorescent labeling of native purine biosynthesis enzymes (An et al. 2008, Zhao et al. 2013). For example, it was observed that while endogenous GART behaved similarly to the GART fusion construct in "purine-depleted" medium, its endogenous behavior did not correspond to that observed for its fusion construct in "purine-rich" medium (An et al. 2008). Previous efforts using cells stably transfected with GART did not yield visible foci (Gooljarsingh et al. 2001), raising the question of whether effects of body assembly were promoter-driven artifacts or possibly artifacts due to transient transfection-associated stresses.

3.6.2.1 Questionable physiological relevance

These results also raise interesting questions regarding how growth conditions and cellular physiology impact the formation of single protein aggregates in general, especially proteins altered from their native state, for example whose levels may no longer correspond to normal, endogenous levels or which are expressed as fusion proteins. While it may be possible that purine enzyme foci may shift to stress body-like protein complexes with increasing levels of protein expression or varying methods of cellular insults, we note that in general, the formation of purine enzyme foci did

not vary strongly amongst the different cell lines, and at least for PAICS-RFP, we observed the recombinant enzyme to be expressed at levels very similar to the endogenous proteins (**Figure A-4** in **Appendix A**).

Due to difficulties with immunofluorescent targeting of endogenous purine enzymes and inability to biochemically capture a native intact purinosome, it is unclear whether the purine enzyme fusion constructs function similarly to endogenous enzymes. Others have observed that fluorescent protein fusions are subject to aggregation and form intracellular foci (Garcia-Mata et al. 1999, Landgraf et al. 2012). In fact, the observation that "cells transfected with Hsp90 or Hsp70 alone yielded a diffuse staining pattern" (French et al. 2013) strongly suggests that native purine biosynthesis enzymes do not form such bodies. Rather, the Hsp foci arise only as a function of the transfected recombinant purine biosynthetic enzymes.

In a similar vein, I have noted numerous unexpected yet striking differences on a protein's subcellular localization and/or interaction partners between similar experiments employing different commonly used protein tags or transfection methods (some cases are presented in the appendix, and is a theme within the next chapter describing TTC4). Here, the study of purine enzyme foci is also a case study in how the methods used to query intracellular bodies potentially influence outcomes. The sole reliance on microscopy with fusion proteins, and the absence of other physical demonstrations of the purinosome as a whole, argues for caution in interpreting the nature of such bodies.

3.6.3 Conclusions

The idea of a purinosome has many compelling features. Assorted evidence ranging from metabolic flux and channeling considerations to pairwise kinetic or physical interactions observed between particular biosynthetic enzymes supports the idea of some form of physical association between purine biosynthetic enzymes. Thus, there is a substantial body of literature suggesting the existence of the purinosome. In the preceding chapters, I addressed confounding issues surrounding the interpretation of punctate foci containing purine biosynthetic enzymes. Although current observations do not fully negate the possible formation of functional purinosomes, they also do not substantiate that foci formed by recombinant purine biosynthetic enzymes must be the purinosome. A far simpler explanation is plausible: that the purine biosynthetic enzymes can aggregate under conditions of cellular stress or recombinant expression. Indeed, it is well known that many other intracellular enzymes form aggregates under similar conditions, and purine biosynthetic enzymes are not known to be special in this regard. This discrepancy between purine enzyme foci interpretations highlights the need for caution regarding reliable methodologies for observing punctate body formation and determining their function and physiological relevance.

3.7 Materials and Methods

3.7.1 Abbreviations of protein names

PPAT, phosphoribosyl pyrophosphate amidotransferase; **TrifGART**, the trifunctional enzyme glycinamide ribonucleotide (GAR) synthetase, GAR transformylase, and aminoimidazole ribonucleotide synthetase; **FGAMS**, formylglycinamidine ribonucleotide synthase; **PAICS**, the bifunctional enzyme carboxyaminoimidazole ribonucleotide synthase and succinylaminoimidazolecarboxamide ribonucleotide synthetase; **ADSL**, adenylosuccinate lyase;**ATIC**, the bifunctional enzyme aminoimidazolecarboxamide ribonucleotide transformylase and IMP cyclohydrolase; **GFP**, green fluorescent protein; **RFP**, red fluorescent protein; **EFGP**, enhanced GFP; **GAPDH**, glyceraldehyde 3-phosphate dehydrogenase; **GLNS**, glutamine synthetase.

3.7.2 Cloning

The PPAT-EGFP and FGAMS-EGFP expression plasmids were generously provided by An and Benkovic (An et al. 2008). The ADSL-EGFP expression plasmid was constructed by Gateway cloning into the FGAMS-EGFP expression plasmid, modified to introduce the Invitrogen pDEST47 Gateway cassette in place of FGAMS. All other cDNAs were obtained from the human ORFeome collection (OpenBiosystems) and cloned using Gateway cloning into either the pcDNA-DEST47 plasmid (Invitrogen) for carboxy-terminal GFP-tagged expression clones or

the pTagRFP-N plasmid (Evrogen; modified to introduce the Invitrogen pDEST47 Gateway cassette) for carboxy-terminal RFP-tagged expression clones.

3.7.3 Cell culture

To replicate conditions reported by An et al. to give rise to purine enzyme foci, HeLa cells obtained from the American Type Culture Collection (ATCC) were cultured at 37°C with 5% CO_2 in the following medium as previously described (An et al. 2008): Purine rich medium: MEM medium supplemented with 10% FBS (fetal bovine serum) and 50 µg/mL gentamycin sulfate (Invitrogen); purine-depleted medium: RPMI medium supplemented with 5% dialyzed FBS and 50 µg/mL gentamycin. FBS was dialyzed against 100-fold volumes of 0.9% NaCl at 4°C for 2-4 days using a 25 kDa molecular weight cut-off (MWCO) dialysis membrane with daily exchange of the dialysis solution.

As a test of alternate cell lines and matched purine-rich and purine-depleted media formulations, HEK293 or HEK293T cells obtained from ATCC were cultured at 37°C with 5% CO_2 in the following media: (1) Optimized purine-rich medium: DMEM medium supplemented with 10% FBS and 35 µM hypoxanthine, or (2) Optimized purine-depleted medium: DMEM medium supplemented with 10% dialyzed FBS. To remove purines but retain other important factors, FBS was dialyzed against 100-fold volumes of 0.9% NaCl at 4°C for 2-4 days using a 1 kDa MWCO dialysis membrane with daily exchange of the dialysis solution (Yamaoka et al. 1997, Yamaoka et al. 2001).

3.7.4 Transfections

One day prior to transfection, HeLa cells were plated in 6-well glass bottom plates in either purine-rich or purine-depleted medium lacking antibiotics. HEK293 or HEK293T cells were plated in 96-well glass bottom plates in the optimized purine-rich medium lacking antibiotics. Plasmids were transfected into cells using Lipofectamine 2000 (Invitrogen) and Opti-MEM reduced serum medium (Invitrogen) according to the manufacturer's instructions. The transfection medium was replaced with fresh purine-rich or purine-depleted medium 5 hours after transfection. Immediately before imaging, cells were washed once with their respective growth medium. Cells were imaged in their respective growth medium to minimize stress imposed by nutrient shifts. We note that the imaging medium therefore contained phenol red, which could potentially reduce the signal-to-noise ratio, but in practice, we did not experience problems in detecting and visualizing foci. Live cells were imaged ~20-24 hours after transfection.

For MG-132 experiments, HeLa cells were plated in 96-well glass bottom plates in DMEM medium supplemented with 10% FBS. Due to the toxic nature of MG-132, we opted to use a less toxic transfection reagent than Lipofectamine 2000, and therefore, we transfected the PPAT-EGFP-expressing plasmid into cells using Fugene HD (Roche Applied Sciences) and Opti-MEM reduced serum medium (Invitrogen) using a 2:4.5 DNA to transfection reagent ratio following the manufacturer's protocol. The transfection medium was replaced with fresh DMEM+10% FBS about 5 hours following transfection.

3.7.5 Immunofluorescence

For immunofluorescence experiments, cells were fixed with 3.7% methanol-free formaldehyde freshly diluted from 16% stock (28908, Thermo Scientific) at 37°C for 15-20 minutes, blocked with 5% goat serum in PBS-T buffer for 30-60 minutes at room temperature, then incubated with primary antibody overnight at 4°C. Cells were washed with PBS buffer, then incubated with secondary antibody for 1 hour. Primary antibodies used: HSP70 (ab5439, Abcam), ubiquitin (ab7780, Abcam), GAPDH (sc-32233, Santa Cruz Biotech.), glutamine synthetase (sc-9067, Santa Cruz Biotech.) Secondary antibodies used: Alexa Fluor 594-conjugated goat anti-mouse (Invitrogen), Alexa Fluor 594-conjugated goat anti-rabbit (Invitrogen). All antibodies were used at the manufacturer's recommended concentrations. We additionally tested the protocol using 2% methanol-free formaldehyde fixation or 2% goat serum in PBS-T block at the appropriate steps to find no appreciable differences.

Tested antibodies that proved unsuitable or inconclusive for immunofluorescence included the anti-PPAT antibodies sc-101892 (Santa Cruz Biotech.) and ab71340 (Abcam), the anti-PAICS antibody sc-16150 (Santa Cruz Biotech.), and the anti-TrifGART antibody H00002618-M01 (Novus Biologicals).

3.7.6 Western blot

For western blotting, cells were grown in DMEM media containing 10% FBS and transfected as described above with 1.6 µg or 200 ng DNA (HeLa or HEK cells,

respectively), in regular 96-well plates (HEK293 and HEK293T) or 6-well plates (HeLa). Cells were washed with cold PBS buffer then extracted with Laemmli sample buffer containing 5% beta-mercaptoethanol. Total protein was separated on a 4–15% Mini-PROTEAN TGX precast gel (BioRad) and blotted with PVDF membrane. Primary antibodies used: anti-PAICS (HPA035895, Sigma) and anti-GAPDH (sc-32233, Santa Cruz Biotech.). Secondary antibodies: goat anti-rabbit IgG $F(ab')_2$-HRP (sc-3837, Santa Cruz Biotech.) and goat anti-mouse IgG $F(ab')_2$-HRP (sc-3697, Santa Cruz Biotech.). Membranes were scanned with an ImageQuant LAS 4000 (GE Healthcare).

3.7.7 Fluorescent cell microscopy

Live cells were imaged using a Nikon Eclipse TE2000-E inverted microscope inside a chamber maintained at 37°C and 5% CO_2. Fixed cells were imaged at room temperature without CO_2 supplement. Images were acquired using a Photometrics Cascade II 512 camera and Nikon Plan Apo 40x/0.95, 60x/0.95, 60x/1.40, or 100x/1.40 objectives. GFP detection was accomplished using a ET490/20x excitation filter (Chroma Tech.), ET525/36 m emission filter (Chroma Tech.) and 89100 bs dichroic (Chroma Technology), and RFP detection was carried out using a ET555/25x excitation filter (Chroma Tech.), ET605/52 m emission filter (Chroma Tech.) and 89100 bs dichroic (Chroma Tech.) The same excitation and emission conditions were used for each GFP variant. We did not detect notable fluorescence emission bleedthrough or photobleaching with these settings. Filter wheels (Sutter

Lambda 10-3), motorized stage (Prior H117), and image acquisition were driven by NIS Elements AR Imaging software.

3.7.8 Cell counting

Cell counts measuring penetrance of foci reflect apparently healthy adherent cells only. Mitotic cells (identified as rounded cells with a visible rod-like central structure spanning the cell's DIC image accompanied by a corresponding decrease of fluorescent signal; see cell m in **Figure 3-8C** and **C-DIC** for an example) and dead cells were not counted for measuring penetrance of punctate body formation. (For the survival rate analyses only, dead HeLa cells were selected based on dramatic cell shrinkage, membrane blebbing, and intense fluorescence, e.g. as seen for the cells marked by # in **Figure 3-8**. We note that it is nonetheless possible that only a subset of cell death events—those accompanied by marked visible cellular reorganization—are counted by this assay.) Cells displaying any morphology of foci were counted, regardless of number of foci per cell or ranging sizes of those foci. We observed a spectrum of morphologies which were dynamic; for example, over the course of unperturbed growth, we observed both increases and decreases in the numbers of foci per cell. Future work may address and characterize the different morphologies and kinetics of foci.

3.7.9 Drug treatments

For all drug treatments, drugs were diluted to their final concentrations in pre-warmed cell growth medium then added to cells for the duration of the treatment. Hydrogen peroxide was added to cells at a concentration of 1 mM for 0.5-1.5 hours before imaging. For short-term treatment, geldanamycin was added to cells at the indicated concentrations at the time of post-transfection medium replacement, and the cells were incubated for 18 hours. For long-term treatment, geldanamycin was added to cells at the indicated concentrations 16 hours prior to transfection, for a total incubation time of 40 hours. MG-132 was added to cells at a final concentration of 20 µM for the indicated time intervals. Cells were fixed at each MG-132 time point with 3.7% formaldehyde for 15-20 minutes and washed 3 times with PBS before imaging.

3.7.10 Immunoprecipitation and mass spectrometry

HeLa cells in "purine-depleted" media (RPMI + 5% 25kDa dialyzed FBS) were transfected with either FGAMS-EGFP or EGFP using Lipofectamine 2000, followed by media replacement after 5 hours. The next day, cells were washed once in PBS then trypsinized and resuspended in media to inactivate trypsin, spun down at 1000-1500 RPM for 5-10 minutes, and then resuspended in 500 µL freshly made ice cold lysis buffer (10mM Tris-HCl pH8.0, 10mM KCl, 1.5mM MgCl2, 0.5mM DTT, 1X protease inhibitor I (Calbiochem)) for 5-10 minutes before submitting the sample to glass dounce homogenization (~30-50 thrusts per sample with pestle size "B"). Cell lysate was incubated on rotation at room temperature with Protein A

Dynabeads (Thermo Scientific) pre-washed with PBS and bound to anti-GFP antibody (Sigma G1544). After washing the Dynabead-antibody-antigen complex with PBS, proteins were eluted from the beads in 50 µL 2,2,2-trifluoroethanol (TFE, Sigma) at 60°C and then stored at -20°C. For mass spectrometry sample preparation, each TFE sample was reduced with freshly made dithiothreitol and heated to 55°C for 45 minutes. After cooling to RT, iodoacetaminde was added to 550mM and incubated in dark for 30 minutes at RT. Sample was then diluted with 50mM Tris with 2mM $CaCl_2$ to 5% TFE and subject to 4-5 hour 1ug trypsin incubation at 37°C. Tryptic digest was quenched with 1% vol/vol formic acid and sample was then stored at -80°C. Sample was then vacuum dried to 60-100 µL and resuspended in Buffer C (95% H_2O, 5% acetonitrile, 0.1% formic acid). Samples were cleaned on C18 tips (Thermo), resuspended in 100 µL Buffer C, then subject to LC-MS (Orbitrap Velos) for 300 minute runs. Fold changes and Z-scores were calculated using pseudo-peptide spectral matches (PSM + 1) normalized against total PSMs (calculated from the sum of total pseudoPSMs) within each sample.

The fold change of protein a is denoted FC_a and calculated as follows:

$$FC_a = \frac{a_{\mathrm{IP}} / \sum_{p \in S_{\mathrm{IP}}} p}{a_{\mathrm{mock}} / \sum_{p \in S_{\mathrm{mock}}} p}$$

Where S_{IP} and S_{mock} are the sets of IP and mock pseudo-PSMs, respectively. Fold change and Z-scores were calculated as an average of three experiments.

Chapter 4: Summary

Large-scale protein interaction mapping and subcellular localization screening provide a wealth of protein characterization data from which researchers can use as starting points for discovering protein function. The first part of my work (**Chapter 1**) focused on building a new platform for large-scale microscopy on human cells that would enable future large-scale functional genetics assays using expression constructs, RNA interference, or chemical methods to induce perturbation. The later parts of the manuscript attempted to provide deeper characterization on a few inferred protein complexes and foci assemblies derived from our large-scale screens, specifically focusing on foci formed by fluorescently tagged purine biosynthesis enzymes.

Observations of cytoplasmic puncta or foci may indicate the compartmentalization of proteins to specific organelles such as lysosomes, endosomes, or autophagosomes. However, the majority of novel cytoplasmic foci-forming proteins discovered in our protein localization screen are not known to be localized to any organelle (Narayanaswamy, Levy, et al. 2009), and the reasons for their foci formation remains largely unexplained. I presented three possible explanations in **Chapter 2**, including functional compartmentalization for improved enzymatic regulation and catalytic efficiency, storage depots for inactive enzymes poised for rapid mobilization, or aggregates of misfolded and/or dysfunctional proteins as a result of stress. Many of these proteins are generally insoluble in foci-forming conditions and thus their macrostructure assemblies may be stress-related

aggregation, as suggested by a proteome-wide mass spectrometry survey (O'Connell et al. 2014). On the other hand, many foci- and fiber-forming metabolic proteins, such CTP synthase and glutamine synthetase, were characterized more in-depth, with their macrostructural assemblies *in vitro* reconstituted and physiologically relevant roles of such assemblies proposed. Such example cases suggest exciting explanations underlying other foci and fiber structures.

A complex formed by enzymes of the purine biosynthesis pathway compartmentalized for catalytic efficiency (purinosome) is an attractive hypothesis. Yet, the full purinosome was not found in our large scale protein complex mapping efforts by fractionation (Wan et al. 2015). Although, sequential pairwise interactions were observed within our dataset (Wan et al. 2015). There is not yet convincing data for the existence of the functional multi-enzyme purinosome, and I showed in **Chapter 3** that the purinosome is not necessarily supported by the observation of foci formed by purine biosynthesis pathway members. I demonstrated that this observation is likely due to stress-related aggregation and explored the possibility that such foci are formed as a by-product of fluorescent tagging. Notably, it is important when interpreting experimental observations to consider alternative explanations, especially to take into account the methodology used to collect observations. Formation of proteinaceous macrostructures such as foci and fibers are frequently seen and are in many cases biologically relevant for compartmentalizing cellular processes, as explained in **Chapter 2**; however, as pointed out many times throughout the manuscript and further supported by several case observations shown later in **Appendix B**, they may arise due to fluorescent tagging of the proteins of

interest. Along the same lines, it is also possible that co-localization of fluorescently tagged proteins into the foci may be the result of aggregation; it is known that many fluorescent proteins oligomerize and aggregate (Snapp 2005, Snapp 2009, Krasowska et al. 2010), driving efforts to produce non-oligomerizing mutagenized variants for researchers to use (Baird, Zacharias, and Tsien 2000, Zacharias et al. 2002, Campbell et al. 2002, Rizzo, Davidson, and Piston 2009, Costantini et al. 2015). Despite these efforts, many "monomeric" derivatives still produce aggregating artifacts (Landgraf et al. 2012, Shaner et al. 2013), especially when local concentrations are high such as within membranes or within cellular compartments in which the pH is non-optimal for the fluorescent protein. An in-depth case of tagging effects on the characterization of TTC4 is explored later in **Appendix C**.

Appendices

Appendix A: Additional figures to accompany Chapter 3

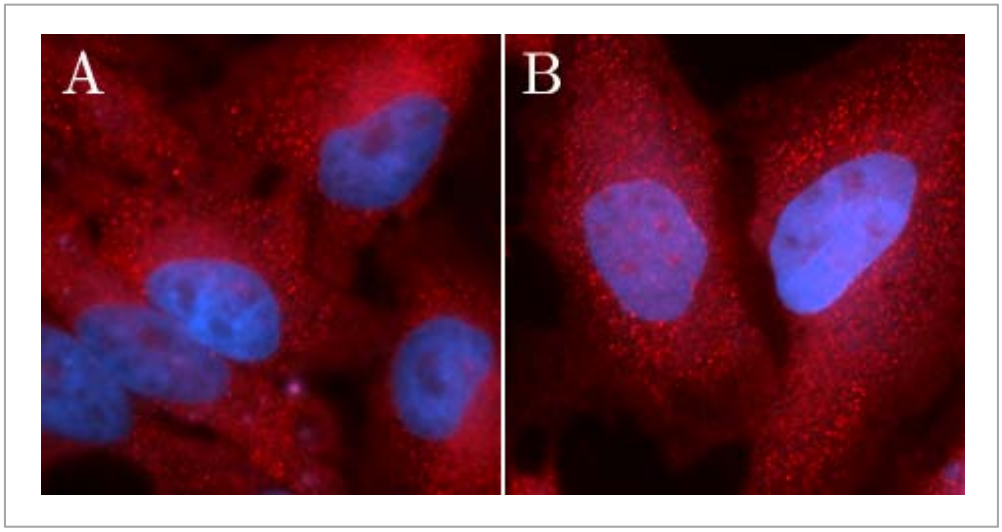

Figure A-1. Immunofluorescence of endogenous PPAT shows high prevalence of puncta regardless of purine availability.

HeLa cells cultured in DMEM + 10% FBS (**A**) before switching to "purine-depleted" media (RPMI + 5% dialyzed FBS) (**B**) for 16 hours. Cells are fixed with 3.7% formaldehyde and permeabilized with 0.2% triton. Antibodies used: anti-Atase (PPAT, sc-101892, Santa Cruz Biotech, 1:100), Alexa Fluor goat anti-rabbit 594 (Invitrogen), nuclear stained with Hoechst 1:1000.

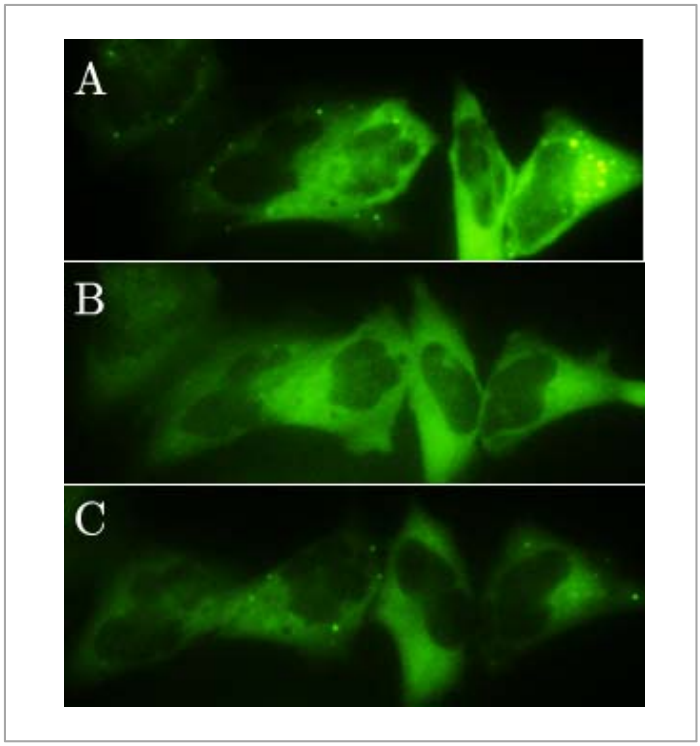

Figure A-2. FGAMS-EGFP foci dynamics are independent of media formulation, but respond to media replacement.

HeLa cells grown in "purine-rich" medium with FGAMS-EGFP foci (**A**) may dissolve foci after normal culture condition incubation for 2 hours without perturbation (**B**). However, upon media replacement with identical "purine-rich" media, foci re-appears within 30 minutes (**C**).

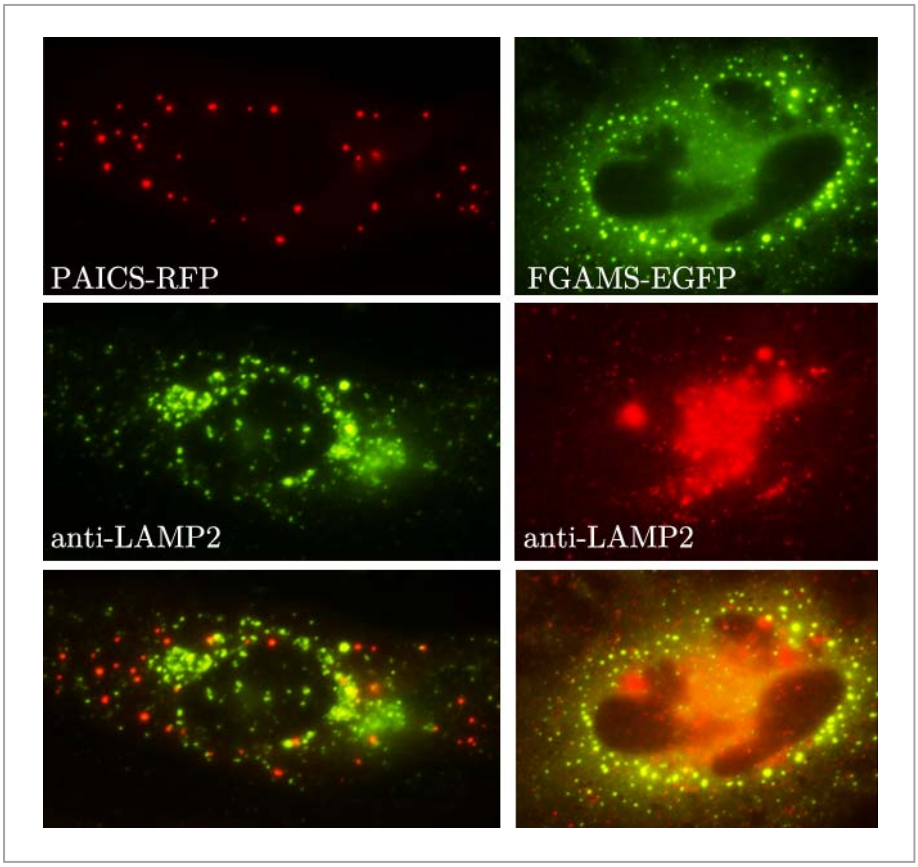

Figure A-3. Purine enzyme foci do not co-localize with lysosomes.

HeLa cells expressing PAICS-RFP (**left column**) or FGAMS-EGFP (**right column**) grown in "purine-depleted" media (RPMI + 5% 25kDa dialyzed FBS) are treated with 0.2mM leupeptin for 19 hours before fixation and immunofluorescence against LAMP2 (anti-LAMP2 antibody, ab25631, Abcam). PPAT-EGFP foci also do not co-localize with foci marked by LAMP2 (data not shown). **Bottom row** shows merged images.

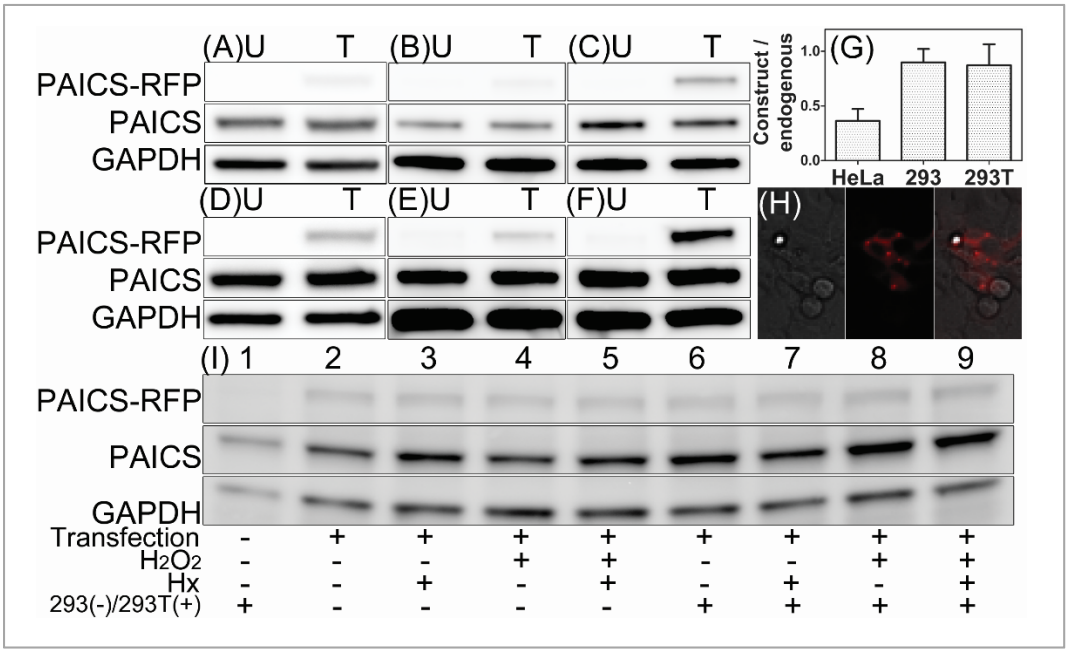

Figure A-4. Representative expression levels of PAICS-RFP transfected constructs in comparison to the endogenous PAICS protein.

HeLa (**A,D**), HEK293 (**B,E**) and HEK293T (**C,F**), untransfected (U) or transfected with PAICS-RFP (T) under representative growth conditions are probed with anti-PAICS antibody. (**D-F**) Representative variation of construct expression between independent transfections. Quantified expression level ratios of PAICS-RFP construct to endogenous PAICS (**G**) across all cells, adjusted with respective transfection rates (ranging from 6-32%) measured by microscopy, show comparable expression (within two-fold) of the construct and endogenous enzyme in transfected cells, under conditions (see Methods) in which PAICS-RFP foci are observed. Expression ratios are calculated as (PAICS-RFP Western blot intensity)/(endogenous PAICS Western blot intensity * transfection efficiency). Bars indicate average +/- 1 s. d. across at least 3 replicates. (**H**) Representative HEK293 cells corresponding to the bar in (**G**) and showing visible foci. (**G**) Equivalent levels of the expressed PAICS construct and the corresponding endogenous protein is in agreement with independent experiments observing transiently transfected purine constructs (personal communication with Dr. Michael Rosen, UTSW). (**I**) Expression levels of PAICS-RFP are not strongly altered by hydrogen peroxide and/or hypoxanthine treatment for HEK293 (lanes 2-5) and HEK293T cells (lanes 1, 6-9).

Appendix B: Examples of tag-influenced protein localization

Although immunofluorescence and fluorescent-protein tagging show high correlation for protein localization, it is the case that for a substantial portion (20%) of proteins, the two methods give disparate results (Stadler et al. 2013). Landgraf et al. (2012) have shown that the behavior of recombinant tagged Clp proteases can be influenced by the type of fluorescent protein it is expressed with. In this Appendix section, I catalog examples of a variety of proteins in which differential tagging unexpectedly produced different protein subcellular localization in yeast (**Figure B-2**) and human cells (**Figure B-1**), and some of these differences seem to be independent of the tag, but accrued through the process of tagging (**Figure B-1**). I also show an example of how altering transfection method and staining technique can also drastically alter a protein's subcellular localization (**Figure B-3**).

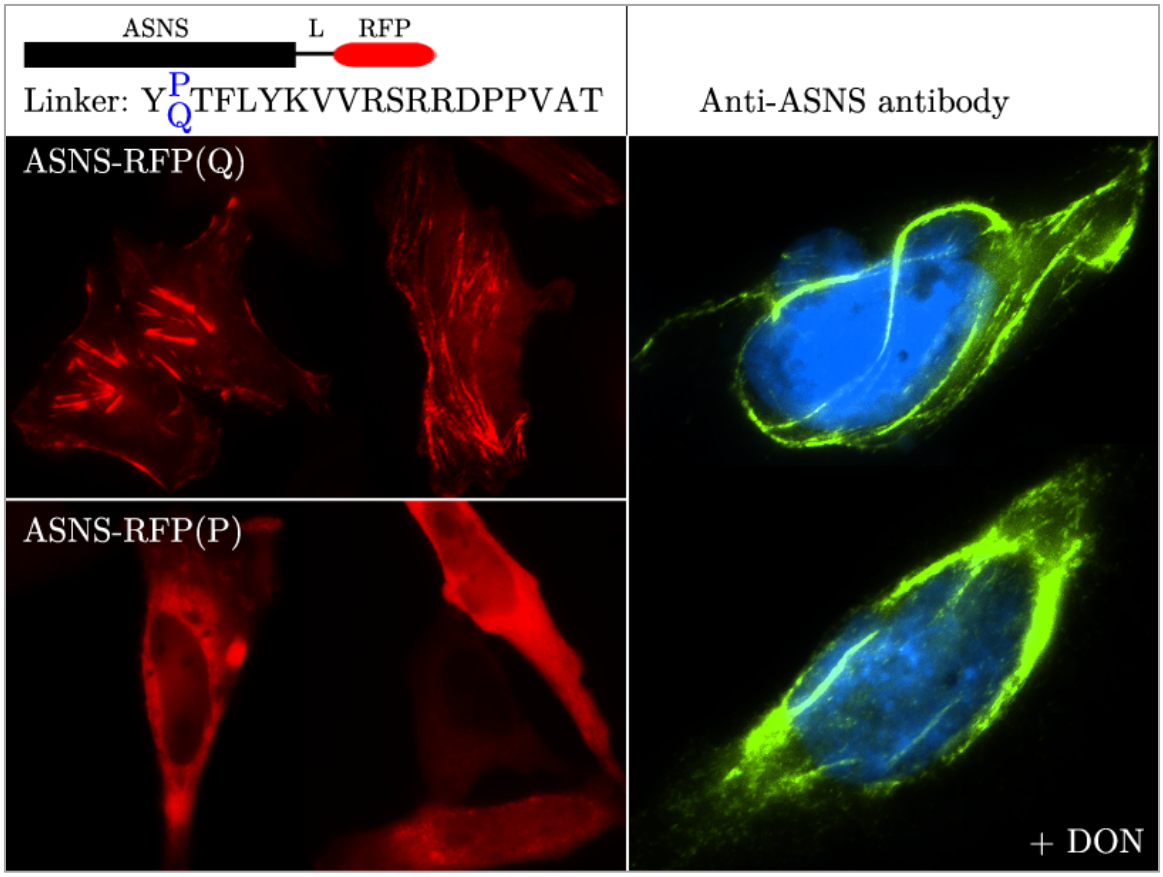

Figure B-1. Asparagine synthetase behavior is altered by a single mutation in linker.

Left: A spontaneous mutation (cca (P) → caa (Q)) in the linker region between asparagine synthetase (ASNS) and red fluorescent protein (RFP) tag results in drastically different subcellular phenotype of the tagged protein. ASNS cDNA was obtained from the human ORFeome collection (OpenBiosystems) and cloned into a modified pTagRFP-N plasmid (Evrogen; modified to introduce the Invitrogen pDEST47 Gateway cassette for carboxy-terminal RFP-tagged expression clones). The P→Q mutation within the linker region lies within the clonase-reactive (*att*L and *att*R) sites and was presumably introduced during the Gateway LR clonase reaction. **Right**: HeLa cells treated with 2mM DON for 24 hours are probed with anti-ASNS antibody (sc-376151, Santa Cruz Biotech.), showing filaments. Secondary antibody: Alexa Fluor goat anti-mouse 488. DON treatment is previously used to modulate CTP synthase filaments (Ingerson-Mahar et al. 2010).

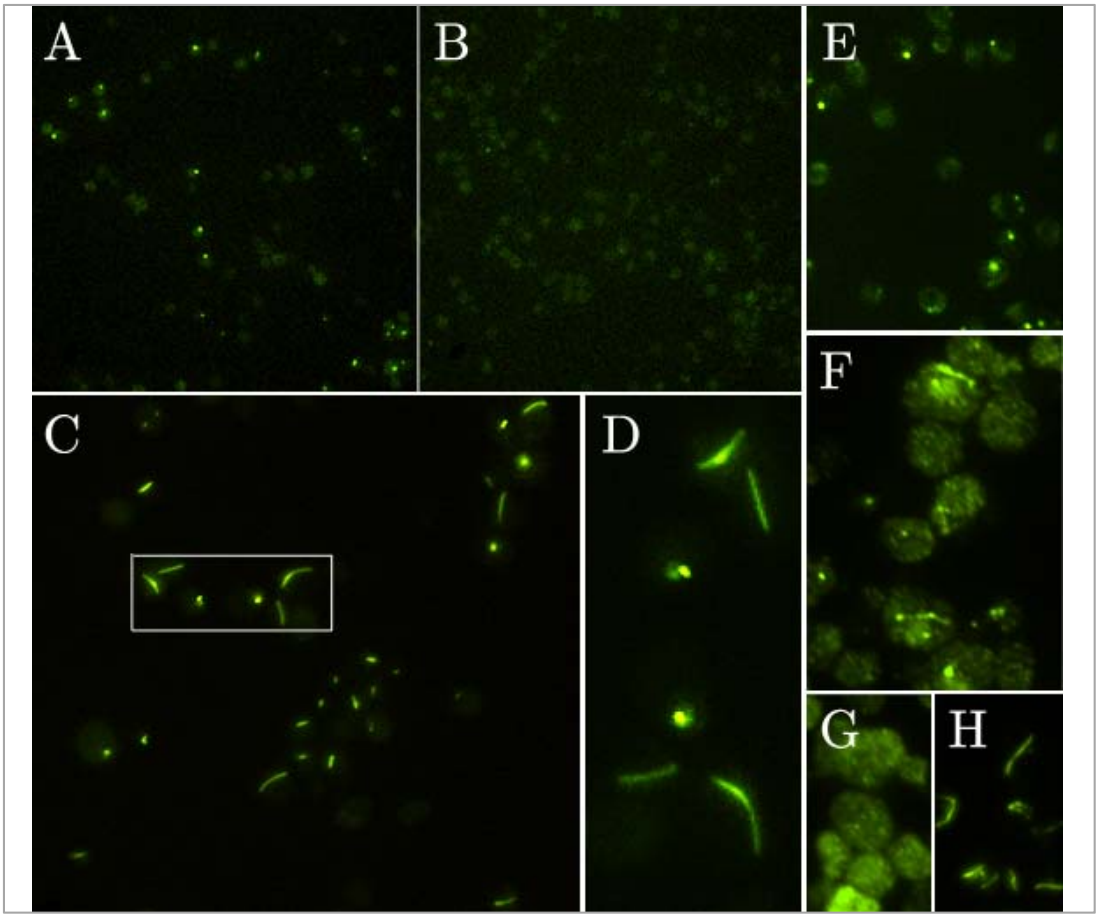

Figure B-2. Fluorescent label-dependent behavior of Ade4 in yeast.

Ade4 genomically labeled with CFP (**A**) or GFP (**B**) were grown in YPD for 6 days. The CFP-labeled foci are more prominent than the GFP counterparts. Camera exposures were identical and LUTs scaled equally. (**C**) Ade4p-CFP formed fibers that are noticeable in a small subset of the population at 5 days in YPD but becomes prominent within the population at an extended 19 days (19 days is shown). A z-stack maximum intensity projection is displayed. (**D**) A magnification of the boxed region in (**C**). Contrarily, Ade4p-GFP maintains a punctate focus throughout extended growth in YPD, examined after up to 21 days (**E**, 21 days is shown). (**F**) Haploid yeast with one copy Ade4-HA showed a mixture of puncta and fiber at 5 days in YPD (although only foci were present in longer growths up to 25 days). 5 days is shown. Structures were not an immunofluorescence artifact, compared to lack of such in (**G**) similarly tagged and probed heterologous haploid Ade17-HA strain. Yet they were not as prominent as the robust fiber forming Ura7p-MYC in a heterologous haploid strain (**H**). In all cases, cells were grown in YPD at 30°C with continuous shaking. Unique puncta morphologies and behaviors were exhibited by other tested tags, such as Ade4p-mCherry (data not shown). Images of stationary phase Ade4p-TAP and Ade4p-YFP can be found in Narayanaswamy, Levy, et al. (2009).

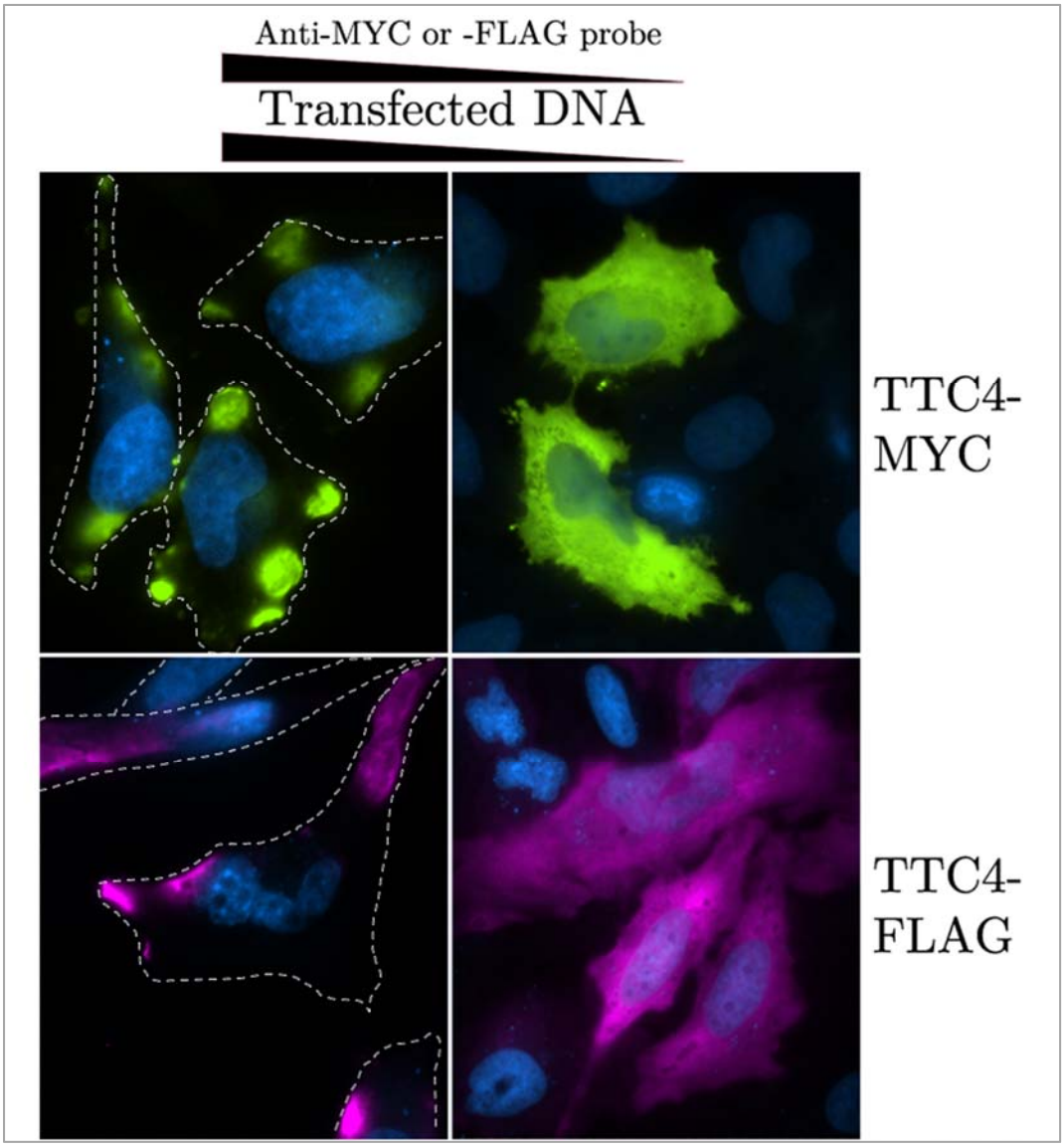

Figure B-3. Transfected DNA and probe concentration affects TTC4 localization.

In a 96 well plate, (**Left; top**) 200ng TTC4-MYC was transfected with 0.6uL FuGENE HD and probed with anti-Myc antibody (mouse, sc-40, Santa Cruz Biotech) at 1:200. The cell boundaries are drawn in. (**Left; bottom**) 200ng TTC4-FLAG was transfected with 0.6uL FuGENE HD and probed with anti-FLAG antibody (mouse, F1804, Sigma) at 1:200. The cell boundaries are drawn in. (**Right; top**) 40ng TTC4-MYC was transfected with 0.12uL FuGENE HD and probed with anti-Myc antibody (1:2000). (**Right; bottom**) 40ng TTC4-FLAG was transfected with 0.12uL FuGENE HD and probed with anti-FLAG antibody (mouse, F1804, Sigma) 1:2000. In all panels, nuclei were stained with Hoechst 1:1000. Secondary antibodies used: Alexa Fluor goat anti-mouse 488 and Alexa Fluor goat anti-mouse 647 (Invitrogen).

Appendix C: Multi-dimensional characterization of TTC4

In **Chapter 3**, I pointed out that the study of foci characterized by purine biosynthetic enzymes is a case study in how the methods used to query intracellular structures potentially influence outcomes. This Appendix section presents TTC4 as another case example of a protein whose characterization has proven difficult due to different outcomes depending on the method of investigation. Prior experiments investigating TTC4's subcellular localization and interacting protein partners have solely relied on the overexpression of recombinant tagged versions of the protein (Dmitriev et al. 2009, Crevel, Bennett, and Cotterill 2008, Dmitriev et al. 2014, Huttlin et al. 2015).

The various methods that we used to investigate TTC4 include: tag-free, label-free methods including (1) mass spectrometry, (2) RNA sequencing, (3) RNA interference, and (4) structural homology prediction; and tag-dependent and label-dependent methods including (5) immunofluorescence and (6) immunoprecipitation with multiple antibodies, and (7) recombinant tagging with various epitopes and (8) fluorescent proteins. Surprisingly, many potential TTC4 characteristics uncovered by these techniques seem to be unique to each employed technique. Such findings urge for caution and critical examination of the biochemical methods before accepting results that came from experiments employing said methods.

For this section, I would like to thank the authors of Wan et al. (2015), and in particular, Kevin Drew, who conducted the TTC4 structural homology model

construction and assisted in data interpretation. Fan Tu designed and conducted all frog experiments. Taejoon Kwon kindly provided the frog mRNA expression data. John Wallingford generously supported all work on frogs. For the TTC4 knockdown transcriptome sequencing experiments, I thank Jessica Podnar and Gabby Huerta at the Genomic Sequencing and Analysis Facility who aided in cDNA library preparation and sequencing, and especially Anna Battenhouse who aided immensely in downstream data processing and analysis. Maria Person and Andre Bui at the Proteomics Facility greatly facilitated mass spectrometry data acquisition. My contributions to the work in this section include the design and execution of all cell culture-based experiments, and processing, analyzing, and interpreting the data.

C1. Introduction

TTC4, or tetratricopeptide repeat (TPR)-containing domain 4, contains a degenerate structural motif identified in many scaffolding proteins that mediate protein-protein interactions (D'Andrea and Regan 2003). TTC4 consists of 3 tandem-repeats of the 34 amino acid residue TPR motif. TPR motif-containing proteins typically serve as scaffolds for the assembly of various multi-protein complexes, including the anaphase promoting complex, peroxisomal import receptor, NADPH oxidase complexes, and HSP70-Hop-HSP90 complex (Alvira et al. 2014, Scheufler et al. 2000, Zeytuni and Zarivach 2012). TPR-containing proteins are involved in a variety of biological processes, such as cell cycle regulation, transcriptional control, mitochondrial and peroxisomal protein transport, neurogenesis, protein folding (D'Andrea and Regan 2003), and ciliogenesis and cilia maintenance (Xu et al. 2015).

C2. Possible TTC4 association with retromer

C2.1 Biochemical co-fractionation of TTC4 with retromer

The motivation for characterizing TTC4 originated from the proteome-wide mapping of conserved physical association networks based on large-scale biochemical fractionations (Wan et al. 2015), in which we found endogenous TTC4 across multiple species strongly associated with VPS26, VPS29, and VPS35 (**Figure C-1**). VPS26-VPS29-VPS35 is a well-established complex known as the retromer (Hierro et al. 2007). The retromer is an endosomal protein sorting complex that is conserved across eukaryotes and functions in endosome-to-trans-Golgi network cargo trafficking. Recently discovered retromer-interacting proteins such as the WASH complex and Wntless/MIG-14 cargo extends its role further in endosome-to-plasma membrane sorting and signaling event regulation (Seaman 2012). Retromer association with endosomes and its transport by molecular motors requires the presence of sorting nexin proteins SNX1, SNX2, SNX5, and SNX6 (Hong et al. 2009, Wassmer et al. 2007, Rojas et al. 2007), which were, along with TTC4, also mapped to the same complex in our work. Importantly, we also saw conservation of these inferred physical interactions preserved across multiple systems, including worm, mouse, fly, and human (**Figure C-2**). Thus, we predicted with high confidence that TTC4 is a novel member of the retromer complex.

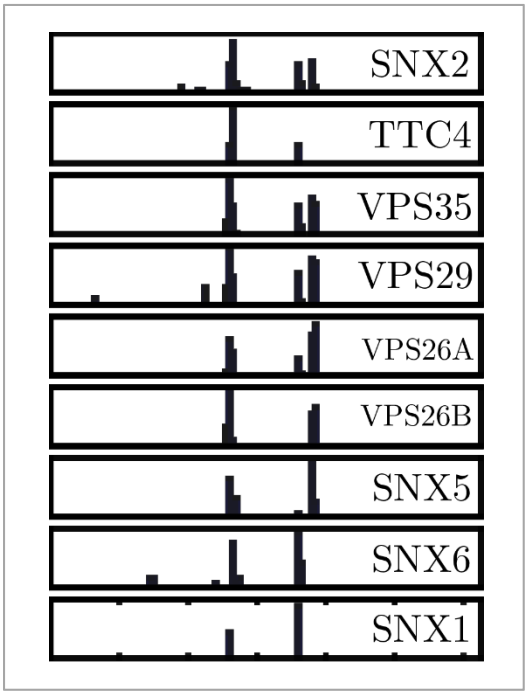

Figure C-1. TTC4 elution profile coincides with that of the retromer complex.

Representative elution profile of TTC4 clustered with members of the retromer complex (VPS26, VPS29, and VPS35 and associated SNX proteins). Peaks indicates protein elution within the 120 biochemical fractionations shown here for a HEK293 prep. The highly correlated elution profiles suggests TTC4 may be physically associated with the retromer.

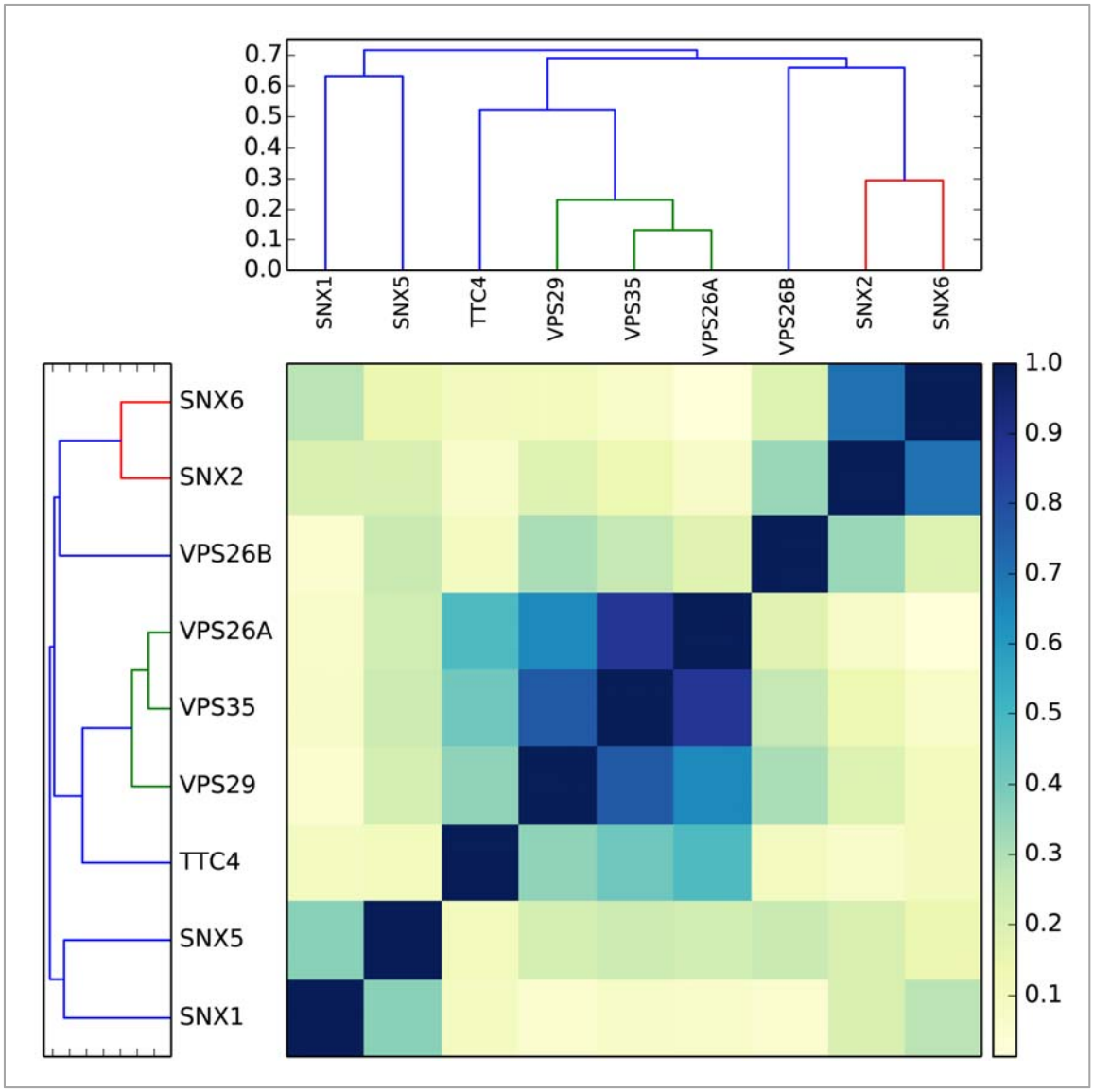

Figure C-2. Heat map elution correlation matrix of phylogenetically conserved TTC4-retromer interactions

Elution profiles of TTC4 and retromer components were integrated across multiple species tested, including worm (*C. elegans*) larvae, fly (*D. melanogaster*) S2 cells, mouse (*M. musculus*) embryonic stem cells, sea urchin (*S. purpuratus*) eggs and human (HEK293/HeLa) cell lines.

C2.2 Homology modeling of TTC4 structure

As previously mentioned, TTC4 contains TPR structural motifs which are known to be evolutionarily highly conserved. The TPR domains are composed of alpha-helix paired repeats; however, structural information for the rest of TTC4 is not available. We used HHpred (http://toolkit.tuebingen.mpg.de/hhpred) to produce query-template alignments that were then used for building a 3D structural homology model with the MODELLER software (Eswar et al. 2007). The TTC4 homology model was compared against structures in the Protein Data Bank using Dali server (Holm and Rosenstrom 2010) to find structural neighbors that may give clues to its molecular function. VPS29 was among one of the identified structural neighbors, strengthening our prediction that TTC4 is associated with the retromer in some capacity.

C2.3 However, directed experiments do not support TTC4-retromer complexing

The three retromer proteins VPS26-VPS29-VPS35 and the associated sorting nexins co-localize as foci in the cell cytoplasm (Rojas et al. 2007)(**Figure C-3**). Logically, I asked whether TTC4 would be co-localized with retromer foci. Endogenous TTC4 appears as foci in the Human Protein Atlas (HPA) (Uhlen et al. 2005), which initially seemed like a promising prospect. However, cytoplasmic TTC4 foci marked by the same HPA antibody (HPA041608) formed foci in HeLa and HEK293T cells that were distinct from those marked by retromer proteins (**Figure C-4**). In addition to cytoplasmic foci which were prominent in most cells (**Figure**

C-5), TTC4 was also sometimes marked in the nucleus and cytoskeleton (**Figure C-6**), consistent with observations gleaned from HPA images available online.

Interestingly, while foci marked by TTC4 was morphologically similar to those seen by the Human Protein Atlas group (**Figure C-5**), epitope- and fluorescent-tagged TTC4 did not form the same foci, instead showing diffuse signal in the cytoplasm (**Figure C-7**), which is consistent with previous literature on TTC4 characterization which heavily relies on recombinantly-tagged TTC4 (Dmitriev et al. 2009, Crevel, Bennett, and Cotterill 2008, Dmitriev et al. 2014, Huttlin et al. 2015). We found that GFP-labeled yeast and frog homologues of TTC4 were also diffuse their respective systems (data not shown).

Immunoprecipitation of endogenous VPS26 or recombinant VPS26 followed by mass spectrometry identified co-purified VPS29 and VPS35, but not SNX proteins nor TTC4 (data not shown). Immunoprecipitation of endogenous or recombinant TTC4 did not recover retromer proteins either, to be discussed in more detail in the next section.

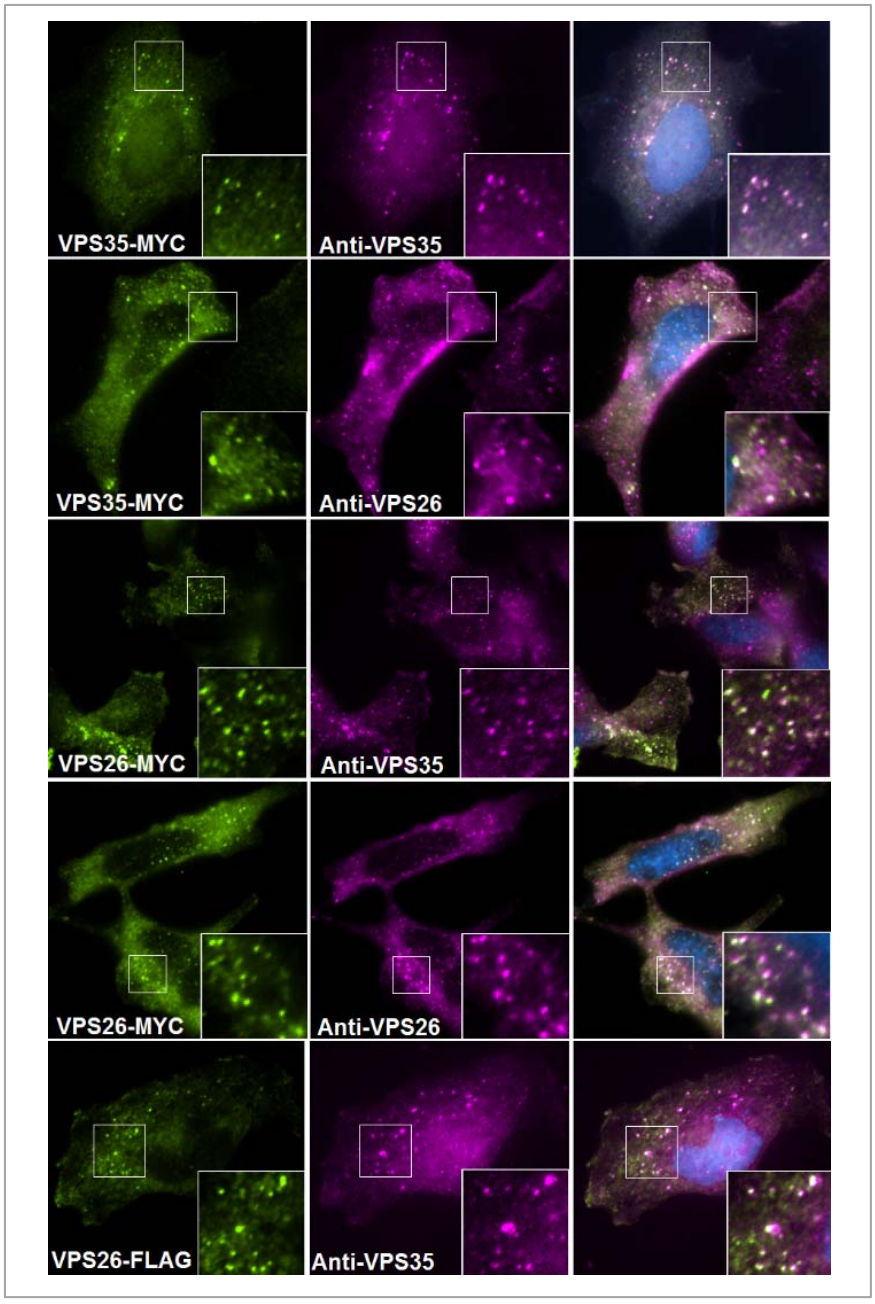

Figure C-3. Retromer components VPS26 and VPS35 co-localize into the same foci

HeLa cells transfected with VPS35-MYC, VPS26-MYC, or VPS26-FLAG formed foci that were marked by endogenous VPS35 or VPS26. The channels are merged in the third column with Hoechst-marked nuclei in blue.

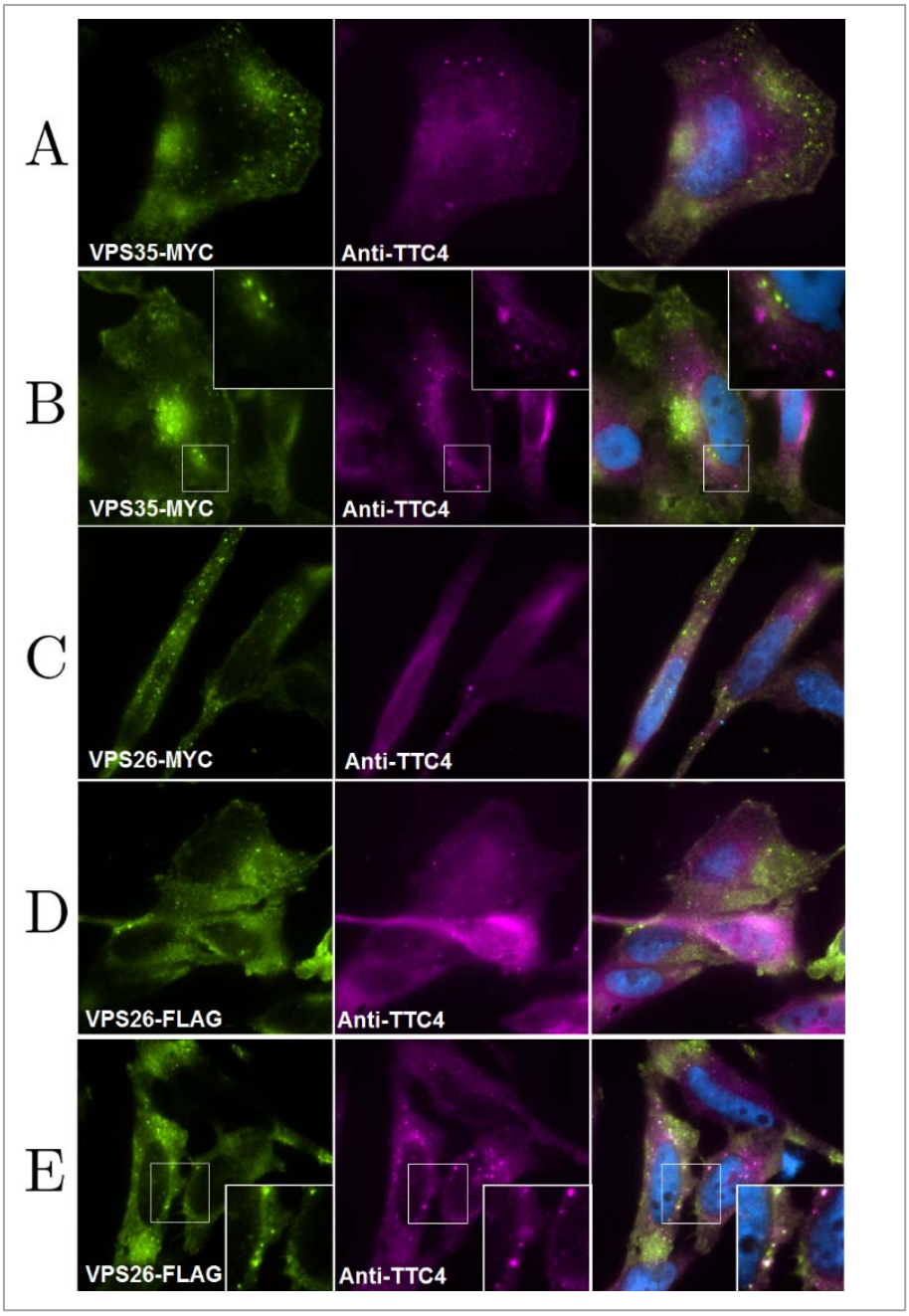

Figure C-4. TTC4 foci do not co-localize with retromer foci

Foci formed by retromer components VPS35 (**A**,**B**) and VPS26 (**C**,**D**) generally do not co-localize with TTC4 foci detected with the HPA041608 antibody. (**E**) Vary rare occurrence of co-localization is seen in <1% of cells; it is not clear whether those represent immunofluorescence aberrations. The channels are merged in the third column with Hoechst-marked nuclei in blue.

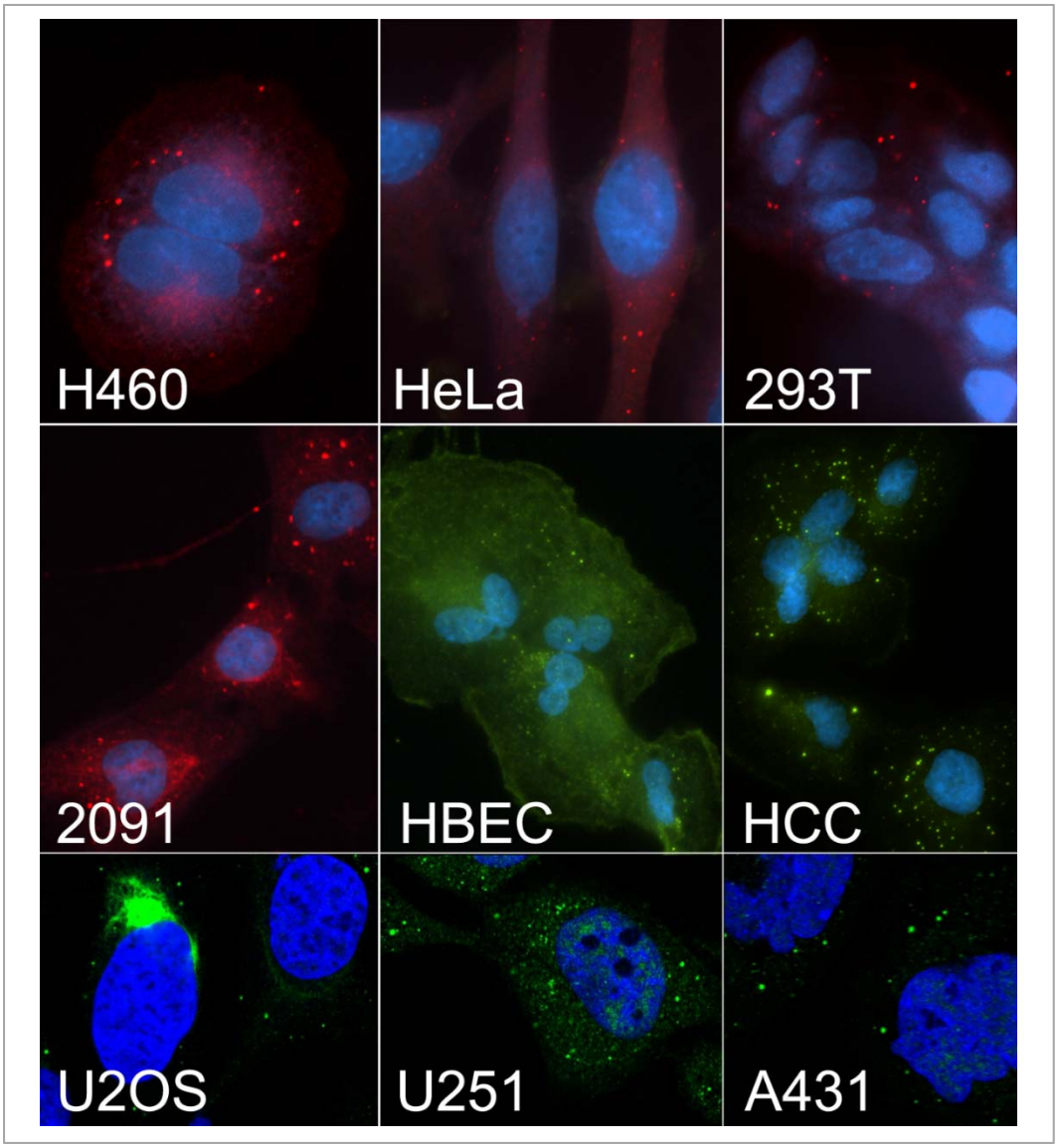

Figure C-5. Endogenous TTC4 shows distinct cytoplasmic puncta.

In all human cells tested, endogenous TTC4 probed with antibody HPA041608 forms cytoplasmic foci. Top row: H-460 large cell lung cancer cells, HeLa cervical cancer cells, HEK293T embryonic kidney cells; middle row: CRL-2091 normal fibroblast cells, HBEC30KT immortalized normal lung epithelial cells, HCC4017 non-small cell lung cancer cells. Similar results were acquired when anti-TTC4 antibody was labelled either with Alexa Fluor goat anti-rabbit 594 (red) or Alexa Fluor goat anti-rabbit 488 (green). This agrees with prior work done by the Swedish Human Protein Atlas project (Uhlen et al. 2005), shown in the bottom row: U-2 osteosarcoma cells, U-251 MG glioblastoma cells, and A-431 epidermoid carcinoma cells. Nuclei are visualized in blue with Hoechst (top and middle rows) or DAPI (bottom row).

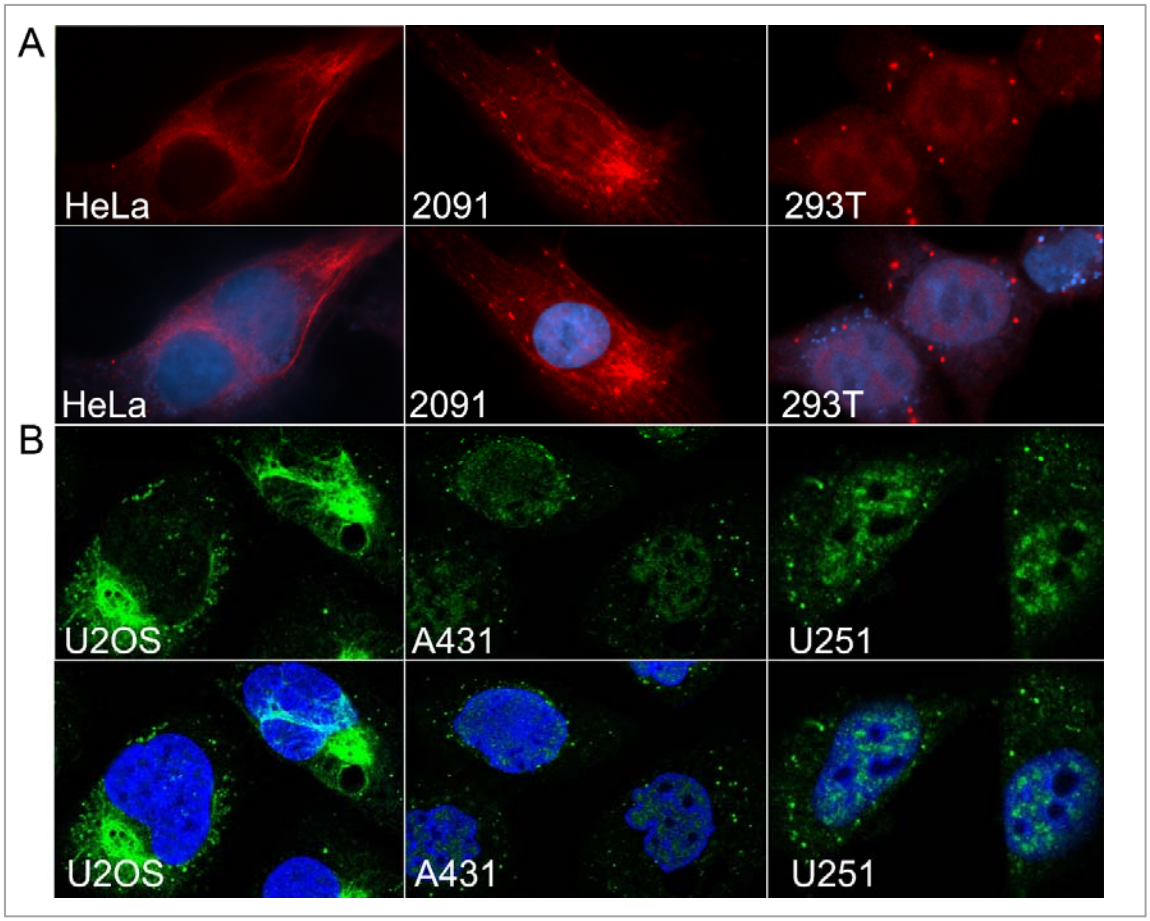

Figure C-6. Secondary localization of endogenous TTC4 in nuclei and cytoskeleton.

In addition to ubiquitous cytoplasmic puncta, TTC4 (probed with antibody HPA041608) sometimes show secondary subcellular localization to the cytoskeleton or nucleus. (**A**) Examples of cytoskeletal staining of TTC4 are shown in HeLa cells and CRL-2091 cells, and nuclear staining of TTC4 shown in HEK293T cells. This agrees with independent work in which (**B**) cytoskeletal staining was detected in U2-OS cells along with nuclear staining in A-431 and U-251MG cells (Uhlen et al. 2005). Nuclei are visualized in blue with Hoechst (**A**) or DAPI (**B**).

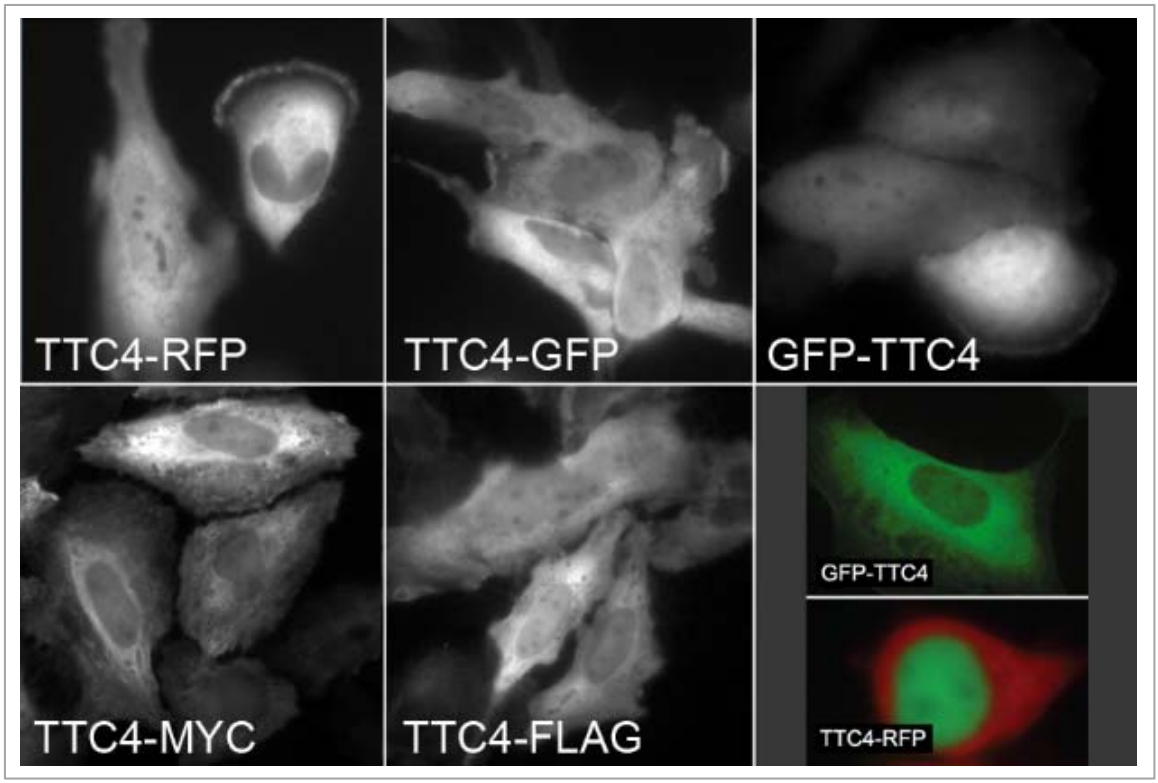

Figure C-7. Epitope and fluorescent-tagged TTC4 is diffuse.

TTC4 tagged with RFP, emGFP, MYC, or FLAG at either the N-terminus (attempted for GFP) or the C-terminus (attempted for all tags) of the protein show diffuse uniform localization in HeLa cells. This phenotype agrees with prior work in literature employing the similar methods of detection; shown in the grey bottom right panel: GFP-TTC4 from Dmitriev et al. (2009) and TTC4-RFP from Dmitriev et al. (2014).

C3. Unbiased search of TTC4 interaction partners

I then set out to identify whether the TTC4 foci marked by the HPA041608 antibody (here forth referred to as antibody 1, or ab1) were enriched for other protein components. Immunoprecipitation of endogenous TTC4 followed by mass spectrometry (IPMS) revealed high abundance of the mRNA decapping machinery

known to be present in mRNA processing bodies (p-bodies), specifically EDC4, DCP1A, EDC3, PATL1, and part of the LSM ring (**Figure C-8C**).

Due to the differential subcellular localization of TTC4 that was dependent on detection methods discussed earlier (**Figure C-5**, **Figure C-7**), IPMS using 2 additional TTC4 antibodies (ab2, HPA042459; and ab3, SAB1402385) and APMS targeting two differently tagged TTC4 constructs was performed (TTC4-FLAG and TTC4-GFP). Anti-TTC4 antibody 3 did not pull down any TTC4 and was then considered a mock pulldown and subsequently analyzed as a negative control. Interestingly, the 4 successful APMS and IPMS methods (endogenous TTC4 via ab1, endogenous TTC4 via ab2, recombinant TTC4 with FLAG, and recombinant TTC4 with GFP) produced 4 strikingly different potential interaction partners (**Figure C-8**). The top ~30 TTC4-interaction partners identified through each method was searched for functional protein association linkages among each set using StringDB, and for enrichment in gene ontology and biological pathways within each set using gProfiler (**Figure C-9**, **Figure C-10**). No overlap was found.

Notably, enriched proteins (prey) found with TTC4-FLAG overlaps considerably with that from Bioplex, a large-scale human interactome effort employing APMS with FLAG-HA-tagged proteins (Huttlin et al. 2015). Of their seven significantly enriched preys captured by TTC4-FLAG-HA, four overlap with my set (EEF2, FAM203B, HSP90AA1, HSP90AB1**)**. A slightly modified APMS protocol (using anti-FLAG bound to Dynabeads, as opposed to anti-FLAG agarose) pulled down a fifth shared prey (GCC2; data not shown). The two TTC4 prey from

Bioplex not seen in my TTC4-FLAG APMS experiments are HSP90AB3P and C10orf137.

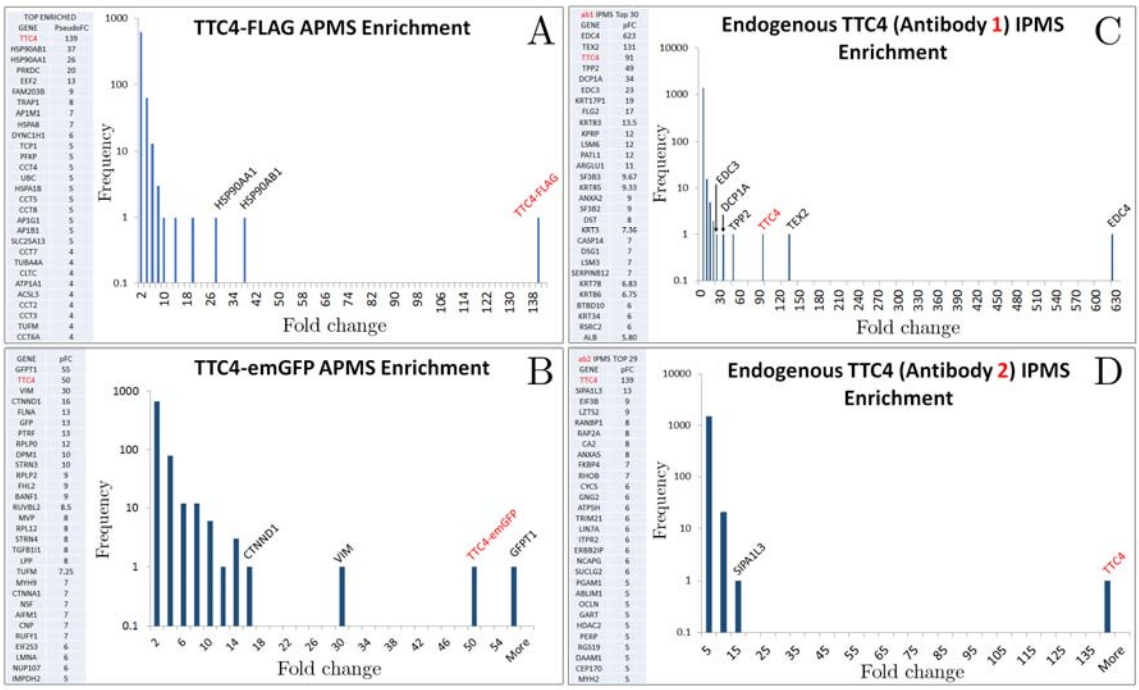

Figure C-8. Various TTC4 bait pulldown techniques with differentially recovered prey.

(**A,B**) TTC4 was constructed into epitope-tagged plasmids and transfected into HEK293T and/or HeLa cells, from which affinity precipitation followed by mass spectrometry (APMS) protein identification was performed on the eluate. The two tags show differential prey proteins. (**A**) Proteins recovered with TTC4-FLAG show high correlation with those prey recovered by the Bioplex group (Huttlin et al. 2015). (**C,D**) Immunoprecipitation followed by mass spectrometry (IPMS) of endogenous TTC4 using two different polyclonal antibodies produced in rabbit from the Human Protein Atlas (antibody 1 is HPA041608, polyclonal produced in rabbit; antibody 2 is HPA042459, polyclonal produced in rabbit). A third TTC4 antibody (antibody 3, SAB1402385, monoclonal produced in mouse) was used for IPMS but did not pull down any TTC4. It is included in the above dataset within the control sets. Fold change (pFC or pseudo-fold change) on the x-axes are calculated as the ratio of pseudoPSMs (peptide spectral matches for each protein + 1) for each protein in the pull down set vs. mock pull-down set. Frequency on the y-axes are the number of individual unique proteins, plotted in log-scale. A list of the most-enriched proteins per IP is shown. For each AP/IP-MS, bait is shown in red.

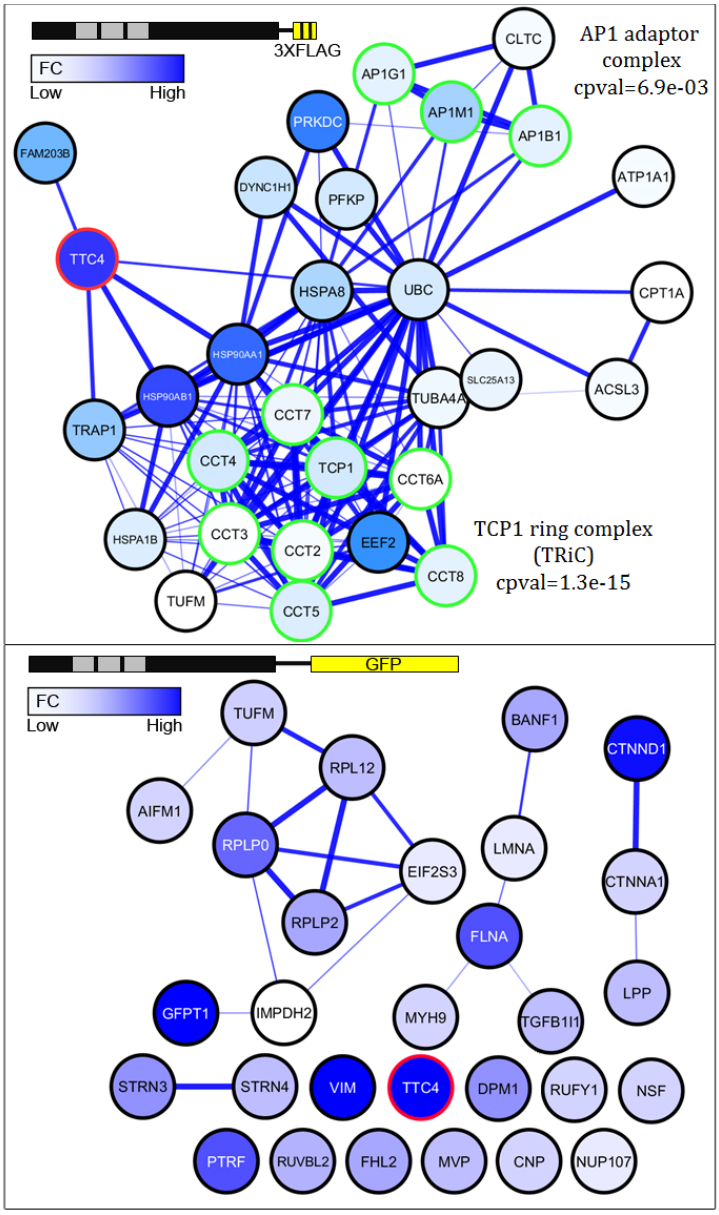

Figure C-9. StringDB interactions of top tagged-TTC4 APMS prey

Bait (yellow) is tagged-TTC4 (represented as a black bar with grey TPR domains). **Top**: StringDB interactions between TTC4-FLAG (bait, red circle outline) APMS top enriched prey proteins are represented with gene names as nodes connected by weighted edges. Darker nodes are genes with higher enrichment (fold change; FC) in the pulldown. Green outlined circles present significant protein complexes found by gProfiler. cpval is cumulative p-value. **Bottom**: StringDB interactions between TTC4-emGFP (bait, red circle outline) APMS top enriched prey proteins, represented similarly. No significant complexes were found. TUFM is shared among the pulldowns.

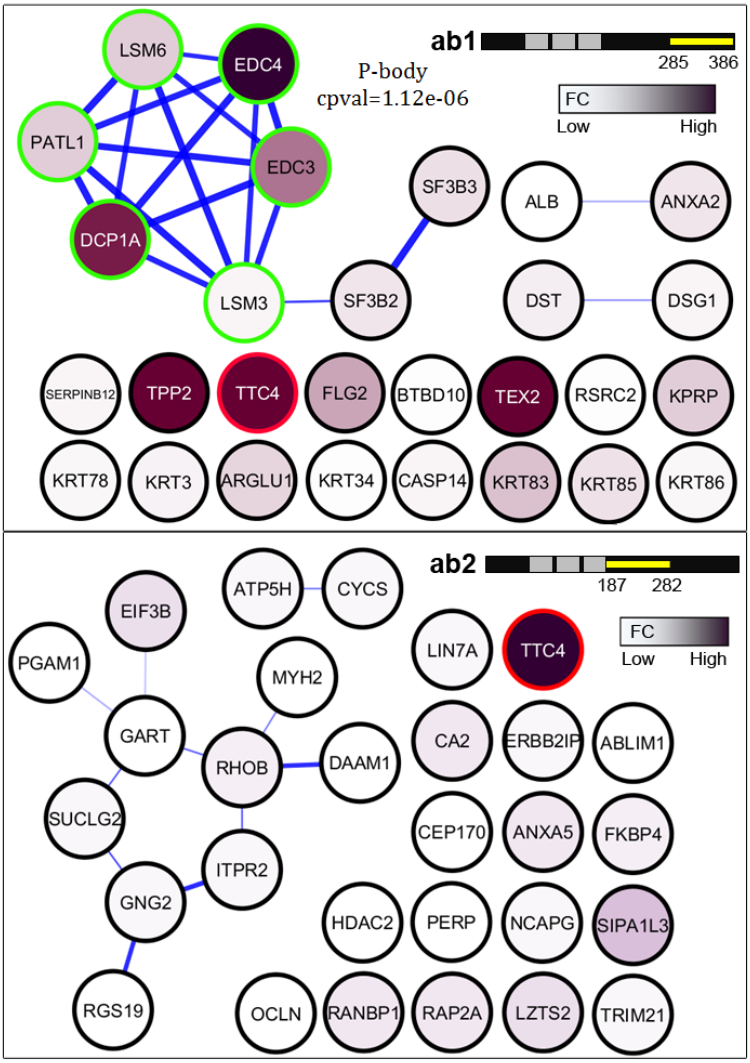

Figure C-10. StringDB interactions of top endogenous TTC4 IPMS prey

Commercial antibodies were made with immunogen sequences corresponding to different regions (yellow) of TTC4 (represented by a black bar with 3 grey TPR domains). **Top**: StringDB interactions between top enriched prey proteins pulled down with "antibody 1" represented with gene names as nodes connected by weighted edges. Darker nodes are genes with higher enrichment scores (fold change; FC) in the pulldown. Green outlined circles present significant protein complexes found by gProfiler, of which one is found: the mRNA decapping complex within p-bodies (cumulative p-value, or cpval = 1.12e-06). **Bottom**: StringDB interactions between top enriched prey proteins pulled down with "antibody 2" represented similarly. No significant complexes were found. No enriched prey proteins are shared between the two pulldowns or with the APMS pulldowns.

C4. Possible TTC4 association with p-bodies

Processing bodies (P-bodies) are cellular structures that regulates the stability of cytoplasmic mRNAs. They can be visualized as cytoplasmic puncta and contain numerous RNA silencing and degradation factors, including deadenylases and decapping complexes. P-bodies are also known as GW bodies, named for the marker protein GW182 which contains glycine (G) and tryptophan (W) rich repeats. Canonical p-body markers also include decapping complex members EDC4 (also known as Ge-1 or HEDLS), EDC3, DCP1A/1B/DCP2, RCK/P54, XRN1, PATL1 and the Lsm1-7 complex (Zheng, Chen, and Shyu 2011). P-body assembly is dependent on HSP90 activity, although HSP90 itself may not be present in p-bodies (Matsumoto et al. 2011). P-bodies are frequently physically juxtaposed to stress granules (Anderson and Kedersha), which possess stalled translational preinitiation complexes that accumulate during cellular stress.

C4.1 TTC4 foci are marked with p-body proteins

Due to the identification of foci marked by anti-TTC4 antibody 1 (**Figure C-5**) and to TTC4's physical interaction with decapping proteins identified with antibody 1 (**Figure C-10**), I asked whether foci marked by TTC4 were also marked by p-body components. Indeed, many TTC4 foci co-localized with three attempted p-body markers: DCP1A, EDC3, and LSM6 (**Figure C-11**), chosen due to their availability in our ORFeome collection.

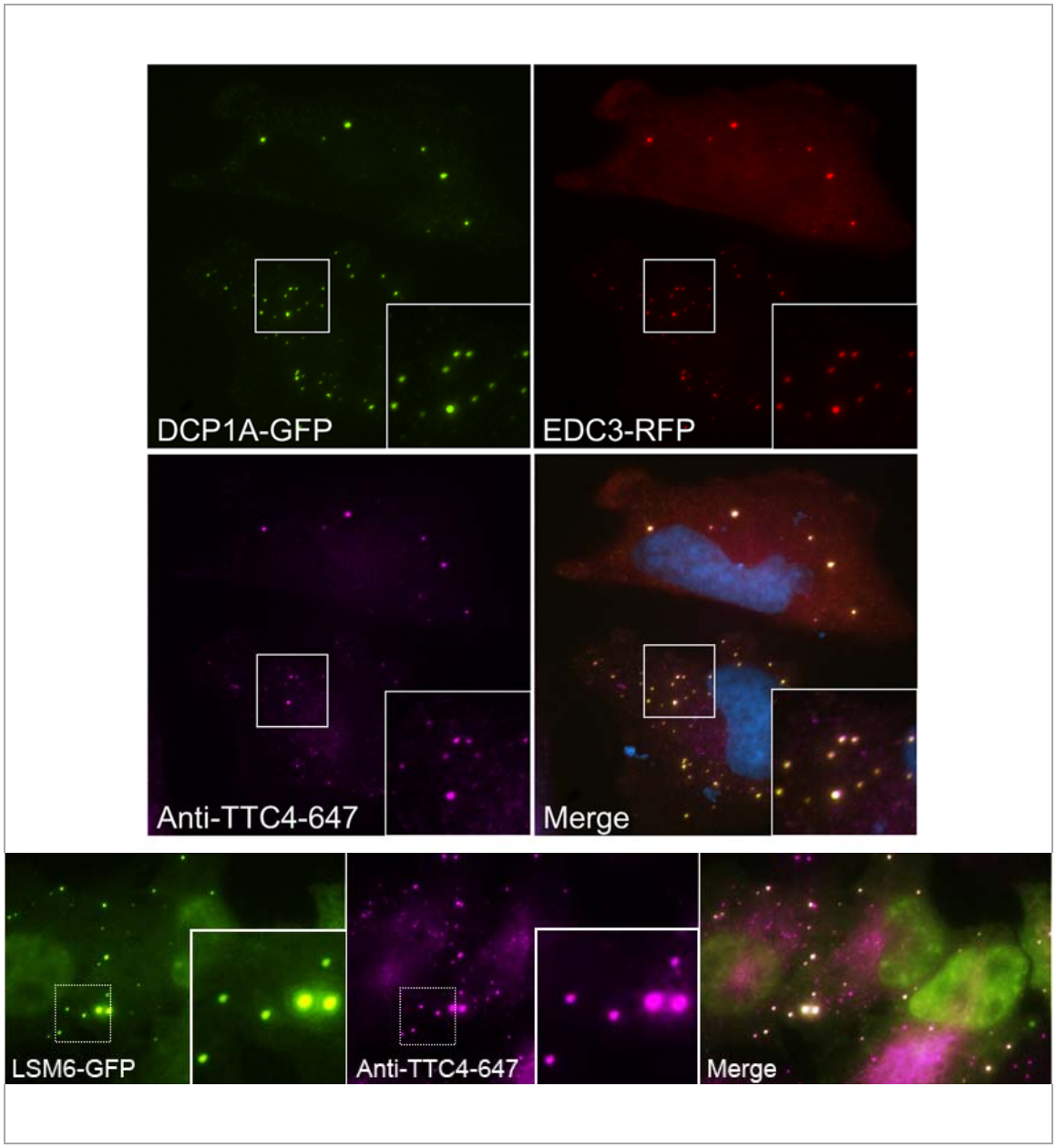

Figure C-11. TTC4 foci co-localize with p-bodies.

Endogenous TTC4 marked by antibody 1 co-localizes with transiently co-transfected p-body markers (**top**) DCP1A and EDC3, and transfected LSM6 (**bottom**) in HeLa cells. Nuclei stained with Hoechst is shown in the top merged image. LSM6 additionally marks nuclei in the bottom panel. Similar co-localization between TTC4 and p-body markers is seen in HEK293T cells (data not shown).

C4.2 Like p-bodies, TTC4 foci are distinct from stress granules

Stress granules are cytoplasmic granules consisting of stalled translational preinitiation complexes that accumulate during stress. They are visualized as cytoplasmic foci distinct from but frequently juxtaposed to p-bodies (Anderson and Kedersha 2009). Some components of p-bodies are shared by stress granules (Decker and Parker 2012), and exchange of material between the two mRNP structures are known. It has been proposed that mRNA released from disassembled polysomes is sorted and remodeled at stress granules, from which selected transcripts are delivered to p-bodies for degradation or to polysomes for re-entry into translation (Kedersha et al. 2005, Buchan, Muhlrad, and Parker 2008). Consistent with literature, stress granules were generally distinct from foci formed by TTC4, similar to behavior expected between stress granules and p-bodies (**Figure C-12**).

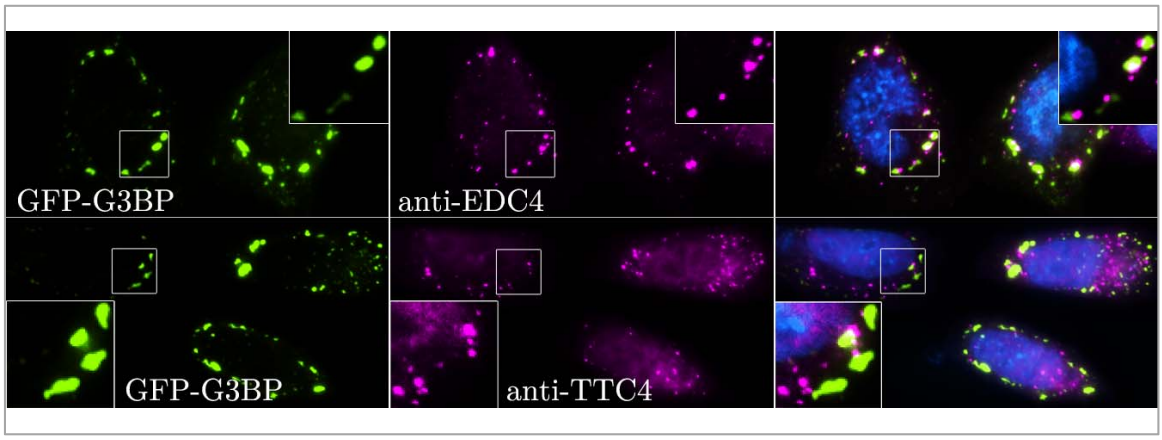

Figure C-12. TTC4 and EDC4 are distinct from stress granules

emGFP-G3BP-marked stress granules are induced in HeLa cells by 500 µM sodium arsenite for 1 hour and generally do not co-localize with p-bodies marked by endogenous EDC4 or TTC4. Many stress granules are juxtaposed to EDC4 or TTC4 foci, consistent with previous literature on the relationship between stress granules and p-bodies.

C4.3 Foci disruption by chemical perturbations

P-body formation requires Hsp90 activity, and inhibition of that activity by geldanamycin is known to disrupt p-body formation (Suzuki et al. 2009). As well, p-body formation requires its association with RNA, and p-bodies are lost upon cycloheximide-induced blockage of translation elongation, which traps mRNAs on ribosomes (Teixeira et al. 2005). These two p-body disrupting agents also inhibited the formation of TTC4 foci marked by antibody 1 (**Figure C-13** and **Figure C-14**).

It is known that geldanamycin can reduce endogenous and transfected expression levels of some p-body components (Johnston et al. 2010). I determined that the effects of geldanamycin and cycloheximide on TTC4 foci disruption was minimally due to reduction in its protein expression levels, though geldanamycin had a much more pronounced effect on suppressing transfected TTC4-emGFP expression (**Figure C-15**). Measurements of EDC4 expression levels were not successful (could not see band on western, data not shown).

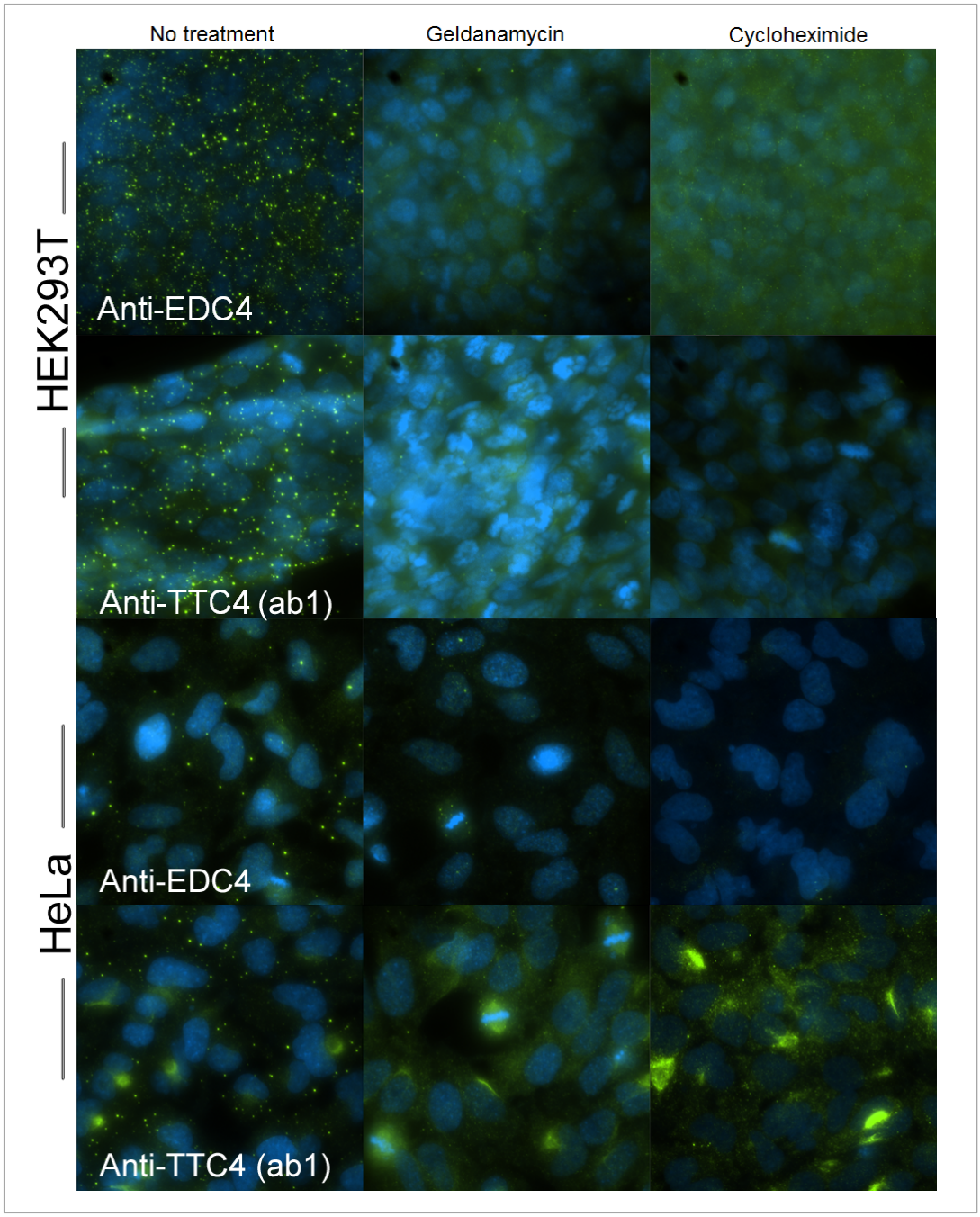

Figure C-13. P-body disruption agents also inhibit TTC4 foci formation

Cells treated with 10 µM geldanamycin for 20 hours or 20ug/mL cycloheximide for 1 hour exhibit markedly reduced EDC4 and TTC4 foci as compared to an untreated control. Antibody signal is shown in green. Nuclei are stained with Hoechst in blue. Images within each cell line were taken with the same exposure and LUTs are scaled equally for both green and blue channels.

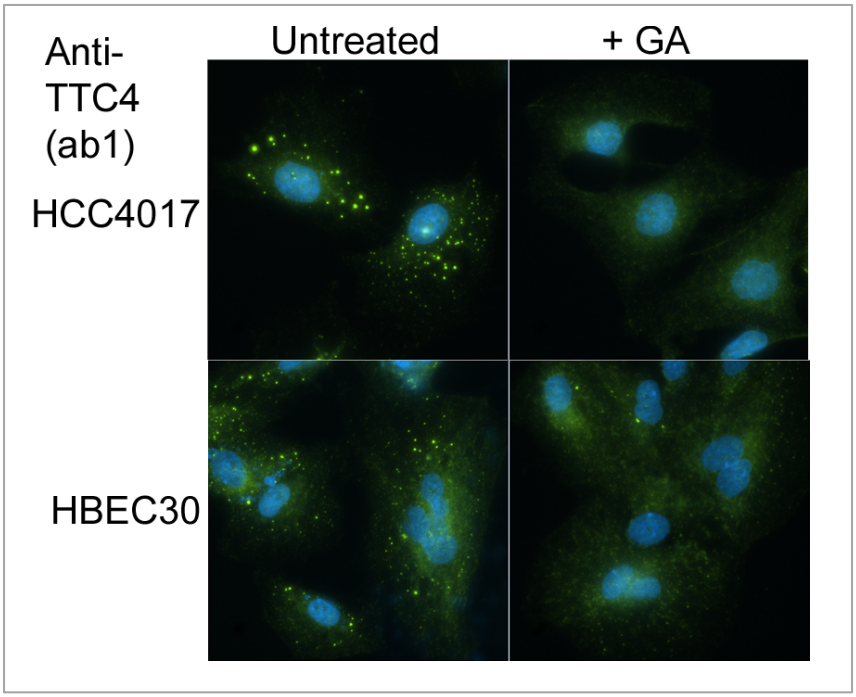

Figure C-14. Geldanamycin inhibits TTC4 foci in patient-derived lung cells

Patient-matched lung cells HCC4017 (cancer) or HBEC30KT (normal) treated with 10 µM geldanamycin for ~20 hours exhibit markedly reduced EDC4 and TTC4 foci as compared to an untreated control. Antibody signal is shown in green. Nuclei are stained with Hoechst in blue. Images within each cell line were taken with the same exposure and LUTs are scaled equally.

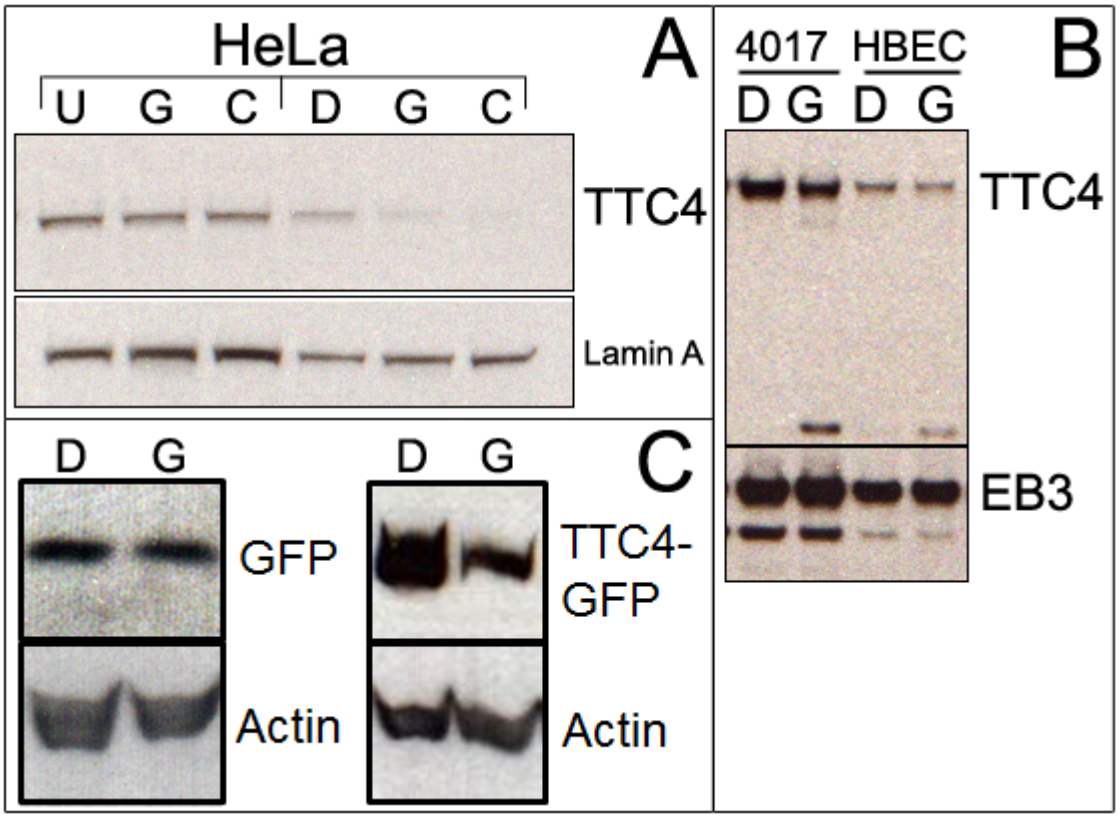

Figure C-15. Geldanamycin and cycloheximide minimally affect TTC4 expression

U=untreated, D=DMSO (carrier), G=geldanamycin (GA; 10 µM, 20-22hrs), C=cycloheximide (20ug/mL, 1hr). (**A**) Two replicates of HeLa cells or (**B**) lung cancer HCC4017 and patient-matched normal HBEC30KT cells treated with or without p-body disrupting agents and probed for endogenous levels of TTC4, shows slight but minimal expression changes in TTC4. In lung cancer cells, a small TTC4 product of ~20kDa is detected in the presence of GA. (**C**) Overexpression of TTC4-emGFP is inhibited with GA compared to the DMSO control, but GFP itself is not affected. TTC4-GFP was probed with anti-TTC4 (shown) and anti-GFP (not shown) with identical results. GFP control was probed with anti-GFP.

C5. Subcellular localization of TTC4 is influenced by method of visual examination

As shown earlier, I employed two commonly used approaches to examine the subcellular localization of TTC4: transient transfection of recombinant TTC4 (tagged with fluorescent proteins or small epitopes) and immunofluorescence of endogenous TTC4 using one antibody from HPA. The two different methods gave very different results. While tagged TTC4 was always uniformly diffuse within the cytoplasm with occasional diffuse nuclear signal (**Figure C-7**), immunofluorescence of endogenous TTC4 (by antibody 1) showed distinct puncta within the cytoplasm (**Figure C-5**) with occasional cytoskeletal and nuclear signal (**Figure C-6**).

Tagging proteins is known to sometimes cause their mis-localization (Ramanathan, Ayyavoo, and Weiner 2001, Skube, Chaverri, and Goodson 2010, Landgraf et al. 2012, Palmer and Freeman 2004, Snapp 2005). Due to the possibility of tag-induced mischaracterization of TTC4, I decided to further consider the antibody-based data gathered for endogenous TTC4. The second HPA antibody targeting TTC4 (antibody 2; HPA042459) identified TTC4's uniformly diffuse distribution within the cytoplasm, although cells also frequently contained one or two foci (**Figure C-16**).

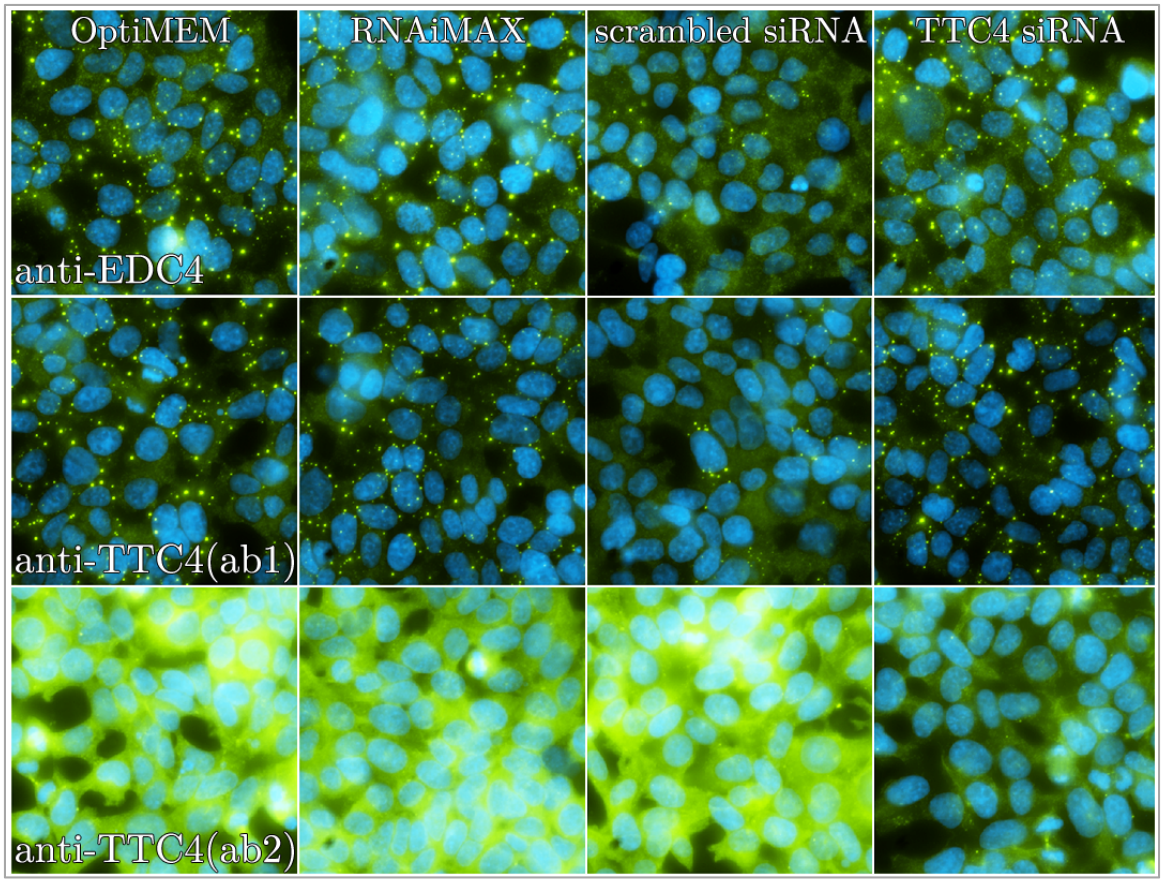

Figure C-16. Anti-TTC4 antibodies ab1 and ab2 give unique staining patterns

Mock knockdowns (OptiMEM: transfection media only, RNAiMAX: transfection media and carrier reagent only, scrambled siRNA: transfection with negative control siRNA) and TTC4 knockdown (transfection with anti-TTC4 siRNA) is immunostained and with anti-EDC4 antibody (top row), and with anti-TTC4 antibodies 1 and 2 (middle and bottom row, respectively), shown in green. Nuclei are stained with Hoechst in blue. Interestingly, foci containing EDC4 and foci containing TTC4 marked by ab1 were inhibited with scrambled siRNA mock knockdown but not with TTC4 knockdown. Diffuse TTC4 marked by antibody 2 was notably decreased in the TTC knockdown compared to all mocks. All images within each row were captured under identical conditions and signal intensities scaled equally for comparison.

C6. Integrity of materials

Due to a lack of consensus between immunofluorescence and IPMS results obtained with different anti-TTC4 antibodies (namely ab1 and ab2, corresponding to reagents HPA041608 and HPA042459, respectively), an important concern is whether the antibody actually targets TTC4 as advertised, as well as the possibility of differences arising due to potential secondary targeting, or off-targets. These two antibodies were designed against two non-overlapping regions of TTC4 outside of its TPR domains (**Figure C-17A**). The immunogen sequence for ab1 does not significantly align to any other protein human (by blastq). The immunogen sequence for ab2, however, does significantly align with TCPR2 and SRGAP2, although neither were captured in the IP using ab2. For completeness, the immunogen sequence for ab3, due to its partial inclusion of the TPR domain, also significantly aligns to a few other proteins. The immunogen sequences are provided in the Methods. In addition to immunoprecipitation discussed earlier, I employed experimental approaches to further address specificity of the antibodies: by western blot against TTC4 overexpression and depletion, and immunofluorescence against TTC4 depletion.

By western blotting, antibody 1 shows a major band at the expected TTC4 size (47 kDa) that is lost when TTC4 is knocked down by siRNA, and also detected overexpressed recombinant TTC4 (**Figure C-17B**). However, antibody 1 did not show loss of foci by immunofluorescence during TTC4 knockdown (**Figure C-16**). Similarly, TTC4 knockdown did not affect foci formed by EDC4 (**Figure C-16**), suggesting that TTC4 is not critical for p-body formation. It is not immediately clear

why EDC4 and TTC4 foci were reduced in the mock knockdown with scrambled siRNA. (Although one likely possibility is that, because p-bodies are also sites of RNA interference, target-less siRNAs may cause disruption of the RNAi machinery leading to p-body disassembly.)

By western blotting, antibody 2 shows several major bands in addition to that corresponding to the expected migration pattern for TTC4, although the major bands do not seem to be reduced in the TTC4 knockdown (**Figure C-17**). However, antibody 2 detects recombinant TTC4. Immunostaining with antibody 2 shows cytoplasmic diffuse staining pattern that is diminished in signal intensity during TTC4 knockdown.

Antibodies 1 and 2 were both demonstrated to successfully target TTC4 by at least two orthogonal methods. Due to antibody 3's inability to immunoprecipitate TTC4 or detect the expected sized band on western blot (**Figure C-17B**), it was not further considered. However, due to the general non-consensus between antibodies 1 and 2 in TTC4's characterization, I decided to further analyze endogenous TTC4 using label-free methods.

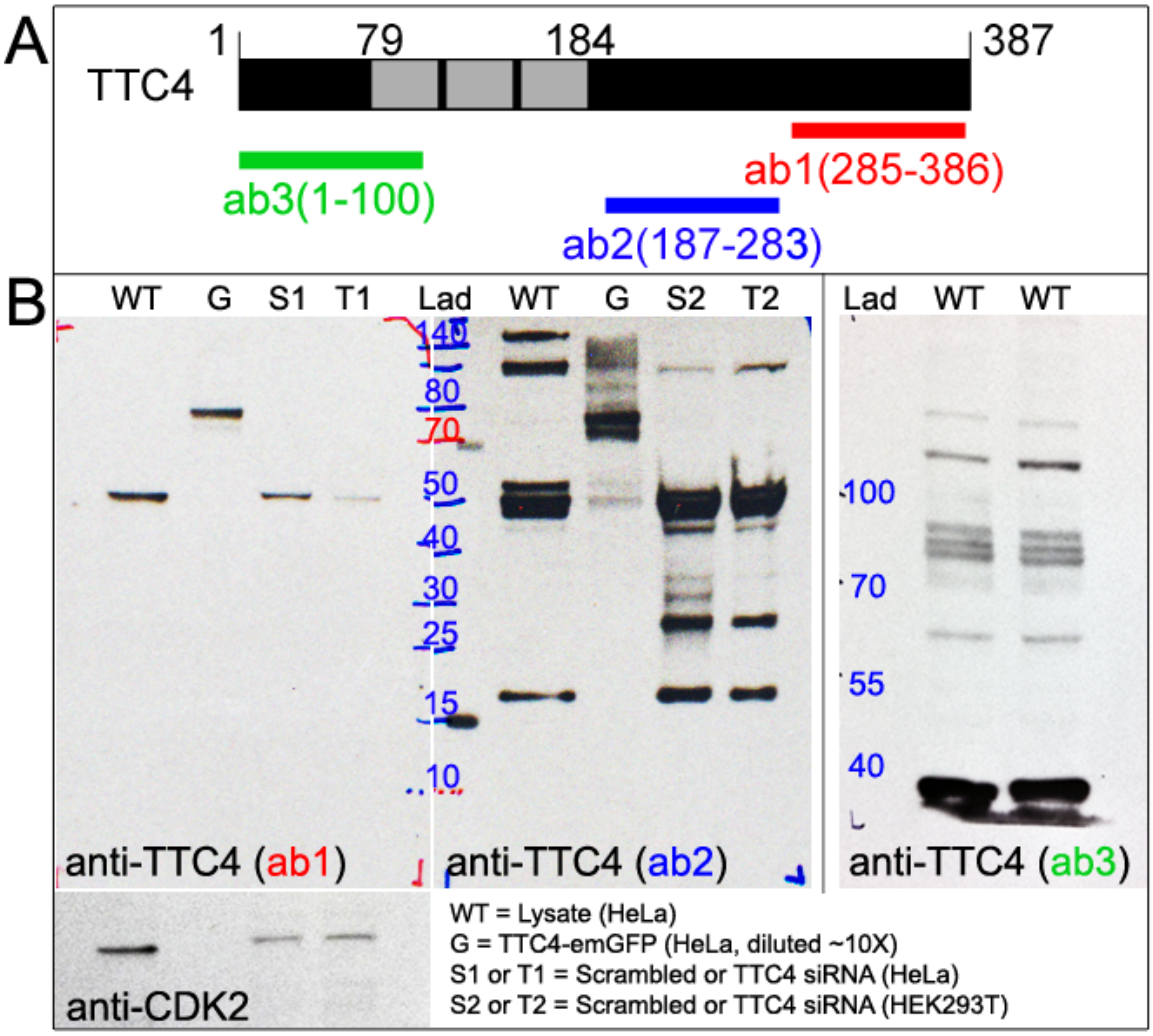

Figure C-17. Western blot detection of TTC4 using various commercial antibodies.

(**A**) Schematic of TTC4 (387 a.a.s), with its three TPR domains in grey (residues 79-112, 117-150, 151-184). Immunogen sequence regions for the production of each antibody are indicated: antibody 1 (ab1, HPA041608, Sigma), antibody 2 (ab2, HPA042459, Sigma), and antibody 3 (ab3, SAB1402385, Sigma). (**B**) Each anti-TTC4 have different western blot band patterns. TTC4 is ~45 kDa. Ab1 and ab2 both detect recombinant TTC4-emGFP (lysate was diluted due to its high expression), but ab1 gives the cleanest signal and the major band is reduced upon TTC4 knockdown by siRNA (anti-CDK2 is shown as loading control). Ab3, which fails to immunoprecipitate TTC4, does not seem to be a successful antibody for western either. The "ab1" and "ab2" gel was run on a Bolt 4-12% Bis-Tris precast gel (Thermo Scientific). "Ab3" gel was run on a 7.5% Mini-PROTEAN TGX precast gel (BioRad).

C7. RNAseq of TTC4 revisits its association with p-bodies

Using transcriptome sequencing, I looked for global mRNA changes upon TTC4 depletion by siRNA (depletion quantified; **Figure C-20**) to identify candidate functionally associated genes in a tag-free, label-free manner. Along with the expected TTC4 transcript depletion, the transcript abundances of many other genes were also reduced, including quite notably those of EIF4G2 and KIAA0430 (**Figure C-18** and **Figure C-19**). The top three strongly affected genes (excluding TTC4), EIF4G2, KIAA0430, and SREK1IP1, are all known to be associated with or directly involved in mRNA processing and RNA splicing (Heese et al. 2004, Bloch et al. 2014, Costello et al. 2015). EIF4G2 and KIAA0430 have previously been found in p-bodies.

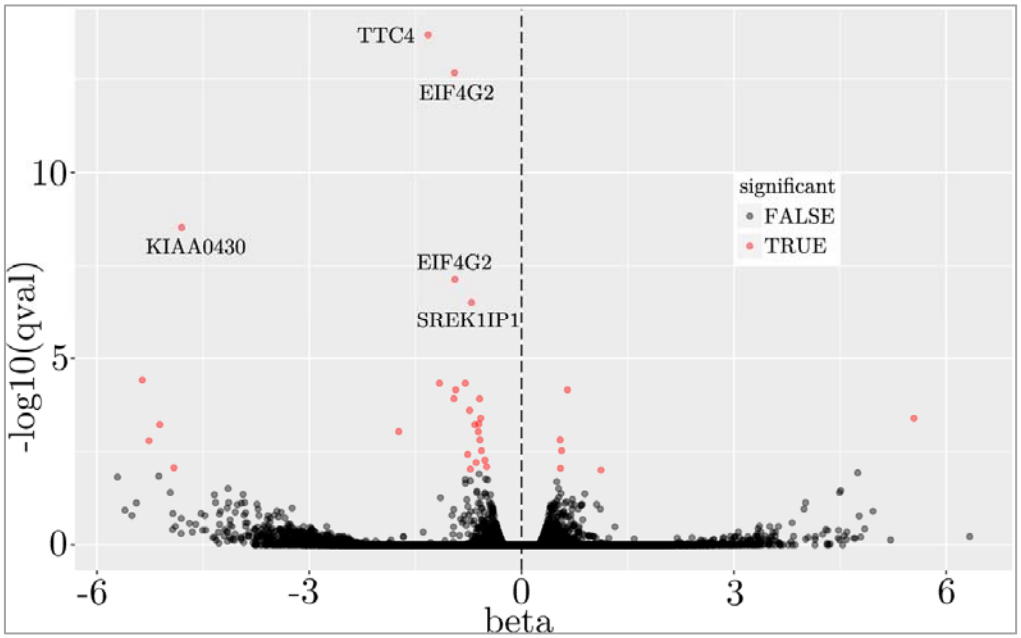

Figure C-18. Differential transcript abundances during TTC4 depletion

Volcano plot of all differential transcript abundances during siRNA-TTC4 knockdown in HEK293T cells. x-axis: beta is a bias estimator (analogous to a log-based measure of fold change). y-axis: qval is the false discovery rate (FDR) adjusted pvalue using Benjamini-Hochberg. Significant data-points are in shown red with a FDR of 0.01, with a few notable gene names labeled.

	Transcript ID	Gene	qval	b	se_b	mean_obs	var_obs
1	**ENST00000371284**	**TTC4**	**2.08E-14**	**-1.3206**	**0.147**	**7.102627**	**0.53518**
2	ENST00000532120	EIF4G2	2.14E-13	-0.9475	0.109	6.590462	0.26548
3	ENST00000621511	KIAA0430	2.99E-09	-4.8031	0.645	3.467837	7.07113
4	ENST00000525681	EIF4G2	7.36E-08	-0.9427	0.135	10.13287	0.26213
5	ENST00000513458	SREK1IP1	3.06E-07	-0.7072	0.105	7.278053	0.15665
6	ENST00000484181	PCCB	3.85E-05	-5.3599	0.899	1.986829	8.27518
7	ENST00000261798	CSNK1A1	4.66E-05	-0.7952	0.135	7.940307	0.212
8	ENST00000361815	PSMD10	4.66E-05	-1.1602	0.197	5.700122	0.45129
9	ENST00000369042	BEND3	7.03E-05	0.6478	0.112	6.463144	0.13317
10	ENST00000396525	EIF4G2	7.03E-05	-0.9302	0.161	6.941308	0.26791
11	ENST00000412431	CSNK1A1	0.0001213	-0.956	0.169	6.904841	0.3099
12	ENST00000282344	USP12	0.0001222	-0.5932	0.105	7.429968	0.11411
13	**ENST00000486621**	**TTC4**	**0.0002504**	**-0.7342**	**0.133**	**5.331031**	**0.18197**
14	ENST00000356978	CALM1	0.0004031	-0.5775	0.107	7.6507	0.11103
15	ENST00000394989	SNCA	0.0004031	5.5434	1.026	2.318746	8.88553
16	ENST00000374704	CCNY	0.0005751	-0.6057	0.114	6.855507	0.127
17	ENST00000396222	LDHA	0.0006012	-5.1125	0.964	1.863099	7.66219
18	ENST00000536441	SESN3	0.0006012	-0.6606	0.125	8.067674	0.15133
19	**ENST00000371281**	**TTC4**	**0.0009182**	**-1.7358**	**0.333**	**6.278466**	**1.05125**
20	ENST00000295958	SMIM14	0.0009222	-0.6118	0.118	5.956021	0.1307
21	ENST00000256925	CABLES1	0.0015411	0.5453	0.107	6.756533	0.10266
22	ENST00000438909	G2E3	0.0015411	-0.5889	0.116	6.99348	0.12209
23	ENST00000327570	RWDD4	0.0016274	-5.2646	1.039	1.939158	8.10203
24	ENST00000305123	IRS1	0.0030015	0.5654	0.115	5.258038	0.10832
25	ENST00000489294	UHMK1	0.0030015	-0.567	0.115	8.580466	0.10519
26	ENST00000366687	CCSAP	0.0037829	-0.7612	0.156	7.942574	0.20729
27	ENST00000367097	TULP4	0.0053786	-0.5182	0.108	6.848835	0.08663
28	ENST00000335327	WASF3	0.0061833	-0.6421	0.135	6.95309	0.14893
29	ENST00000298130	SPTSSA	0.008031	-0.4932	0.105	6.812264	0.07305
30	ENST00000578386	SMURF2	0.0085882	-4.9125	1.049	1.763113	6.94794
31	ENST00000533683	SAMD1	0.0088715	0.5497	0.118	6.223507	0.11006
32	ENST00000328268	CRELD2	0.0093147	-0.7237	0.155	6.359092	0.19086
33	ENST00000351017	MLLT4	0.0099046	1.1219	0.242	4.639511	0.46001

Figure C-19. Transcriptome profiling upon TTC4 knockdown

HEK293T cells were subject to RNAi against TTC4 followed by transcriptome analysis. The genes for which the multiple hypothesis-corrected p-value (qval) was <0.01 are ranked by qval; all 33 genes under this condition are shown. All 4 TTC4 transcripts (3 are shown here, bolded) are reduced in the knockdown group. Multiple EIF4G2 transcripts are also reduced (3 are shown here, blue). 4 replicates comprised the knockdown group, and 4 replicates consisting of 2 "untransfected" and 2 "mock-transfected" replicates comprised the control group. RNAseq reads were mapped with Kallisto (Bray et al. 2016) and analyzed with sleuth (http://pachterlab.github.io/sleuth/about.html). Qval = FDR adjusted p-value; b = beta (a bias estimator) analogous to fold change; se_b = standard error of b; mean_obs = mean of observations used for smoothing; var_obs = variance of observations.

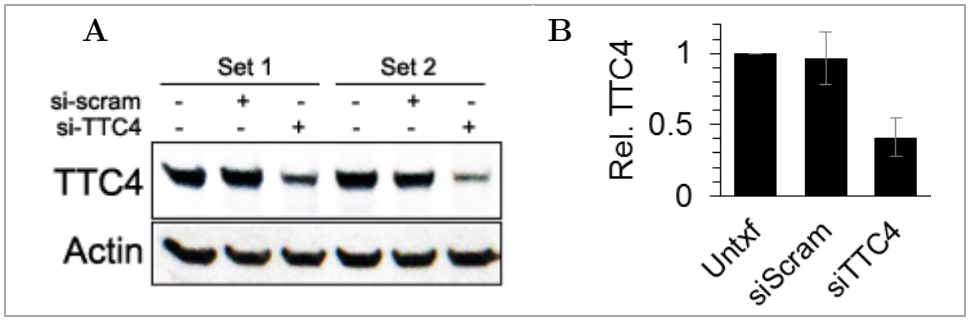

Figure C-20. TTC4 knockdown efficiency for RNAseq

Two sets of reverse-transfected knockdown experiments in HEK293T cells are performed for RNAseq. (**A**) Western blot of untransfected (OptiMEM only) cells, mock knockdown (scrambled siRNA), and TTC4 knockdown (TTC4 siRNA) cells. TTC4 is detected by anti-TTC4 antibody 1, and equal loading is detected with actin. (**B**) Plot of TTC4 levels normalized to actin for 'mock knockdown' and 'TTC4 knockdown' samples relative to TTC4 levels normalized to actin of 'untransfected' samples. Bars represent average +/- 1 s.d. across 2 replicates. By definition, the relative TTC4 levels for 'untransfected' is 1. Quantified western blot signal was acquired with ImageJ.

C7.1 Functional association of TTC4 with EIF4G2

Eukaryotic translation initiation factor 4-Gamma 2 (eIF4G2) is also known as DAP-5 (death-associated protein 5), p97, and NAT1. Unlike EIF4G1 (a scaffolding component of the EIF4F mRNA cap binding complex and involved in eukaryotic translation initiation) which supports cap-dependent and independent translation, EIF4G2 functions as a general repressor of both cap-dependent and independent translation by forming translationally inactive complexes (Imataka, Olsen, and Sonenberg 1997). EIF4G2, which exhibits 27% identity to the C-terminal two-thirds of EIF4G1, binds to EIF4A and EIF3, but not to EIF4E (Imataka, Olsen, and Sonenberg 1997).

As previously mentioned, stress granules and p-bodies are spatially, compositionally, and functionally linked. Although translation initiation factors (with the exception of EIF4E) known to be associated with stress granules were historically considered to be absent from p-bodies (Kedersha et al. 2005, Kedersha and Anderson 2007, Kedersha et al. 2008), studies have shown that translation initiation factors including EIF4G2 localize to p-bodies under certain environmental or growth conditions (e.g., during glucose deprivation and stationary phase in yeast) (Brengues and Parker 2007). It is hypothesized that the EIF4G2-containing mRNPs seen associated with p-bodies conditionally is actually formed during normal cycling of mRNAs between p-bodies and polysomes, but is too transient to be detected under normal cell conditions (Brengues and Parker 2007).

Interestingly, EIF4G2 (as well as EIF4G1) seem to form cytoplasmic foci in A-431 cells (Human Protein Atlas subcell atlas). Additionally, its known binding partner, EIF3B (a component of EIF3 complex), strongly immunoprecipitated with TTC4 by antibody 2. Future work can involve further examination of the functional association between TTC4 and EIF4G2.

C7.2 Functional association of TTC4 with KIAA0430/LMKB

KIAA0430, also known as murine MARF1 (Meiosis arrest female 1), and its human orthologue Limkain B1 (LMKB1, also known as LKAP), is localized to p-bodies (Bloch et al. 2014). Although it is not considered a canonical p-body marker, LMKB formed cytoplasmic puncta that co-localized with EDC4 and DCP1, and the

immunoprecipitation of either LMKB or EDC4 co-precipitated the other (Bloch et al. 2014). Large-scale protein interaction mapping approaches have also found KIAA0430 strongly associated with p-bodies (Hein et al. 2015).

C8. TTC4 role in embryonic development

Whether TTC4 is associated with p-bodies or the retromer complex (or both) implicates its potential involvement in embryonic development. P-body components are known to have many roles in early stages of embryonic development (Schneider et al. 2006, Noble et al. 2008, Squirrell et al. 2006, Gallo et al. 2008). The retromer complex plays an important role in Wnt secretion by recycling Wntless (Belenkaya et al. 2008, Eaton 2008, de Groot et al. 2013), and proper Wnt signaling is essential for many stages of development. Thus, we decided to explore whether TTC4 is necessary for proper embryonic development using the frog *X. laevis* model system.

C8.1 Embryonic TTC4 expression

TTC4 mRNA is expressed in all stages of early embryonic development, as are p-body components (**Figure C-21**). We then used *in situ* hybridization to look for tissue-specific expression of TTC4 (**Figure C-22**). We found that TTC4 expression in NF stage 15 embryos was localized to areas involved in the formation of neural patterning. As expected, in later stage embryos, we found that TTC4 expression was strongly localized to the head and spinal column regions, suggesting an important role in vertebrate embryonic development.

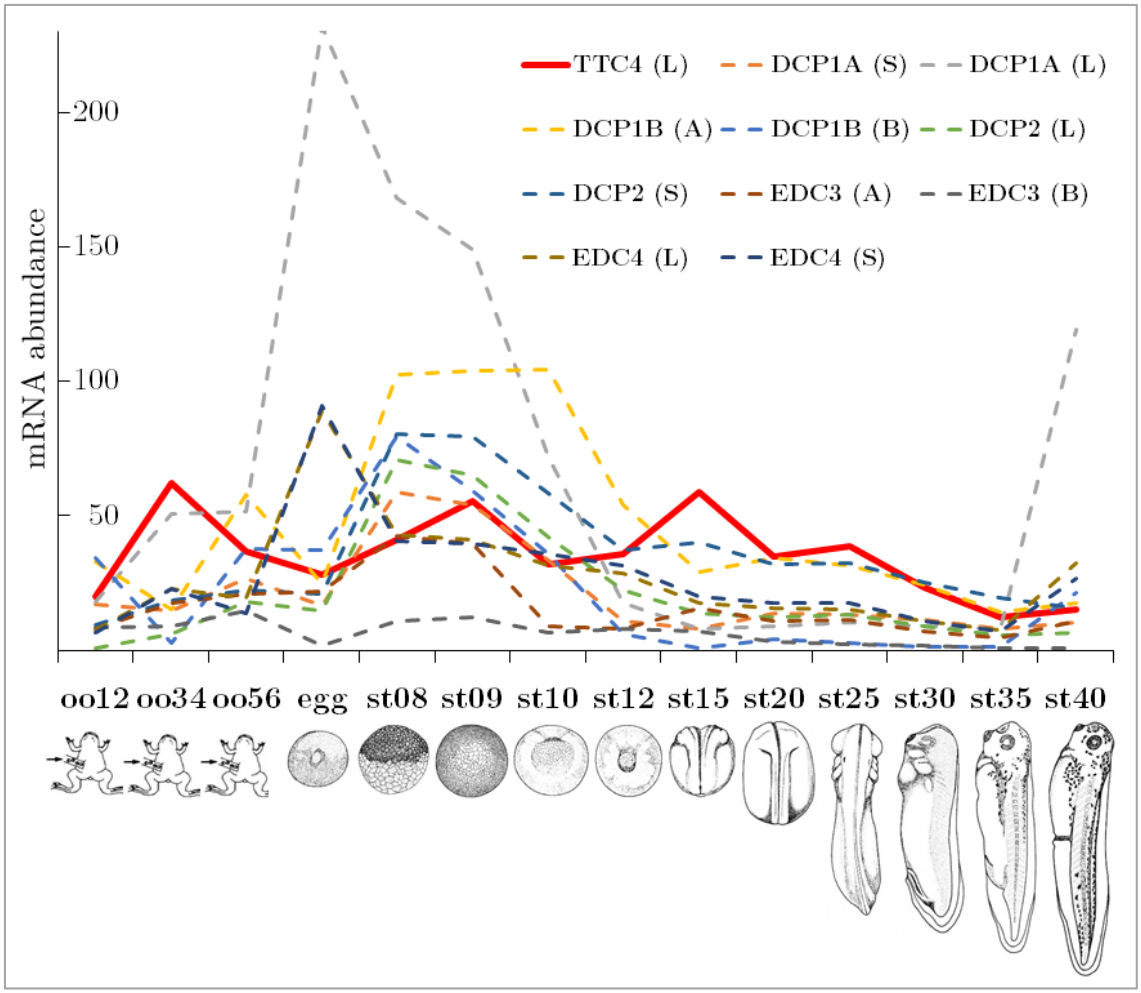

Figure C-21. Levels of TTC4 and p-body markers in developing *X. laevis*

mRNA expression levels of TTC4 and p-body markers DCP1A, DCP1B, DCP2, EDC3, and EDC4 are measured by RNA sequencing. *X. laevis* is allotetraploid; many genes are duplicated as homeologs. L and S correspond to homeologs from *laevis* subgenomes (long and short chromosomes, respectively). A and B designation are for cases where sub-genome determination has not been determined. Diagrams of NF stages are from XenBase.

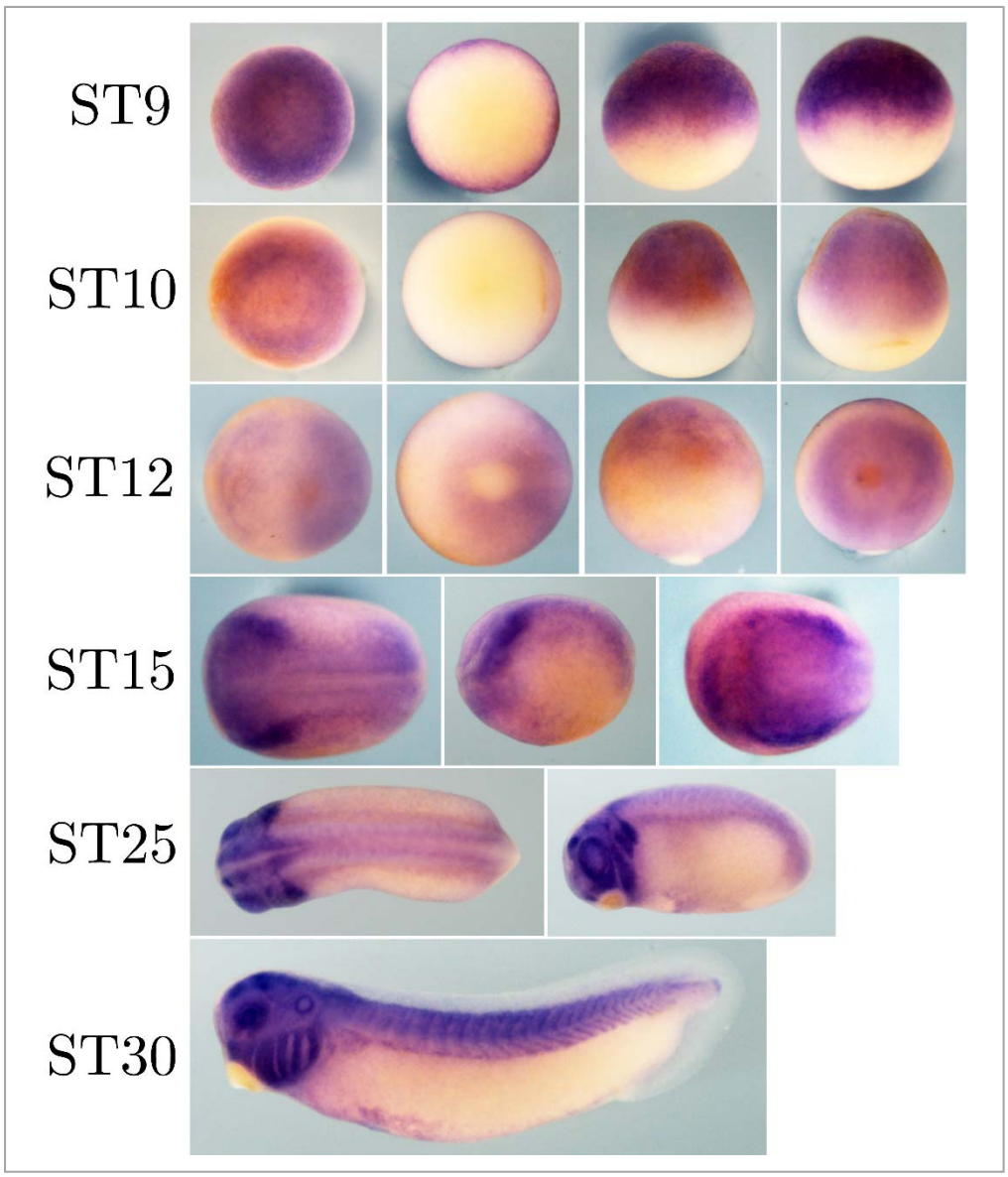

Figure C-22. TTC4 expression in embryonic development

Whole-mount *in situ* hybridization of TTC4 transcripts during *X. laevis* development with indicated Nieuwenkoop-Faber stages 9-30. TTC4 expression is indicated by digoxigenin-labeled riboprobe. Stages 9-12, left to right: animal view, vegetal view, ventral view, dorsal view. Stages 9, 10: TTC4 expression is evenly distributed in the animal top. Stage 12: TTC4 expression is enriched in neural pattern formation. Anterior is to the left in ST15+. Stage 15, left to right: dorsal view, lateral view, anterior view. Stage 25, left and right: dorsal view and lateral view. Expression is seen in the neural tube in stage 15, in the head region in stage 25, and prominent in the head and spinal cord in stage 30 embryos.

C8.2 TTC4 is indispensable for head and eye formation

TTC4 knockdown during early embryogenesis produced embryos with severe morphological defects (**Figure C-23**). Most notably, eyes were absent in these embryos. Further experimentation should address the mechanisms and pathways in which TTC4 is involved during embryonic development.

Figure C-23. Knockdown and overexpression of TTC4 in frog embryos.

Top: Knockdown analysis. Compared to controls (**A**), knockdown of TTC4 (**B**) with 8ng anti-sense morpholino injected into the dorsal cells at the 4-cell stage produced viable late-stage embryos that lacked eyes and displayed slightly abnormal overall morphology. All were injected or co-injected with 100pg memGFP mRNA. **Bottom**: Overexpression analysis. Embryos injected at the 4 cell stage with 100pg memGFP (**C**) or with 500pg GFP-TTC4 (**D**) look similar. (**E**) Positive control embryos injected with 50pg DKK1-GFP (dickkopf wnt signaling pathway inhibitor 1; a wnt antagonist) show severe developmental defects. All injections were done at the dorsal side of the 4-cell stage.

C9. Discussion

Protein tagging is a powerful and widely used method to study protein localization, interaction, and function. Through the efforts of characterizing TTC4, I demonstrated that the many commonly used fluorescent and small epitope tags may have interfered with the system that was being studied. Whether by immunofluorescence imaging to determine subcellular localization or by biochemical assays such as co-purification to determine interaction partners, tagged TTC4 produced vastly different results compared to the untagged form of the protein. Fluorescent imaging of tagged homologous TTC4 proteins in other systems (yeast and frog) suggests that the tag may be problematic as well, although this has not been studied in depth to the extent shown here for human cells. Past literature characterizing the TTC4 protein product have all used tagged recombinant TTC4, and should be reconsidered. For the same reasons, many localization datasets or interactome maps employing fluorescent tags (Huh et al. 2003, Hein et al. 2015) or small epitopes (Huttlin et al. 2015) should be carefully considered.

Orthogonal tag-free, label-free methods characterizing endogenous TTC4 has also produced confounding data. The biochemical fractionations that motivated us to characterize TTC4 inferred that TTC4 physically interacts with retromers, however RNAseq under TTC4 depletion suggested genetic interactions with a number of proteins previously shown to associate with mRNA processing bodies. This latter association between TTC4 and p-bodies is supported by several immunologic assays shown here, which suggests that TTC4 and p-bodies are likely a biologically true association.

Although interactions between retromer and TTC4 were not observed in any experimental capacity presented here, it does not necessarily rule out their possible interaction suggested from the large-scale biochemical fractionation dataset. For a variety of reasons, such as the inability of some interactions to survive our experimental conditions post-cell lysis, failure to observe many protein interactions is to be expected. For example, co-fractionation of canonical p-body members were not observed in our large-scale biochemical fractionations. Another limitation due to the methodology used in our fractionation experiments results in the poor observation of membrane complexes, for example, the robust nuclear pore complex was not observed.

However, it is also the case that some physical interactions inferred from the co-elution profiles are false positives, meaning their inferred interactions have no true biological basis. Two functionally unrelated or non-interacting proteins may have correlated elution profiles simply due to chance that they exhibit similar biophysical properties (properties that are exploited to achieve protein complex separation). While this may be the case for the TTC4-retromer observation, it is also possible that, like many moonlighting proteins, TTC4 participates in many biological processes, and that the experimental data presented here only samples a subpopulation or a context-dependent role of TTC4. Immunoprecipitation of TTC4 also strongly co-precipitated numerous uncharacterized proteins, such as TPP2 and TEX2, providing interesting starting points for further characterization. Further experiments can also focus on whether the different antibodies that produced differential results—within this manuscript and the differential antibody-derived

TTC4 data in literature (Dmitriev et al. 2009)—capture separate subpopulations of TTC4.

C10. Methods

C10.1 Cloning

Human cDNAs were obtained from the human ORFeome collection (OpenBiosystems) and cloned using Gateway cloning into either the Vivid Colors pcDNA6.2/C-EmGFP-DEST Vector (Thermo Scientific) for carboxy-terminal GFP-tagged expression clones, or the Vivid Colors pcDNA6.2/N-EmGFP-DEST Vector (Thermo Scientific) for amino-terminal GFP-tagged expression clones (both emGFP vectors gifted by Dr. Kyle Miller), or the pTagRFP-N plasmid (Evrogen; modified to introduce the Invitrogen pDEST47 Gateway cassette) for carboxy-terminal RFP-tagged expression clones, or the pcDNA3-6MYC plasmid modified to introduce the Invitrogen pDEST47 Gateway cassette for carboxy-terminal MYC expression clones (acquired from Dr. Zhihua Li), or the pcDNA3-3FLAG plasmid modified to introduce the Invitrogen pDEST47 Gateway cassette for carboxy-terminal FLAG expression clones (acquired from Dr. Zhihua Li).

For *in situ* hybridization, digoxigenin (DIG)-labeled RNA probes against *X. laevis* TTC4 were generated with the following primers:
TTC4 forward primer: caccATGGATCCAAAAGACCAAGAGG
TTC4 reverse primer: CTATGCAGAAGGCAGTCTATAAACTCTC

For TTC4 over-expression in *X. laevis*, the same TTC4 primers above were used to clone TTC4 from a cDNA library (generated by *in vitro* reverse transcription of total RNA extracted from frog embryos) that was then cloned into the pENTR/D-TOPO kit (Thermo Scientific) and then into the expression vector 223-pCS-EGFP-DEST (Addgene).

C10.2 Human cell culture

HeLa, HEK293T, and CRL-2091 cells were obtained from the American Type Culture Collection (ATCC) were cultured at 37°C with 5% CO_2 in DMEM medium (Sigma D6429) supplemented with 10% FBS (fetal bovine serum). H-460, HCC4017, and HBEC30KT cells were gifted from Dr. Deepak Nijhawan (UT Southwestern) and cultured in RPMI with 10% FBS. Cells were treated with geldanamycin, cycloheximide, sodium arsenite, or mocks, as described in the text.

C10.3 Transfections

For plasmid transfections: One day prior to transfection, HeLa or HEK293T cells were plated in 96-well or 6-well glass bottom plates (for imaging) or normal plates (for lysate collection). Plasmids were transfected into cells using FuGENE HD and cells were imaged the next day.

For RNAi transfections, both forward and reverse transfections were attempted with similar knockdown efficiencies. For forward transfections, HeLa or

HEK293T cells were plated the previous day in 96-well or 6-well glass bottom plates (for imaging) or normal tissue culture plates (for sample collection). Transfection was carried out with Lipofectamine RNAiMAX or 2000 (Thermo Scientific) the following day according to the manufacturer's manual. For reverse transfections, HEK293T cells were plated onto transfection mixtures following the manufacturer's protocol with Lipofectamine RNAiMAX or 2000. HEK293T samples collected for transcriptome profiling were all transfected in the reverse fashion with RNAiMAX in 6-well tissue culture plates. Samples were collected ~27-29 hours post-transfection. siRNAs used were: anti-TTC4 (human, sc-88730, Santa Cruz Biotech.) and Silencer Negative Control No. 1 siRNA (AM4611, Thermo Scientific).

C10.4 Immunofluorescence

For immunofluorescence experiments, cells were fixed with 3.7% methanol-free formaldehyde freshly diluted from 16% stock (28908, Thermo Scientific) at 37°C for 15-20 minutes, blocked with 5% goat serum in PBS-T buffer for 30-60 minutes at room temperature, then incubated with primary antibody overnight at 4°C. Cells were washed with PBS buffer, then incubated with secondary antibody for 1-2 hours. Primary antibodies used: anti-TTC4 (HPA041608, Sigma), anti-TTC4 (HPA042459, Sigma), anti-TTC4 (SAB1402385, Sigma), anti-EDC4 (HPA041164, Sigma). The immunogen sequences for anti-TTC4 antibodies are listed below. Secondary antibodies used: Alexa Fluor 488-conjugated goat anti-rabbit (Invitrogen), Alexa Fluor 647-conjugated goat anti-rabbit (Invitrogen). All antibodies were used at the manufacturer's recommended concentrations.

Anti-TTC4 Ab 1 (HPA041608, Sigma) immunogen sequence:
AQSDFISAFHEDSRFIDHLMVMFGETPSWDLEQKYCPDNLEVYFEDEDRAEL
YRVPAKSTLLQVLQHQRYFVKALTPAFLVCVGSSPFCKNFLRGRKVYQI
Anti-TTC4 Ab 2 (HPA042459, Sigma) immunogen sequence:
LEMRAKADKLKRIEQRDVRKANLKEKKERNQNEALLQAIKARNIRLSEAACE
DEDSASEGLGELFLDGLSTENPHGARLSLDGQGRLSWPVLFLYP
Anti-TTC4 Ab 3 (SAB1402385, Sigma) immunogen sequence:
MEQPGQDPTSDDVMDSFLEKFQSQPYRGGFHEDQWEKEFEKVPLFMTRAP
SEIDPRENPDLACLQSIIFDEERSPEEQAKTYKDEGNDYFKEKDYKKAVI

C10.5 Western blot

For western blotting, cells were washed with PBS buffer then sampled collected with Laemmli sample buffer containing 5% beta-mercaptoethanol. Total protein were separated on commercial precast gels, then blotted with PVDF membrane, followed by blocking with 5% milk in TBS-T (tris-buffered saline with 0.1% Tween 20). Primary antibodies used include those listed above, and additionally: anti-End-binding protein 3 (EB3, AB6033, Millipore), anti-Lamin A (sc-20680, Santa Cruz Biotech.), anti-actin (sc-1616, Santa Cruz Biotech.) Secondary antibodies: goat anti-rabbit IgG F(ab')$_2$-HRP (sc-3837, Santa Cruz Biotech.) and goat anti-goat IgG F(ab')$_2$-HRP (sc-3697, Santa Cruz Biotech.), and donkey anti-goat IgG F(ab')$_2$-HRP (sc-3851, Santa Cruz Biotech.) Primary antibodies were incubated overnight at 4°C or at ambient temperature for 1-2 hours. Secondary

antibodies were incubated at ambient temperature for 1-2 hours, with TBS-T washes before and after. Signal was illuminated with luminol and film-developed.

C10.6 Fluorescent cell microscopy

Live cells were imaged using a Nikon Eclipse TE2000-E inverted microscope. Images were acquired using a Photometrics Cascade II 512 camera and Nikon Plan Apo 40x/0.95, 60x/0.95, 60x/1.40, or 100x/1.40 objectives. The 89000 - ET - Sedat Quad (Chroma Tech.) filter set was used for all fluorescence detection. Image acquisition were driven by NIS Elements AR Imaging software.

C10.7 Immunoprecipitation and mass spectrometry

HeLa or HEK293T cells were transfected with appropriate expression vectors for APs or left untransfected for endogenous IPs. The next day, cells were washed once in PBS then trypsinized and resuspended in media to inactivate trypsin, spun down at 1000-1500RPM for 5-10 minutes, and then resuspended in 1X cell lysis buffer (Cell Signaling) supplemented with 1X protease inhibitor I (Roche) for 10-30 minutes before sonicating with Diagenode Bioruptor until at least ~90% of cells were visibly broken. Cell lysate was incubated on rotation at room temperature with Protein A Dynabeads (Thermo Scientific) pre-washed with PBS and bound to the appropriate antibody. Antibodies used include those previously mentioned, in addition to anti-GFP (G1544, Sigma) and anti-FLAG (F1804, Sigma). After washing the Dynabead-antibody-antigen complex with PBS, proteins were eluted from the beads in 2% SDS

in 0.1M Tris/HCl pH7.6 with 100mM DTT at 60C and then stored at -20°C. For FLAG pulldowns, anti-FLAG M2 affinity gel (A2220, Sigma) was also used and gave better results than the antibody/Dynabeads method. For anti-FLAG gel, proteins were eluted with 3X FLAG peptides.

For mass spectrometry sample preparation, DTT was added to 0.1M to each sample, and washes of the sample in 8M urea (UA; in 0.1M Tris/HCl pH8.5) were passed over Amicon Ultra 30 kD spin filtration columns (pre-conditioned with by washing with 0.5% PEG20K). Iodoacetaminde was then added to 55mM and incubated in dark for 40 minutes at RT. Sample was then washed again in UA, then equilibrated with washes of 0.05M ammonium bicarbonate (ABC). Samples were then subject to overnight trypsin digest at 37°C. Tryptic peptides were filtered through the spin filtration columns and bound, washed, and eluted from HyperSep C18 SpinTips (Thermo Scientific). Eluted peptides were dried by speed-vac and resuspended in Buffer C (5% acetonitrile, 0.1% formic acid) for analysis by LC-MS/MS. For LC-MS/MS analysis, peptides were subjected to separation by C18 reverse phase chromatography on a Dionex Ultimate 3000 RSLCnano UHPLC system (Thermo Scientific). Peptides were loaded onto an Acclaim C18 PepMap RSLC column (Dionex; Thermo Scientific) and eluted using an acetonitrile gradient. Eluted peptides were directly injected into an Orbitrap Classic or Orbitrap Elite mass spectrometer (Thermo Scientific) by nano-electrospray and subject to data-dependent tandem mass spectrometry.

Spectra were searched against a human protein sequence database and common contaminant proteins (MaxQuant using SEQUEST (Proteome Discoverer 1.4; Thermo Scientific)). Fully-tryptic peptides were considered, with up to two missed cleavages. High-confidence peptide-spectral matches (PSMs) were filtered at <1% false discovery rate determined by Percolator (Proteome Discoverer 1.4; Thermo Scientific). Fold changes for each protein were calculated as the ratio of pseudo-PSMs (PSM + 1) in the pulldown sample over that in the mock.

C10.8 Transcriptome sequencing of human cells

RNA was extracted from HEK293T cells grown in 6 well-plates using the Trizol-chloroform method. Sets of OptiMEM-only, scrambled knockdown, and TTC4 knockdown samples were prepared as described above and collected twice on separate occasions. 8 individual samples were subject to RNAseq, including 4 control samples consisting of two OptiMEM-only and two mock knockdown with scrambled siRNA, and 4 TTC4 knockdown samples. Samples were depleted of DNA by DNase extraction and RNeasy column purification. Eluted RNA was processed with Ribo-Zero Gold rRNA removal kit (Illumina) followed by additional RNeasy MinElute clean up. RNA quality was evaluated with Agilent Bioanalyzer (Agilent Technologies). cDNA libraries were constructed using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina. Single-end (1 x 75 base) sequencing was performed using a NextSeq 500 instrument (Illumina). RNA-seq data was processed with Kallisto (Bray et al. 2016) to generate abundances which were then

analyzed and mapped to transcript identities using sleuth (http://pachterlab.github.io/sleuth/).

C10.9 *In situ* hybridization

X. laevis embryos from various developmental stages were fixed in MEMFA overnight and *in situ* hybridization and data acquisition was performed as previously described (Hayes et al. 2007).

C10.10 *Xenopus laevis* perturbation experiments

For live animal perturbation experiments, frog embryos were injected with 100pg memGFP or indicated concentrations of recombinant constructs into dorsal cells at the 4-cell stage. Morpholinos (MOs) were injected in the same manner.
Anti-TTC4 ATG-MO: TCAATGGTTTTCACGATTTCTGCAC
Anti-TTC4 SP-MO: TACTATGGCATCTCCCTTACCTAAA

C10.11 *Xenopus laevis* RNA-seq

Collection of large scale J-strain *X. laevis* transcriptome resources

We complemented the genome sequence with more than 1 billion RNAseq reads that sample a useful range of developmental stages and adult organs and tissues, as summarized in Table S6. For RNAseq, RNA was extracted from a series of developmental stages, or from a collection of adult tissues. Both stages and tissue

samples were collected twice independently. Embryos from fourteen different developmental stages (including 3 oocyte stages, unfertilized egg, and st8 to NF stage 40) (J-strain 34th generation, cultured at 20 degrees) were collected from one pair of 33rd generation frogs. Thirteen adult tissues and oocytes of different stages (stages I & II, III & IV, V & VI) were collected from a single female, and a testis was harvested from a single male (J-strain 33rd generation).

RNA-seq

Total RNA was extracted using Isogen (Nippon Gene). Quality of the total RNA was evaluated by a spectrophotometer and Agilent 2100 Bioanalyzer (Agilent Technologies). cDNA libraries were constructed using Illumina Truseq RNA sample prep kit V2 (Illumina), with the standard non-strand specific mRNA library preparation protocol. Independent samplings were performed from embryos of two crossings or from two female and male adults, separately, providing cDNA library sets for Taira201203_stage, Taira201203_tissue, Ueno201210_stage, Ueno201210_tissue series of RNAseqs. Additionally, to add reads to Ueno201210_stage (for stage 35), their siblings were analyzed to produce Ueno201302_stage series.

Paired-end (100 bp × 2 101 bp × 2) sequencing was performed using an Illumina HiSeq 2000 instrument (Illumina). Datasets of the short reads were deposited in NCBI Gene Expression Omnibus (GEO) database (accession number

GSE73430 for stages, GSE73419 for tissues). These RNA-seq data were used for expression analysis described below.

Quantification of gene expression levels with RNA-seq

We analyzed gene expression of the RNAseq data described above for a developmental time series and selected adult tissues. After filtering (1) reads with no call ('N') and (2) reads with low complexity (not having all of 'A', 'C', 'G, and 'T') from raw J-strain RNA-seq reads, we mapped them to primary transcript sequences using bwa mem (version 0.7.10) with paired-end option. We quantified the expression of each transcript using Transcripts Per Million (TPM) values estimated by RSEM (version 1.2.19) (Li and Dewey 2011).

To prevent the noise derived from reads of homeologous transcripts, we removed hits either (1) with additional targets including homeologs, or (2) with partial alignment with insertions/deletions (indels), before running RSEM. As a result, we used highly specific reads mapped only on one copy of homeolog transcripts in this analysis. This approach may underestimate the expression of homeologous transcripts by ignoring reads from identical regions, although the expression of most homeologs can be measured by taking advantage of paired-end reads. All redundant sequences were removed in the database before mapping, to measure at least group-level expression of redundant sequences.

References

Abraham, V. C., D. L. Taylor, and J. R. Haskins. 2004. "High content screening applied to large-scale cell biology." *Trends Biotechnol* 22 (1):15-22. doi: 10.1016/j.tibtech.2003.10.012.

Ahmad, F., P. M. Ahmad, L. Pieretti, and G. T. Watters. 1978. "Purification and subunit structure of rat mammary gland acetyl coenzyme A carboxylase." *J Biol Chem* 253 (5):1733-7.

Ahmad, K. A., G. Wang, G. Unger, J. Slaton, and K. Ahmed. 2008. "Protein kinase CK2--a key suppressor of apoptosis." *Adv Enzyme Regul* 48:179-87. doi: 10.1016/j.advenzreg.2008.04.002.

Ahmed, M., J. Davis, D. Aucoin, T. Sato, S. Ahuja, S. Aimoto, J. I. Elliott, W. E. Van Nostrand, and S. O. Smith. 2010. "Structural conversion of neurotoxic amyloid-beta(1-42) oligomers to fibrils." *Nat Struct Mol Biol* 17 (5):561-7. doi: 10.1038/nsmb.1799.

Aimi, J., H. Qiu, J. Williams, H. Zalkin, and J. E. Dixon. 1990. "De novo purine nucleotide biosynthesis: cloning of human and avian cDNAs encoding the trifunctional glycinamide ribonucleotide synthetase-aminoimidazole ribonucleotide synthetase-glycinamide ribonucleotide transformylase by functional complementation in E. coli." *Nucleic Acids Res* 18 (22):6665-72.

Albeck, J. G., G. MacBeath, F. M. White, P. K. Sorger, D. A. Lauffenburger, and S. Gaudet. 2006. "Collecting and organizing systematic sets of protein data." *Nat Rev Mol Cell Biol* 7 (11):803-12. doi: 10.1038/nrm2042.

Alberti, S., R. Halfmann, O. King, A. Kapila, and S. Lindquist. 2009. "A systematic survey identifies prions and illuminates sequence features of prionogenic proteins." *Cell* 137 (1):146-58. doi: 10.1016/j.cell.2009.02.044.

Alvira, S., J. Cuellar, A. Rohl, S. Yamamoto, H. Itoh, C. Alfonso, G. Rivas, J. Buchner, and J. M. Valpuesta. 2014. "Structural characterization of the substrate transfer mechanism in Hsp70/Hsp90 folding machinery mediated by Hop." *Nat Commun* 5:5484. doi: 10.1038/ncomms6484.

An, S., Y. Deng, J. W. Tomsho, M. Kyoung, and S. J. Benkovic. 2010. "Microtubule-assisted mechanism for functional metabolic macromolecular complex formation." *Proc Natl Acad Sci U S A* 107 (29):12872-6. doi: 10.1073/pnas.1008451107.

An, S., R. Kumar, E. D. Sheets, and S. J. Benkovic. 2008. "Reversible compartmentalization of de novo purine biosynthetic complexes in living cells." *Science* 320 (5872):103-6. doi: 10.1126/science.1152241.

An, S., M. Kyoung, J. J. Allen, K. M. Shokat, and S. J. Benkovic. 2010. "Dynamic regulation of a metabolic multi-enzyme complex by protein kinase CK2." *J Biol Chem* 285 (15):11093-9. doi: 10.1074/jbc.M110.101139.

Anderson, Paul, and Nancy Kedersha. 2009. "Stress granules." *Current Biology* 19 (10):R397-R398. doi: 10.1016/j.cub.2009.03.013.

Angermuller, S., G. Bruder, A. Volkl, H. Wesch, and H. D. Fahimi. 1987. "Localization of xanthine oxidase in crystalline cores of peroxisomes. A cytochemical and biochemical study." *Eur J Cell Biol* 45 (1):137-44.

Aouffen, M., J. Paquin, A. Furtos, K. C. Waldron, and M. A. Mateescu. 2004. "Oxidative aggregation of ceruloplasmin induced by hydrogen peroxide is prevented by pyruvate." *Free Radic Res* 38 (1):19-26.

Aughey, G. N., S. J. Grice, Q. J. Shen, Y. Xu, C. C. Chang, G. Azzam, P. Y. Wang, L. Freeman-Mills, L. M. Pai, L. Y. Sung, J. Yan, and J. L. Liu. 2014. "Nucleotide synthesis is regulated by cytoophidium formation during neurodevelopment and adaptive metabolism." *Biol Open* 3 (11):1045-56. doi: 10.1242/bio.201410165.

Aughey, Gabriel N., and Ji-Long Liu. 2016. "Metabolic regulation via enzyme filamentation." *Critical Reviews in Biochemistry and Molecular Biology*:1-12. doi: 10.3109/10409238.2016.1172555.

Bader, G. D., and C. W. Hogue. 2002. "Analyzing yeast protein-protein interaction data obtained from different sources." *Nat Biotechnol* 20 (10):991-7. doi: 10.1038/nbt1002-991.

Bailey, S. N., S. M. Ali, A. E. Carpenter, C. O. Higgins, and D. M. Sabatini. 2006. "Microarrays of lentiviruses for gene function screens in immortalized and primary cells." *Nat Methods* 3 (2):117-22. doi: 10.1038/nmeth848.

Baird, G. S., D. A. Zacharias, and R. Y. Tsien. 2000. "Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral." *Proc Natl Acad Sci U S A* 97 (22):11984-9. doi: 10.1073/pnas.97.22.11984.

Bak, D., G. R. Cutting, and M. Milewski. 2007. "The CFTR-derived peptides as a model of sequence-specific protein aggregation." *Cell Mol Biol Lett* 12 (3):435-47. doi: 10.2478/s11658-007-0014-1.

Baresova, V., V. Skopova, J. Sikora, D. Patterson, J. Sovova, M. Zikanova, and S. Kmoch. 2012. "Mutations of ATIC and ADSL affect purinosome assembly in cultured skin fibroblasts from patients with AICA-ribosiduria and ADSL deficiency." *Hum Mol Genet* 21 (7):1534-43. doi: 10.1093/hmg/ddr591.

Barry, R. M., A. F. Bitbol, A. Lorestani, E. J. Charles, C. H. Habrian, J. M. Hansen, H. J. Li, E. P. Baldwin, N. S. Wingreen, J. M. Kollman, and Z. Gitai. 2014. "Large-scale filament formation inhibits the activity of CTP synthetase." *Elife* 3:e03638. doi: 10.7554/eLife.03638.

Beaty, N. B., and M. D. Lane. 1983a. "Kinetics of activation of acetyl-CoA carboxylase by citrate. Relationship to the rate of polymerization of the enzyme." *J Biol Chem* 258 (21):13043-50.

Beaty, N. B., and M. D. Lane. 1983b. "The polymerization of acetyl-CoA carboxylase." *J Biol Chem* 258 (21):13051-5.

Becker, M. A., and M. Kim. 1987. "Regulation of purine synthesis de novo in human fibroblasts by purine nucleotides and phosphoribosylpyrophosphate." *J Biol Chem* 262 (30):14531-7.

Beguin, P., and M. Lemaire. 1996. "The cellulosome: an exocellular, multiprotein complex specialized in cellulose degradation." *Crit Rev Biochem Mol Biol* 31 (3):201-36. doi: 10.3109/10409239609106584.

Belenkaya, T. Y., Y. Wu, X. Tang, B. Zhou, L. Cheng, Y. V. Sharma, D. Yan, E. M. Selva, and X. Lin. 2008. "The retromer complex influences Wnt secretion by recycling wntless from endosomes to the trans-Golgi network." *Dev Cell* 14 (1):120-31. doi: 10.1016/j.devcel.2007.12.003.

Berti, C., B. Fontanella, R. Ferrentino, and G. Meroni. 2004. "Mig12, a novel Opitz syndrome gene product partner, is expressed in the embryonic ventral midline and co-operates with Mid1 to bundle and stabilize microtubules." *BMC Cell Biol* 5:9. doi: 10.1186/1471-2121-5-9.

Bibby, A. C., and D. W. Litchfield. 2005. "The multiple personalities of the regulatory subunit of protein kinase CK2: CK2 dependent and CK2 independent roles reveal a secret identity for CK2beta." *Int J Biol Sci* 1 (2):67-79.

Bjorkoy, G., T. Lamark, A. Brech, H. Outzen, M. Perander, A. Overvatn, H. Stenmark, and T. Johansen. 2005. "p62/SQSTM1 forms protein aggregates degraded by autophagy and has a protective effect on huntingtin-induced cell death." *J Cell Biol* 171 (4):603-14. doi: 10.1083/jcb.200507002.

Bloch, D. B., P. Li, E. G. Bloch, D. F. Berenson, R. L. Galdos, P. Arora, R. Malhotra, C. Wu, and W. Yang. 2014. "LMKB/MARF1 localizes to mRNA processing bodies, interacts with Ge-1, and regulates IFI44L gene expression." *PLoS One* 9 (4):e94784. doi: 10.1371/journal.pone.0094784.

Bobik, T. A., B. P. Lehman, and T. O. Yeates. 2015. "Bacterial microcompartments: widespread prokaryotic organelles for isolation and optimization of metabolic pathways." *Mol Microbiol* 98 (2):193-207. doi: 10.1111/mmi.13117.

Bock, E., and G. Heinrich. 1971. "[Structural and functional changes in reactivating cells of Nitrobacter winogradskyi Buch]." *Arch Mikrobiol* 77 (4):349-65.

Bonacci, W., P. K. Teng, B. Afonso, H. Niederholtmeyer, P. Grob, P. A. Silver, and D. F. Savage. 2012. "Modularity of a carbon-fixing protein organelle." *Proc Natl Acad Sci U S A* 109 (2):478-83. doi: 10.1073/pnas.1108557109.

Bray, N. L., H. Pimentel, P. Melsted, and L. Pachter. 2016. "Near-optimal probabilistic RNA-seq quantification." *Nat Biotechnol*. doi: 10.1038/nbt.3519.

Brengues, M., and R. Parker. 2007. "Accumulation of polyadenylated mRNA, Pab1p, eIF4E, and eIF4G with P-bodies in Saccharomyces cerevisiae." *Mol Biol Cell* 18 (7):2592-602. doi: 10.1091/mbc.E06-12-1149.

Brengues, M., D. Teixeira, and R. Parker. 2005. "Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies." *Science* 310 (5747):486-9. doi: 10.1126/science.1115791.

Brinsmade, S. R., T. Paldon, and J. C. Escalante-Semerena. 2005. "Minimal functions and physiological conditions required for growth of salmonella enterica on ethanolamine in the absence of the metabolosome." *J Bacteriol* 187 (23):8039-46. doi: 10.1128/JB.187.23.8039-8046.2005.

Broadley, S. A., and F. U. Hartl. 2009. "The role of molecular chaperones in human misfolding diseases." *FEBS Lett* 583 (16):2647-53. doi: 10.1016/j.febslet.2009.04.029.

Buchan, J. Ross, Denise Muhlrad, and Roy Parker. 2008. "P bodies promote stress granule assembly in Saccharomyces cerevisiae." *The Journal of Cell Biology* 183 (3):441-455. doi: 10.1083/jcb.200807043.

Camp, R. L., V. Neumeister, and D. L. Rimm. 2008. "A decade of tissue microarrays: progress in the discovery and validation of cancer biomarkers." *J Clin Oncol* 26 (34):5630-7. doi: 10.1200/JCO.2008.17.3567.

Campbell, R. E., O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias, and R. Y. Tsien. 2002. "A monomeric red fluorescent protein." *Proc Natl Acad Sci U S A* 99 (12):7877-82. doi: 10.1073/pnas.082243699.

Campbell, S. G., N. P. Hoyle, and M. P. Ashe. 2005. "Dynamic cycling of eIF2 through a large eIF2B-containing cytoplasmic body: implications for translation control." *J Cell Biol* 170 (6):925-34. doi: 10.1083/jcb.200503162.

Cannon, G. C., R. S. English, and J. M. Shively. 1991. "In situ assay of ribulose-1,5-bisphosphate carboxylase/oxygenase in Thiobacillus neapolitanus." *J Bacteriol* 173 (4):1565-8.

Caperelli, C. A. 1985. "Mammalian glycinamide ribonucleotide transformylase: purification and some properties." *Biochemistry* 24 (6):1316-20.

Caperelli, C. A., P. A. Benkovic, G. Chettur, and S. J. Benkovic. 1980. "Purification of a complex catalyzing folate cofactor synthesis and transformylation in de novo purine biosynthesis." *J Biol Chem* 255 (5):1885-90.

Carcamo, W. C., M. Satoh, H. Kasahara, N. Terada, T. Hamazaki, J. Y. Chan, B. Yao, S. Tamayo, G. Covini, C. A. von Muhlen, and E. K. Chan. 2011. "Induction of cytoplasmic rods and rings structures by inhibition of the CTP and GTP synthetic pathway in mammalian cells." *PLoS One* 6 (12):e29690. doi: 10.1371/journal.pone.0029690.

Carcamo, Wendy C., S. John Calise, Carlos A. von Mühlen, Minoru Satoh, and Edward K. L. Chan. 2014. "Chapter Two - Molecular Cell Biology and Immunobiology of Mammalian Rod/Ring Structures." In *International Review of Cell and Molecular Biology*, edited by W. Jeon Kwang, 35-74. Academic Press.

Carpenter, A. E., T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini. 2006. "CellProfiler: image analysis software for identifying and quantifying cell phenotypes." *Genome Biol* 7 (10):R100. doi: 10.1186/gb-2006-7-10-r100.

Castel, D., A. Pitaval, M. A. Debily, and X. Gidrol. 2006. "Cell microarrays in drug discovery." *Drug Discov Today* 11 (13-14):616-22. doi: 10.1016/j.drudis.2006.05.015.

Cerami, A., and C. M. Peterson. 1975. "Cyanate and sickle-cell disease." *Sci Am* 232 (4):44-50.

Chai, Y., S. L. Koppenhafer, N. M. Bonini, and H. L. Paulson. 1999. "Analysis of the role of heat shock protein (Hsp) molecular chaperones in polyglutamine disease." *J Neurosci* 19 (23):10338-47.

Chang, C. C., W. C. Lin, L. M. Pai, H. S. Lee, S. C. Wu, S. T. Ding, J. L. Liu, and L. Y. Sung. 2015. "Cytoophidium assembly reflects upregulation of IMPDH activity." *J Cell Sci* 128 (19):3550-5. doi: 10.1242/jcs.175265.

Chen, K., J. Zhang, O. Y. Tastan, Z. A. Deussen, M. Y. Siswick, and J. L. Liu. 2011. "Glutamine analogs promote cytoophidium assembly in human and Drosophila cells." *J Genet Genomics* 38 (9):391-402. doi: 10.1016/j.jgg.2011.08.004.

Chen, P., D. I. Andersson, and J. R. Roth. 1994. "The control region of the pdu/cob regulon in Salmonella typhimurium." *J Bacteriol* 176 (17):5474-82.

Choder, M. 1991. "A general topoisomerase I-dependent transcriptional repression in the stationary phase in yeast." *Genes Dev* 5 (12A):2315-26.

Cioni, P., and G. B. Strambini. 1989. "Dynamical structure of glutamate dehydrogenase as monitored by tryptophan phosphorescence. Signal transmission following binding of allosteric effectors." *J Mol Biol* 207 (1):237-47.

Clarke, B. A., and S. D. Clarke. 1982. "Polymer-protomer transition of acetyl-CoA carboxylase as a regulator of lipogenesis in rat liver." *Arch Biochem Biophys* 218 (1):92-100.

Costantini, L. M., M. Baloban, M. L. Markwardt, M. Rizzo, F. Guo, V. V. Verkhusha, and E. L. Snapp. 2015. "A palette of fluorescent proteins optimized for diverse cellular environments." *Nat Commun* 6:7670. doi: 10.1038/ncomms8670.

Costello, Joseph, Lydia M. Castelli, William Rowe, Christopher J. Kershaw, David Talavera, Sarah S. Mohammad-Qureshi, Paul F. G. Sims, Christopher M. Grant, Graham D. Pavitt, Simon J. Hubbard, and Mark P. Ashe. 2015. "Global mRNA selection mechanisms for translation initiation." *Genome Biology* 16 (1):1-21. doi: 10.1186/s13059-014-0559-z.

Crevel, G., D. Bennett, and S. Cotterill. 2008. "The human TPR protein TTC4 is a putative Hsp90 co-chaperone which interacts with CDC6 and shows alterations in transformed cells." *PLoS One* 3 (3):e0001737. doi: 10.1371/journal.pone.0001737.

Cromwell, M. E., E. Hilario, and F. Jacobson. 2006. "Protein aggregation and bioprocessing." *AAPS J* 8 (3):E572-9. doi: 10.1208/aapsj080366.

D'Andrea, L. D., and L. Regan. 2003. "TPR proteins: the versatile helix." *Trends Biochem Sci* 28 (12):655-62. doi: 10.1016/j.tibs.2003.10.007.

da Fonseca, P. C., and E. P. Morris. 2008. "Structure of the human 26S proteasome: subunit radial displacements open the gate into the proteolytic core." *J Biol Chem* 283 (34):23305-14. doi: 10.1074/jbc.M802716200.

Dai, Q., S. B. Qian, H. H. Li, H. McDonough, C. Borchers, D. Huang, S. Takayama, J. M. Younger, H. Y. Ren, D. M. Cyr, and C. Patterson. 2005. "Regulation of the cytoplasmic quality control protein degradation pathway by BAG2." *J Biol Chem* 280 (46):38673-81. doi: 10.1074/jbc.M507986200.

de Groot, R. E., H. F. Farin, M. Macurkova, J. H. van Es, H. C. Clevers, and H. C. Korswagen. 2013. "Retromer dependent recycling of the Wnt secretion factor Wls is dispensable for stem cell maintenance in the mammalian intestinal epithelium." *PLoS One* 8 (10):e76971. doi: 10.1371/journal.pone.0076971.

Decker, Carolyn J., and Roy Parker. 2012. "P-Bodies and Stress Granules: Possible Roles in the Control of Translation and mRNA Degradation." *Cold Spring Harbor Perspectives in Biology* 4 (9):a012286. doi: 10.1101/cshperspect.a012286.

DeRisi, J. L., V. R. Iyer, and P. O. Brown. 1997. "Exploring the metabolic and genetic control of gene expression on a genomic scale." *Science* 278 (5338):680-6.

Diacovich, L., D. L. Mitchell, H. Pham, G. Gago, M. M. Melgar, C. Khosla, H. Gramajo, and S. C. Tsai. 2004. "Crystal structure of the beta-subunit of acyl-CoA carboxylase: structure-based engineering of substrate specificity." *Biochemistry* 43 (44):14027-36. doi: 10.1021/bi049065v.

Ding, D. Q., Y. Tomita, A. Yamamoto, Y. Chikashige, T. Haraguchi, and Y. Hiraoka. 2000. "Large-scale screening of intracellular protein localization in living fission yeast cells by the use of a GFP-fusion genomic DNA library." *Genes Cells* 5 (3):169-90.

Dmitriev, R. I., I. A. Okkelman, R. A. Abdulin, M. I. Shakhparonov, and N. B. Pestov. 2009. "Nuclear transport of protein TTC4 depends on the cell cycle." *Cell Tissue Res* 336 (3):521-7. doi: 10.1007/s00441-009-0785-y.

Dmitriev, R. I., N. B. Pestov, M. I. Shakhparonov, and I. A. Okkelman. 2014. "Two distinct nuclear localization signals in mammalian MSL1 regulate its function." *J Cell Biochem* 115 (11):1967-73. doi: 10.1002/jcb.24868.

Dobler, J., and J. F. Bertles. 1968. "The physical state of hemoglobin in sickle-cell anemia erythrocytes in vivo." *J Exp Med* 127 (4):711-4.

Dou, Z., S. Heinhorst, E. B. Williams, C. D. Murin, J. M. Shively, and G. C. Cannon. 2008. "CO2 fixation kinetics of Halothiobacillus neapolitanus mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO2." *J Biol Chem* 283 (16):10377-84. doi: 10.1074/jbc.M709285200.

Drews, G., and W. Niklowitz. 1956. "[Cytology of Cyanophycea. II. Centroplasm and granular inclusions of Phormidium uncinatum]." *Arch Mikrobiol* 24 (2):147-62.

Drummond, D. A., and C. O. Wilke. 2008. "Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution." *Cell* 134 (2):341-52. doi: 10.1016/j.cell.2008.05.042.

Drummond, D. A., and C. O. Wilke. 2009. "The evolutionary consequences of erroneous protein synthesis." *Nat Rev Genet* 10 (10):715-24. doi: 10.1038/nrg2662.

Eaton, S. 2008. "Retromer retrieves wntless." *Dev Cell* 14 (1):4-6. doi: 10.1016/j.devcel.2007.12.014.

Edskes, H. K., L. M. McCann, A. M. Hebert, and R. B. Wickner. 2009. "Prion variants and species barriers among Saccharomyces Ure2 proteins." *Genetics* 181 (3):1159-67. doi: 10.1534/genetics.108.099929.

Eisenberg, D., H. S. Gill, G. M. Pfluegl, and S. H. Rotstein. 2000. "Structure-function relationships of glutamine synthetases." *Biochim Biophys Acta* 1477 (1-2):122-45.

Eisenberg, H., and E. Reisler. 1971. "Angular dependence of scattered light, rotary frictional coefficients, and distribution of sizes of associated oligomers in solutions of bovine liver glutamate dehydrogenase." *Biopolymers* 10 (12):2363-76. doi: 10.1002/bip.360101202.

Eising, R, M Heinze, S Kleff, and KB Tenberge. 1998. "Subcellular distribution and photooxidation of catalase in sunflower." *Antioxidants in Higher Plants: Biosynthesis, characteristics, actions and specific functions in stress defense*:53-63.

Endo, S., N. Hiramatsu, K. Hayakawa, M. Okamura, A. Kasai, Y. Tagawa, N. Sawada, J. Yao, and M. Kitamura. 2007. "Geranylgeranylacetone, an inducer of the 70-kDa heat shock protein (HSP70), elicits unfolded protein response and coordinates cellular fate independently of HSP70." *Mol Pharmacol* 72 (5):1337-48. doi: 10.1124/mol.107.039164.

Engelbrecht, A. H., and K. Esau. 1963. "Occurrence of Inclusions of Beet Yellows Viruses in Chloroplasts." *Virology* 21:43-7.

Eswar, N., B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper, and A. Sali. 2007. "Comparative protein structure modeling using MODELLER." *Curr Protoc Protein Sci* Chapter 2:Unit 2 9. doi: 10.1002/0471140864.ps0209s50.

Ewing, R. M., P. Chu, F. Elisma, H. Li, P. Taylor, S. Climie, L. McBroom-Cerajewski, M. D. Robinson, L. O'Connor, M. Li, R. Taylor, M. Dharsee, Y. Ho, A. Heilbut, L. Moore, S. Zhang, O. Ornatsky, Y. V. Bukhman, M. Ethier, Y. Sheng, J. Vasilescu, M. Abu-Farha, J. P. Lambert, H. S. Duewel, Stewart, II, B. Kuehl, K. Hogue, K. Colwill, K. Gladwish, B. Muskat, R. Kinach, S. L. Adams, M. F. Moran, G. B. Morin, T. Topaloglou, and D. Figeys. 2007. "Large-scale mapping of human protein-protein interactions by mass spectrometry." *Mol Syst Biol* 3:89. doi: 10.1038/msb4100134.

Fan, C., S. Cheng, Y. Liu, C. M. Escobar, C. S. Crowley, R. E. Jefferson, T. O. Yeates, and T. A. Bobik. 2010. "Short N-terminal sequences package proteins into bacterial microcompartments." *Proc Natl Acad Sci U S A* 107 (16):7509-14. doi: 10.1073/pnas.0913199107.

Fernandez-Escamilla, A. M., F. Rousseau, J. Schymkowitz, and L. Serrano. 2004. "Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins." *Nat Biotechnol* 22 (10):1302-6. doi: 10.1038/nbt1012.

Fiebitz, A., and D. Vanhecke. 2011. "High-throughput mammalian two-hybrid screening for protein-protein interactions using transfected cell arrays (CAPPIA)." *Methods Mol Biol* 723:165-83. doi: 10.1007/978-1-61779-043-0_11.

Field, M. S., D. D. Anderson, and P. J. Stover. 2011. "Mthfs is an Essential Gene in Mice and a Component of the Purinosome." *Front Genet* 2:36. doi: 10.3389/fgene.2011.00036.

Freeman, B. C., D. O. Toft, and R. I. Morimoto. 1996. "Molecular chaperone machines: chaperone activities of the cyclophilin Cyp-40 and the steroid aporeceptor-associated protein p23." *Science* 274 (5293):1718-20.

French, J. B., H. Zhao, S. An, S. Niessen, Y. Deng, B. F. Cravatt, and S. J. Benkovic. 2013. "Hsp70/Hsp90 chaperone machinery is involved in the assembly of the purinosome." *Proc Natl Acad Sci U S A* 110 (7):2528-33. doi: 10.1073/pnas.1300173110.

Frey, T. G., D. Eisenberg, and F. A. Eiserling. 1975. "Glutamine synthetase forms three- and seven-stranded helical cables." *Proc Natl Acad Sci U S A* 72 (9):3402-6.

Frieden, C. 1958. "The dissociation of glutamic dehydrogenase by reduced diphosphopyridine nucleotide (DPNH)." *Biochim Biophys Acta* 27 (2):431-2.

Fuge, E. K., E. L. Braun, and M. Werner-Washburne. 1994. "Protein synthesis in long-term stationary-phase cultures of Saccharomyces cerevisiae." *J Bacteriol* 176 (18):5802-13.

Gallo, C. M., E. Munro, D. Rasoloson, C. Merritt, and G. Seydoux. 2008. "Processing bodies and germ granules are distinct RNA granules that interact in C. elegans embryos." *Dev Biol* 323 (1):76-87. doi: 10.1016/j.ydbio.2008.07.008.

Garcia-Mata, R., Z. Bebok, E. J. Sorscher, and E. S. Sztul. 1999. "Characterization and dynamics of aggresome formation by a cytosolic GFP-chimera." *J Cell Biol* 146 (6):1239-54.

Garcia-Mata, R., Y. S. Gao, and E. Sztul. 2002. "Hassles with taking out the garbage: aggravating aggresomes." *Traffic* 3 (6):388-96.

Garcia-Sastre, A., and C. A. Biron. 2006. "Type 1 interferons and the virus-host relationship: a lesson in detente." *Science* 312 (5775):879-82. doi: 10.1126/science.1125676.

Glory, E., and R. F. Murphy. 2007. "Automated subcellular location determination and high-throughput microscopy." *Dev Cell* 12 (1):7-16. doi: 10.1016/j.devcel.2006.12.007.

Goldschmidt, L., P. K. Teng, R. Riek, and D. Eisenberg. 2010. "Identifying the amylome, proteins capable of forming amyloid-like fibrils." *Proc Natl Acad Sci U S A* 107 (8):3487-92. doi: 10.1073/pnas.0915166107.

Gooljarsingh, L. T., J. Ramcharan, S. Gilroy, and S. J. Benkovic. 2001. "Localization of GAR transformylase in Escherichia coli and mammalian cells." *Proc Natl Acad Sci U S A* 98 (12):6565-70. doi: 10.1073/pnas.121182998.

Goswami, A., P. Dikshit, A. Mishra, S. Mulherkar, N. Nukina, and N. R. Jana. 2006. "Oxidative stress promotes mutant huntingtin aggregation and mutant huntingtin-dependent cell death by mimicking proteasomal malfunction." *Biochem Biophys Res Commun* 342 (1):184-90. doi: 10.1016/j.bbrc.2006.01.136.

Gou, K. M., C. C. Chang, Q. J. Shen, L. Y. Sung, and J. L. Liu. 2014. "CTP synthase forms cytoophidia in the cytoplasm and nucleus." *Exp Cell Res* 323 (1):242-53. doi: 10.1016/j.yexcr.2014.01.029.

Goure, William F., Grant A. Krafft, Jasna Jerecic, and Franz Hefti. 2014. "Targeting the proper amyloid-beta neuronal toxins: a path forward for Alzheimer's disease immunotherapeutics." *Alzheimer's Research & Therapy* 6 (4):1-15. doi: 10.1186/alzrt272.

Gray, J. V., G. A. Petsko, G. C. Johnston, D. Ringe, R. A. Singer, and M. Werner-Washburne. 2004. ""Sleeping beauty": quiescence in Saccharomyces cerevisiae." *Microbiol Mol Biol Rev* 68 (2):187-206. doi: 10.1128/MMBR.68.2.187-206.2004.

Gregolin, C., E. Ryder, A. K. Kleinschmidt, R. C. Warner, and M. D. Lane. 1966. "Molecular characteristics of liver acetyl CoA carboxylase." *Proc Natl Acad Sci U S A* 56 (1):148-55.

Gunning, B. E. 1965. "The Fine Structure of Chloroplast Stroma Following Aldehyde Osmium-Tetroxide Fixation." *J Cell Biol* 24:79-93.

Halfmann, R., D. F. Jarosz, S. K. Jones, A. Chang, A. K. Lancaster, and S. Lindquist. 2012. "Prions are a common mechanism for phenotypic inheritance in wild yeasts." *Nature* 482 (7385):363-8. doi: 10.1038/nature10875.

Harding, R. L., S. Howley, L. J. Baker, T. R. Murphy, W. E. Archer, G. Wistow, D. R. Hyde, and T. S. Vihtelic. 2008. "Lengsin expression and function during zebrafish lens formation." *Exp Eye Res* 86 (5):807-18. doi: 10.1016/j.exer.2008.02.009.

Harrington, D. J., K. Adachi, and W. E. Royer, Jr. 1997. "The high resolution crystal structure of deoxyhemoglobin S." *J Mol Biol* 272 (3):398-407. doi: 10.1006/jmbi.1997.1253.

Hart, Traver, Alice Zhao, Ankit Garg, Swetha Bolusani, and Edward M. Marcotte. 2009. "Human Cell Chips: Adapting DNA Microarray Spotting Technology to Cell-Based Imaging Assays." *PLoS ONE* 4 (10):e7088. doi: 10.1371/journal.pone.0007088.

Hashimoto, M., L. J. Hsu, Y. Xia, A. Takeda, A. Sisk, M. Sundsmo, and E. Masliah. 1999. "Oxidative stress induces amyloid-like aggregate formation of NACP/alpha-synuclein in vitro." *Neuroreport* 10 (4):717-21.

Havemann, G. D., and T. A. Bobik. 2003. "Protein content of polyhedral organelles involved in coenzyme B12-dependent degradation of 1,2-propanediol in Salmonella enterica serovar Typhimurium LT2." *J Bacteriol* 185 (17):5086-95.

Havemann, G. D., E. M. Sampson, and T. A. Bobik. 2002. "PduA is a shell protein of polyhedral organelles involved in coenzyme B(12)-dependent degradation of 1,2-propanediol in Salmonella enterica serovar typhimurium LT2." *J Bacteriol* 184 (5):1253-61.

Havugimana, P. C., G. T. Hart, T. Nepusz, H. Yang, A. L. Turinsky, Z. Li, P. I. Wang, D. R. Boutz, V. Fong, S. Phanse, M. Babu, S. A. Craig, P. Hu, C. Wan, J. Vlasblom, V. U. Dar, A. Bezginov, G. W. Clark, G. C. Wu, S. J. Wodak, E. R. Tillier, A. Paccanaro, E. M. Marcotte, and A. Emili. 2012. "A census of human soluble protein complexes." *Cell* 150 (5):1068-81. doi: 10.1016/j.cell.2012.08.011.

Hayashi, A., D. Q. Ding, C. Tsutsumi, Y. Chikashige, H. Masuda, T. Haraguchi, and Y. Hiraoka. 2009. "Localization of gene products using a chromosomally tagged GFP-

fusion library in the fission yeast Schizosaccharomyces pombe." *Genes Cells* 14 (2):217-25. doi: 10.1111/j.1365-2443.2008.01264.x.

Hayes, J. M., S. K. Kim, P. B. Abitua, T. J. Park, E. R. Herrington, A. Kitayama, M. W. Grow, N. Ueno, and J. B. Wallingford. 2007. "Identification of novel ciliogenesis factors using a new in vivo model for mucociliary epithelial development." *Dev Biol* 312 (1):115-30. doi: 10.1016/j.ydbio.2007.09.031.

He, Y. X., L. Gui, Y. Z. Liu, Y. Du, Y. Zhou, P. Li, and C. Z. Zhou. 2009. "Crystal structure of Saccharomyces cerevisiae glutamine synthetase Gln1 suggests a nanotube-like supramolecular assembly." *Proteins* 76 (1):249-54. doi: 10.1002/prot.22403.

Heese, K., M. Fujita, H. Akatsu, T. Yamamoto, K. Kosaka, Y. Nagai, and T. Sawada. 2004. "The splicing regulatory protein p18SRP is down-regulated in Alzheimer's disease brain." *J Mol Neurosci* 24 (2):269-76. doi: 10.1385/JMN:24:2:269.

Hein, M. Y., N. C. Hubner, I. Poser, J. Cox, N. Nagaraj, Y. Toyoda, I. A. Gak, I. Weisswange, J. Mansfeld, F. Buchholz, A. A. Hyman, and M. Mann. 2015. "A human interactome in three quantitative dimensions organized by stoichiometries and abundances." *Cell* 163 (3):712-23. doi: 10.1016/j.cell.2015.09.053.

Heinze, M., R. Reichelt, S. Kleff, and R. Eising. 2000. "High Resolution Scanning Electron Microscopy of Protein Inclusions (Cores) purified from Peroxisomes of Sunflower (Helianthus annuus L.) Cotyledons." *Crystal Research and Technology* 35 (6-7):877-886. doi: 10.1002/1521-4079(200007)35:6/7<877::AID-CRAT877>3.0.CO;2-S.

Henikoff, S., M. A. Keene, J. S. Sloan, J. Bleskan, R. Hards, and D. Patterson. 1986. "Multiple purine pathway enzyme activities are encoded at a single genetic locus in Drosophila." *Proc Natl Acad Sci U S A* 83 (3):720-4.

Hierro, A., A. L. Rojas, R. Rojas, N. Murthy, G. Effantin, A. V. Kajava, A. C. Steven, J. S. Bonifacino, and J. H. Hurley. 2007. "Functional architecture of the retromer cargo-recognition complex." *Nature* 449 (7165):1063-7. doi: 10.1038/nature06216.

Hirakura, Y., and B. L. Kagan. 2001. "Pore formation by beta-2-microglobulin: a mechanism for the pathogenesis of dialysis associated amyloidosis." *Amyloid* 8 (2):94-100.

Holm, L., and P. Rosenstrom. 2010. "Dali server: conservation mapping in 3D." *Nucleic Acids Res* 38 (Web Server issue):W545-9. doi: 10.1093/nar/gkq366.

Hong, Z., Y. Yang, C. Zhang, Y. Niu, K. Li, X. Zhao, and J. J. Liu. 2009. "The retromer component SNX6 interacts with dynactin p150(Glued) and mediates endosome-to-TGN transport." *Cell Res* 19 (12):1334-49. doi: 10.1038/cr.2009.130.

Howard, R. A., P. Sharma, C. Hajjar, K. A. Caldwell, G. A. Caldwell, R. du Breuil, R. Moore, and L. Boyd. 2007. "Ubiquitin conjugating enzymes participate in polyglutamine protein aggregation." *BMC Cell Biol* 8:32. doi: 10.1186/1471-2121-8-32.

HS, Gill, and Eisenberg D. 2001. "The crystal structure of phosphinothricin in the active site of glutamine synthetase illuminates the mechanism of enzymatic inhibition." *Biochemistry* 40:1903.

Huh, W. K., J. V. Falvo, L. C. Gerke, A. S. Carroll, R. W. Howson, J. S. Weissman, and E. K. O'Shea. 2003. "Global analysis of protein localization in budding yeast." *Nature* 425 (6959):686-91. doi: 10.1038/nature02026.

Huttlin, E. L., L. Ting, R. J. Bruckner, F. Gebreab, M. P. Gygi, J. Szpyt, S. Tam, G. Zarraga, G. Colby, K. Baltier, R. Dong, V. Guarani, L. P. Vaites, A. Ordureau, R. Rad, B. K. Erickson, M. Wuhr, J. Chick, B. Zhai, D. Kolippakkam, J. Mintseris, R. A. Obar, T. Harris, S. Artavanis-Tsakonas, M. E. Sowa, P. De Camilli, J. A. Paulo, J. W. Harper, and S. P. Gygi. 2015. "The BioPlex Network: A Systematic Exploration of the Human Interactome." *Cell* 162 (2):425-40. doi: 10.1016/j.cell.2015.06.043.

Iancu, C. V., H. J. Ding, D. M. Morris, D. P. Dias, A. D. Gonzales, A. Martino, and G. J. Jensen. 2007. "The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography." *J Mol Biol* 372 (3):764-73. doi: 10.1016/j.jmb.2007.06.059.

Imataka, H., H. S. Olsen, and N. Sonenberg. 1997. "A new translational regulator with homology to eukaryotic translation initiation factor 4G." *EMBO J* 16 (4):817-25. doi: 10.1093/emboj/16.4.817.

Ingerson-Mahar, M., A. Briegel, J. N. Werner, G. J. Jensen, and Z. Gitai. 2010. "The metabolic enzyme CTP synthase forms cytoskeletal filaments." *Nat Cell Biol* 12 (8):739-46. doi: 10.1038/ncb2087.

Jana, N. R., M. Tanaka, Gh Wang, and N. Nukina. 2000. "Polyglutamine length-dependent interaction of Hsp40 and Hsp70 family chaperones with truncated N-terminal huntingtin: their role in suppression of aggregation and cellular toxicity." *Hum Mol Genet* 9 (13):2009-18.

Jedd, G., and N. H. Chua. 2000. "A new self-assembled peroxisomal vesicle required for efficient resealing of the plasma membrane." *Nat Cell Biol* 2 (4):226-31. doi: 10.1038/35008652.

Johnston, J. A., C. L. Ward, and R. R. Kopito. 1998. "Aggresomes: a cellular response to misfolded proteins." *J Cell Biol* 143 (7):1883-98.

Johnston, M., M. C. Geoffroy, A. Sobala, R. Hay, and G. Hutvagner. 2010. "HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells." *Mol Biol Cell* 21 (9):1462-9. doi: 10.1091/mbc.E09-10-0885.

Jones, Mary, Richard Fosbery, Jennifer Gregory, and Dennis Taylor. 2014. *Cambridge International AS and A level biology. Coursebook*. Fourth edition. ed. Cambridge: Cambridge University Press.

Josephs, R., and G. Borisy. 1972. "Self-assembly of glutamic dehydrogenase into ordered superstructures: multichain tubes formed by association of single molecules." *J Mol Biol* 65 (1):127-55.

Juda, P., J. Smigova, L. Kovacik, E. Bartova, and I. Raska. 2014. "Ultrastructure of cytoplasmic and nuclear inosine-5'-monophosphate dehydrogenase 2 "rods and

rings" inclusions." *J Histochem Cytochem* 62 (10):739-50. doi: 10.1369/0022155414543853.

Kadowaki, M., R. Venerando, G. Miotto, and G. E. Mortimore. 1994. "De novo autophagic vacuole formation in hepatocytes permeabilized by Staphylococcus aureus alpha-toxin. Inhibition by nonhydrolyzable GTP analogs." *J Biol Chem* 269 (5):3703-10.

Kaganovich, D., R. Kopito, and J. Frydman. 2008. "Misfolded proteins partition between two distinct quality control compartments." *Nature* 454 (7208):1088-95. doi: 10.1038/nature07195.

Kalia, L. V., S. K. Kalia, H. Chau, A. M. Lozano, B. T. Hyman, and P. J. McLean. 2011. "Ubiquitinylation of alpha-synuclein by carboxyl terminus Hsp70-interacting protein (CHIP) is regulated by Bcl-2-associated athanogene 5 (BAG5)." *PLoS One* 6 (2):e14695. doi: 10.1371/journal.pone.0014695.

Kalia, S. K., S. Lee, P. D. Smith, L. Liu, S. J. Crocker, T. E. Thorarinsdottir, J. R. Glover, E. A. Fon, D. S. Park, and A. M. Lozano. 2004. "BAG5 inhibits parkin and enhances dopaminergic neuron degeneration." *Neuron* 44 (6):931-45. doi: 10.1016/j.neuron.2004.11.026.

Kaminosono, S., T. Saito, F. Oyama, T. Ohshima, A. Asada, Y. Nagai, N. Nukina, and S. Hisanaga. 2008. "Suppression of mutant Huntingtin aggregate formation by Cdk5/p35 through the effect on microtubule stability." *J Neurosci* 28 (35):8747-55. doi: 10.1523/JNEUROSCI.0973-08.2008.

Kato, M., T. W. Han, S. Xie, K. Shi, X. Du, L. C. Wu, H. Mirzaei, E. J. Goldsmith, J. Longgood, J. Pei, N. V. Grishin, D. E. Frantz, J. W. Schneider, S. Chen, L. Li, M. R. Sawaya, D. Eisenberg, R. Tycko, and S. L. McKnight. 2012. "Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels." *Cell* 149 (4):753-67. doi: 10.1016/j.cell.2012.04.017.

Kaul, D. K., M. E. Fabry, P. Windisch, S. Baez, and R. L. Nagel. 1983. "Erythrocytes in sickle cell anemia are heterogeneous in their rheological and hemodynamic characteristics." *J Clin Invest* 72 (1):22-31.

Kedersha, N., and P. Anderson. 2007. "Mammalian stress granules and processing bodies." *Methods Enzymol* 431:61-81. doi: 10.1016/S0076-6879(07)31005-7.

Kedersha, N., G. Stoecklin, M. Ayodele, P. Yacono, J. Lykke-Andersen, M. J. Fritzler, D. Scheuner, R. J. Kaufman, D. E. Golan, and P. Anderson. 2005. "Stress granules and processing bodies are dynamically linked sites of mRNP remodeling." *J Cell Biol* 169 (6):871-84. doi: 10.1083/jcb.200502088.

Kedersha, N., S. Tisdale, T. Hickman, and P. Anderson. 2008. "Real-time and quantitative imaging of mammalian stress granules and processing bodies." *Methods Enzymol* 448:521-52. doi: 10.1016/S0076-6879(08)02626-8.

Kerfeld, C. A., M. R. Sawaya, S. Tanaka, C. V. Nguyen, M. Phillips, M. Beeby, and T. O. Yeates. 2005. "Protein structures forming the shell of primitive bacterial organelles." *Science* 309 (5736):936-8. doi: 10.1126/science.1113397.

Kim, C. W., Y. A. Moon, S. W. Park, D. Cheng, H. J. Kwon, and J. D. Horton. 2010. "Induced polymerization of mammalian acetyl-CoA carboxylase by MIG12

provides a tertiary level of regulation of fatty acid synthesis." *Proc Natl Acad Sci U S A* 107 (21):9626-31. doi: 10.1073/pnas.1001292107.

Kim, S., E. A. Nollen, K. Kitagawa, V. P. Bindokas, and R. I. Morimoto. 2002. "Polyglutamine protein aggregates are dynamic." *Nat Cell Biol* 4 (10):826-31. doi: 10.1038/ncb863.

Kim, S. Y., Y. W. Kim, R. Hegerl, M. Cyrklaff, and I. S. Kim. 2005. "Novel type of enzyme multimerization enhances substrate affinity of oat beta-glucosidase." *J Struct Biol* 150 (1):1-10. doi: 10.1016/j.jsb.2004.07.007.

Klein, J., and P. Dhurjati. 1995. "Protein aggregation kinetics in an Escherichia coli strain overexpressing a Salmonella typhimurium CheY mutant gene." *Appl Environ Microbiol* 61 (4):1220-5.

Kourie, J. I., A. L. Culverson, P. V. Farrelly, C. L. Henry, and K. N. Laohachai. 2002. "Heterogeneous amyloid-formed ion channels as a common cytotoxic mechanism: implications for therapeutic strategies against amyloidosis." *Cell Biochem Biophys* 36 (2-3):191-207. doi: 10.1385/CBB:36:2-3:191.

Krasowska, Joanna, Monika Olasek, Agnieszka Bzowska, Patricia L. Clark, and Beata Wielgus-Kutrowska. 2010. "The comparison of aggregation and folding of enhanced green fluorescent protein (EGFP) by spectroscopic studies." *Spectroscopy* 24 (3-4). doi: 10.3233/spe-2010-0445.

Krobitsch, S., and S. Lindquist. 2000. "Aggregation of huntingtin in yeast varies with the length of the polyglutamine expansion and the expression of chaperone proteins." *Proc Natl Acad Sci U S A* 97 (4):1589-94.

Krutzik, P. O., J. M. Crane, M. R. Clutter, and G. P. Nolan. 2008. "High-content single-cell drug screening with phosphospecific flow cytometry." *Nat Chem Biol* 4 (2):132-42. doi: 10.1038/nchembio.2007.59.

Krutzik, P. O., J. M. Irish, G. P. Nolan, and O. D. Perez. 2004. "Analysis of protein phosphorylation and cellular signaling events by flow cytometry: techniques and clinical applications." *Clin Immunol* 110 (3):206-21. doi: 10.1016/j.clim.2003.11.009.

Kuusisto, E., A. Salminen, and I. Alafuzoff. 2001. "Ubiquitin-binding protein p62 is present in neuronal and glial inclusions in human tauopathies and synucleinopathies." *Neuroreport* 12 (10):2085-90.

Kuusisto, E., A. Salminen, and I. Alafuzoff. 2002. "Early accumulation of p62 in neurofibrillary tangles in Alzheimer's disease: possible role in tangle formation." *Neuropathol Appl Neurobiol* 28 (3):228-37.

Labesse, G., T. Alexandre, L. Vaupre, I. Salard-Arnaud, J. L. Him, B. Raynal, P. Bron, and H. Munier-Lehmann. 2013. "MgATP regulates allostery and fiber formation in IMPDHs." *Structure* 21 (6):975-85. doi: 10.1016/j.str.2013.03.011.

Lam, C., E. Santore, E. Lavoie, L. Needleman, N. Fiacco, C. Kim, and A. M. Neiman. 2014. "A visual screen of protein localization during sporulation identifies new components of prospore membrane-associated complexes in budding yeast." *Eukaryot Cell* 13 (3):383-91. doi: 10.1128/EC.00333-13.

Landgraf, D., B. Okumus, P. Chien, T. A. Baker, and J. Paulsson. 2012. "Segregation of molecules at cell division reveals native protein localization." *Nat Methods* 9 (5):480-2. doi: 10.1038/nmeth.1955.

Legname, G., H. O. Nguyen, D. Peretz, F. E. Cohen, S. J. DeArmond, and S. B. Prusiner. 2006. "Continuum of prion protein structures enciphers a multitude of prion isolate-specified phenotypes." *Proc Natl Acad Sci U S A* 103 (50):19105-10. doi: 10.1073/pnas.0608970103.

Lehninger, Albert L., David L. Nelson, and Michael M. Cox. 2005. *Lehninger principles of biochemistry*. 4th ed. 1 vols. New York: W.H. Freeman.

Levitzki, A., W. B. Stallcup, and D. E. Koshland, Jr. 1971. "Half-of-the-sites reactivity and the conformational states of cytidine triphosphate synthetase." *Biochemistry* 10 (18):3371-8.

Li, B., and C. N. Dewey. 2011. "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome." *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323.

Litchfield, D. W. 2003. "Protein kinase CK2: structure, regulation and role in cellular decisions of life and death." *Biochem J* 369 (Pt 1):1-15. doi: 10.1042/BJ20021469.

Liu, J. L. 2010. "Intracellular compartmentation of CTP synthase in Drosophila." *J Genet Genomics* 37 (5):281-96. doi: 10.1016/S1673-8527(09)60046-1.

MacKeigan, J. P., L. O. Murphy, and J. Blenis. 2005. "Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance." *Nat Cell Biol* 7 (6):591-600. doi: 10.1038/ncb1258.

Marcotte, E. M., M. Pellegrini, H. L. Ng, D. W. Rice, T. O. Yeates, and D. Eisenberg. 1999. "Detecting protein function and protein-protein interactions from genome sequences." *Science* 285 (5428):751-3.

Margolin, W. 2012. "The price of tags in protein localization studies." *J Bacteriol* 194 (23):6369-71. doi: 10.1128/JB.01640-12.

Matsumoto, K., M. Minami, F. Shinozaki, Y. Suzuki, K. Abe, S. Zenno, S. Matsumoto, and Y. Minami. 2011. "Hsp90 is involved in the formation of P-bodies and stress granules." *Biochem Biophys Res Commun* 407 (4):720-4. doi: 10.1016/j.bbrc.2011.03.088.

Matsuyama, A., R. Arai, Y. Yashiroda, A. Shirai, A. Kamata, S. Sekido, Y. Kobayashi, A. Hashimoto, M. Hamamoto, Y. Hiraoka, S. Horinouchi, and M. Yoshida. 2006. "ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe." *Nat Biotechnol* 24 (7):841-7. doi: 10.1038/nbt1222.

Mayer, M., and J. Buchner. 2004. "Refolding of inclusion body proteins." *Methods Mol Med* 94:239-54.

Mazumder, A., L. Q. Pesudo, S. McRee, M. Bathe, and L. D. Samson. 2013. "Genome-wide single-cell-level screen for protein abundance and localization changes in response to DNA damage in S. cerevisiae." *Nucleic Acids Res* 41 (20):9310-24. doi: 10.1093/nar/gkt715.

McCairns, E., D. Fahey, D. Sauer, and P. B. Rowe. 1983. "De novo purine synthesis in human lymphocytes. Partial co-purification of the enzymes and some properties of the pathway." *J Biol Chem* 258 (3):1851-6.

McLean, P. J., H. Kawamata, and B. T. Hyman. 2001. "Alpha-synuclein-enhanced green fluorescent protein fusion proteins form proteasome sensitive inclusions in primary neurons." *Neuroscience* 104 (3):901-12.

McLean, P. J., H. Kawamata, S. Shariff, J. Hewett, N. Sharma, K. Ueda, X. O. Breakefield, and B. T. Hyman. 2002. "TorsinA and heat shock proteins act as molecular chaperones: suppression of alpha-synuclein aggregation." *J Neurochem* 83 (4):846-54.

McLean, P. J., J. Klucken, Y. Shin, and B. T. Hyman. 2004. "Geldanamycin induces Hsp70 and prevents alpha-synuclein aggregation and toxicity in vitro." *Biochem Biophys Res Commun* 321 (3):665-9. doi: 10.1016/j.bbrc.2004.07.021.

Meggio, F., and L. A. Pinna. 2003. "One-thousand-and-one substrates of protein kinase CK2?" *FASEB J* 17 (3):349-68. doi: 10.1096/fj.02-0473rev.

Meredith, M. J., and M. D. Lane. 1978. "Acetyl-CoA carboxylase. Evidence for polymeric filament to protomer transition in the intact avian liver cell." *J Biol Chem* 253 (10):3381-3.

Meyer, E., T. J. Kappock, C. Osuji, and J. Stubbe. 1999. "Evidence for the direct transfer of the carboxylate of N5-carboxyaminoimidazole ribonucleotide (N5-CAIR) to generate 4-carboxy-5-aminoimidazole ribonucleotide catalyzed by Escherichia coli PurE, an N5-CAIR mutase." *Biochemistry* 38 (10):3012-8. doi: 10.1021/bi9827159.

Michnick, S. W. 2004. "Proteomics in living cells." *Drug Discov Today* 9 (6):262-7. doi: 10.1016/S1359-6446(03)03022-8.

Mignot, C., C. Delarasse, S. Escaich, B. Della Gaspera, E. Noe, E. Colucci-Guyon, C. Babinet, M. Pekny, P. Vicart, O. Boespflug-Tanguy, A. Dautigny, D. Rodriguez, and D. Pham-Dinh. 2007. "Dynamics of mutated GFAP aggregates revealed by real-time imaging of an astrocyte model of Alexander disease." *Exp Cell Res* 313 (13):2766-79. doi: 10.1016/j.yexcr.2007.04.035.

Miller, R. E., E. Shelton, and E. R. Stadtman. 1974. "Zinc-induced paracrystalline aggregation of glutamine synthetase." *Arch Biochem Biophys* 163 (1):155-71.

Muchowski, P. J., K. Ning, C. D'Souza-Schorey, and S. Fields. 2002. "Requirement of an intact microtubule cytoskeleton for aggregation and inclusion body formation by a mutant huntingtin fragment." *Proc Natl Acad Sci U S A* 99 (2):727-32. doi: 10.1073/pnas.022628699.

Muchowski, P. J., and J. L. Wacker. 2005. "Modulation of neurodegeneration by molecular chaperones." *Nat Rev Neurosci* 6 (1):11-22. doi: 10.1038/nrn1587.

Murphy, R. F. 2008. "Automated Proteome-Wide Determination of Subcellular Location Using High Throughput Microscopy." *Proc IEEE Int Symp Biomed Imaging* 2008:308-311. doi: 10.1109/ISBI.2008.4540994.

Nagaoka, U., K. Kim, N. R. Jana, H. Doi, M. Maruyama, K. Mitsui, F. Oyama, and N. Nukina. 2004. "Increased expression of p62 in expanded polyglutamine-expressing cells and its association with polyglutamine inclusions." *J Neurochem* 91 (1):57-68. doi: 10.1111/j.1471-4159.2004.02692.x.

Narayanaswamy, R., M. Levy, M. Tsechansky, G. M. Stovall, J. D. O'Connell, J. Mirrielees, A. D. Ellington, and E. M. Marcotte. 2009. "Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation." *Proc Natl Acad Sci U S A* 106 (25):10147-52. doi: 10.1073/pnas.0812771106.

Narayanaswamy, R., E. K. Moradi, W. Niu, G. T. Hart, M. Davis, K. L. McGary, A. D. Ellington, and E. M. Marcotte. 2009. "Systematic definition of protein constituents along the major polarization axis reveals an adaptive reuse of the polarization machinery in pheromone-treated budding yeast." *J Proteome Res* 8 (1):6-19. doi: 10.1021/pr800524g.

Narayanaswamy, R., W. Niu, A. D. Scouras, G. T. Hart, J. Davies, A. D. Ellington, V. R. Iyer, and E. M. Marcotte. 2006. "Systematic profiling of cellular phenotypes with spotted cell microarrays reveals mating-pheromone response genes." *Genome Biol* 7 (1):R6. doi: 10.1186/gb-2006-7-1-r6.

Nelson, D. E., A. E. Ihekwaba, M. Elliott, J. R. Johnson, C. A. Gibney, B. E. Foreman, G. Nelson, V. See, C. A. Horton, D. G. Spiller, S. W. Edwards, H. P. McDowell, J. F. Unitt, E. Sullivan, R. Grimley, N. Benson, D. Broomhead, D. B. Kell, and M. R. White. 2004. "Oscillations in NF-kappaB signaling control the dynamics of gene expression." *Science* 306 (5696):704-8. doi: 10.1126/science.1099962.

Neumann, B., M. Held, U. Liebel, H. Erfle, P. Rogers, R. Pepperkok, and J. Ellenberg. 2006. "High-throughput RNAi screening by time-lapse imaging of live human cells." *Nat Methods* 3 (5):385-90. doi: 10.1038/nmeth876.

Noble, S. L., B. L. Allen, L. K. Goh, K. Nordick, and T. C. Evans. 2008. "Maternal mRNAs are regulated by diverse P body-related mRNP granules during early Caenorhabditis elegans development." *J Cell Biol* 182 (3):559-72. doi: 10.1083/jcb.200802128.

Noree, C., E. Monfort, A. K. Shiau, and J. E. Wilhelm. 2014. "Common regulatory control of CTP synthase enzyme activity and filament formation." *Mol Biol Cell* 25 (15):2282-90. doi: 10.1091/mbc.E14-04-0912.

Noree, C., B. K. Sato, R. M. Broyer, and J. E. Wilhelm. 2010. "Identification of novel filament-forming proteins in Saccharomyces cerevisiae and Drosophila melanogaster." *J Cell Biol* 190 (4):541-51. doi: 10.1083/jcb.201003001.

O'Connell, J. D., M. Tsechansky, A. Royall, D. R. Boutz, A. D. Ellington, and E. M. Marcotte. 2014. "A proteomic survey of widespread protein aggregation in yeast." *Mol Biosyst* 10 (4):851-61. doi: 10.1039/c3mb70508k.

Ogier-Denis, E., J. J. Houri, C. Bauvy, and P. Codogno. 1996. "Guanine nucleotide exchange on heterotrimeric Gi3 protein controls autophagic sequestration in HT-29 cells." *J Biol Chem* 271 (45):28593-600.

Ohtsuki, M., S. L. White, E. Zeitler, T. E. Wellems, S. D. Fuller, M. Zwick, M. W. Makinen, and P. B. Sigler. 1977. "Electron microscopy of fibers and discs of hemoglobin S having sixfold symmetry." *Proc Natl Acad Sci U S A* 74 (12):5538-42.

Olson, J. A., and C. B. Anfinsen. 1952. "The crystallization and characterization of L-glutamic acid dehydrogenase." *J Biol Chem* 197 (1):67-79.

Oughtred, R., A. Chatr-Aryamontri, B. J. Breitkreutz, C. S. Chang, J. M. Rust, C. L. Theesfeld, S. Heinicke, A. Breitkreutz, D. Chen, J. Hirschman, N. Kolas, M. S. Livstone, J. Nixon, L. O'Donnell, L. Ramage, A. Winter, T. Reguly, A. Sellam, C. Stark, L. Boucher, K. Dolinski, and M. Tyers. 2016. "Use of the BioGRID Database for Analysis of Yeast Protein and Genetic Interactions." *Cold Spring Harb Protoc* 2016 (1):pdb prot088880. doi: 10.1101/pdb.prot088880.

Pagano, M. A., J. Bain, Z. Kazimierczuk, S. Sarno, M. Ruzzene, G. Di Maira, M. Elliott, A. Orzeszko, G. Cozza, F. Meggio, and L. A. Pinna. 2008. "The selectivity of inhibitors of protein kinase CK2: an update." *Biochem J* 415 (3):353-65. doi: 10.1042/BJ20080309.

Palmer, E., and T. Freeman. 2004. "Investigation into the use of C- and N-terminal GFP fusion proteins for subcellular localization studies using reverse transfection microarrays." *Comp Funct Genomics* 5 (4):342-53. doi: 10.1002/cfg.405.

Pan, P., E. Woehl, and M. F. Dunn. 1997. "Protein architecture, dynamics and allostery in tryptophan synthase channeling." *Trends Biochem Sci* 22 (1):22-7.

Paweletz, C. P., L. Charboneau, V. E. Bichsel, N. L. Simone, T. Chen, J. W. Gillespie, M. R. Emmert-Buck, M. J. Roth, Iii Ef Petricoin, and L. A. Liotta. 2001. "Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front." *Oncogene* 20 (16):1981-9. doi: 10.1038/sj.onc.1204265.

Petiot, A., E. Ogier-Denis, C. Bauvy, F. Cluzeaud, A. Vandewalle, and P. Codogno. 1999. "Subcellular localization of the Galphai3 protein and G alpha interacting protein, two proteins involved in the control of macroautophagy in human colon cancer HT-29 cells." *Biochem J* 337 (Pt 2):289-95.

Petrovska, I., E. Nuske, M. C. Munder, G. Kulasegaran, L. Malinovska, S. Kroschwald, D. Richter, K. Fahmy, K. Gibson, J. M. Verbavatz, and S. Alberti. 2014. "Filament formation by metabolic enzymes is a specific adaptation to an advanced state of cellular starvation." *Elife*. doi: 10.7554/eLife.02409.

Piatigorsky, J. 1993. "Puzzle of crystallin diversity in eye lenses." *Dev Dyn* 196 (4):267-72. doi: 10.1002/aja.1001960408.

Polianskyte, Z., N. Peitsaro, A. Dapkunas, J. Liobikas, R. Soliymani, M. Lalowski, O. Speer, J. Seitsonen, S. Butcher, G. M. Cereghetti, M. D. Linder, M. Merckel, J. Thompson, and O. Eriksson. 2009. "LACTB is a filament-forming protein localized in mitochondria." *Proc Natl Acad Sci U S A* 106 (45):18960-5. doi: 10.1073/pnas.0906734106.

Price, W. C., A. P. Martinez, and H. E. Warmke. 1966. "Crystalline inclusions in chloroplasts of the coconut palm." *J Ultrastruct Res* 14 (5):618-21.

Probst, C., C. Radzimski, I. M. Blocker, B. Teegen, D. P. Bogdanos, W. Stocker, and L. Komorowski. 2013. "Development of a recombinant cell-based indirect immunofluorescence assay (RC-IFA) for the determination of autoantibodies against "rings and rods"-associated inosine-5'-monophosphate dehydrogenase 2 in viral hepatitis C." *Clin Chim Acta* 418:91-6. doi: 10.1016/j.cca.2013.01.003.

Proctor, C. J., and I. A. Lorimer. 2011. "Modelling the role of the Hsp70/Hsp90 system in the maintenance of protein homeostasis." *PLoS One* 6 (7):e22038. doi: 10.1371/journal.pone.0022038.

Ramanathan, M. P., V. Ayyavoo, and D. B. Weiner. 2001. "Choice of expression vector alters the localization of a human cellular protein." *DNA Cell Biol* 20 (2):101-5. doi: 10.1089/104454901750070300.

Ren, P. H., J. E. Lauckner, I. Kachirskaia, J. E. Heuser, R. Melki, and R. R. Kopito. 2009. "Cytoplasmic penetration and persistent infection of mammalian cells by polyglutamine aggregates." *Nat Cell Biol* 11 (2):219-25. doi: 10.1038/ncb1830.

Reuther-Madrid, J. Y., D. Kashatus, S. Chen, X. Li, J. Westwick, R. J. Davis, H. S. Earp, C. Y. Wang, and A. S. Baldwin Jr, Jr. 2002. "The p65/RelA subunit of NF-kappaB suppresses the sustained, antiapoptotic activity of Jun kinase induced by tumor necrosis factor." *Mol Cell Biol* 22 (23):8175-83.

Riedel, M., O. Goldbaum, L. Schwarz, S. Schmitt, and C. Richter-Landsberg. 2010. "17-AAG induces cytoplasmic alpha-synuclein aggregate clearance by induction of autophagy." *PLoS One* 5 (1):e8753. doi: 10.1371/journal.pone.0008753.

Riesner, D. 2003. "Biochemistry and structure of PrP(C) and PrP(Sc)." *Br Med Bull* 66:21-33.

Rinas, U., and J. E. Bailey. 1992. "Protein compositional analysis of inclusion bodies produced in recombinant Escherichia coli." *Appl Microbiol Biotechnol* 37 (5):609-14.

Rizzo, M. A., M. W. Davidson, and D. W. Piston. 2009. "Fluorescent protein tracking and detection: fluorescent protein structure and color variants." *Cold Spring Harb Protoc* 2009 (12):pdb top63. doi: 10.1101/pdb.top63.

Robertson, J. G. 1995. "Determination of subunit dissociation constants in native and inactivated CTP synthetase by sedimentation equilibrium." *Biochemistry* 34 (22):7533-41.

Rojas, R., S. Kametaka, C. R. Haft, and J. S. Bonifacino. 2007. "Interchangeable but essential functions of SNX1 and SNX2 in the association of retromer with endosomes and the trafficking of mannose 6-phosphate receptors." *Mol Cell Biol* 27 (3):1112-24. doi: 10.1128/MCB.00156-06.

Ross, C. A., and M. A. Poirier. 2004. "Protein aggregation and neurodegenerative disease." *Nat Med* 10 Suppl:S10-7. doi: 10.1038/nm1066.

Rowe, P. B., E. McCairns, G. Madsen, D. Sauer, and H. Elliott. 1978. "De novo purine synthesis in avian liver. Co-purification of the enzymes and properties of the pathway." *J Biol Chem* 253 (21):7711-21.

Rudolph, J., and J. Stubbe. 1995. "Investigation of the mechanism of phosphoribosylamine transfer from glutamine phosphoribosylpyrophosphate amidotransferase to glycinamide ribonucleotide synthetase." *Biochemistry* 34 (7):2241-50.

Ruepp, A., B. Waegele, M. Lechner, B. Brauner, I. Dunger-Kaltenbach, G. Fobo, G. Frishman, C. Montrone, and H. W. Mewes. 2010. "CORUM: the comprehensive resource of mammalian protein complexes--2009." *Nucleic Acids Res* 38 (Database issue):D497-501. doi: 10.1093/nar/gkp914.

Sabbaghian, M., A. Ebrahim-Habibi, and M. Nemat-Gorgani. 2009. "Thermal aggregation of a model allosteric protein in different conformational states." *Int J Biol Macromol* 44 (2):156-62. doi: 10.1016/j.ijbiomac.2008.11.011.

Sagot, I., B. Pinson, B. Salin, and B. Daignan-Fornier. 2006. "Actin bodies in yeast quiescent cells: an immediately available actin reserve?" *Mol Biol Cell* 17 (11):4645-55. doi: 10.1091/mbc.E06-04-0282.

Sampathu, D. M., B. I. Giasson, A. C. Pawlyk, J. Q. Trojanowski, and V. M. Lee. 2003. "Ubiquitination of alpha-synuclein is not required for formation of pathological inclusions in alpha-synucleinopathies." *Am J Pathol* 163 (1):91-100.

Savage, D. F., B. Afonso, A. H. Chen, and P. A. Silver. 2010. "Spatially ordered dynamics of the bacterial carbon fixation machinery." *Science* 327 (5970):1258-61. doi: 10.1126/science.1186090.

Sawin, K. E., and P. Nurse. 1996. "Identification of fission yeast nuclear markers using random polypeptide fusions with green fluorescent protein." *Proc Natl Acad Sci U S A* 93 (26):15146-51.

Schendel, F. J., Y. S. Cheng, J. D. Otvos, S. Wehrli, and J. Stubbe. 1988. "Characterization and chemical properties of phosphoribosylamine, an unstable intermediate in the de novo purine biosynthetic pathway." *Biochemistry* 27 (7):2614-23.

Scheufler, C., A. Brinker, G. Bourenkov, S. Pegoraro, L. Moroder, H. Bartunik, F. U. Hartl, and I. Moarefi. 2000. "Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine." *Cell* 101 (2):199-210. doi: 10.1016/S0092-8674(00)80830-2.

Schmidt, T., K. S. Lindenberg, A. Krebs, L. Schols, F. Laccone, J. Herms, M. Rechsteiner, O. Riess, and G. B. Landwehrmeyer. 2002. "Protein surveillance machinery in brains with spinocerebellar ataxia type 3: redistribution and differential recruitment of 26S proteasome subunits and chaperones to neuronal intranuclear inclusions." *Ann Neurol* 51 (3):302-10.

Schneider, C. C., A. Hessenauer, C. Gotz, and M. Montenarh. 2009. "DMAT, an inhibitor of protein kinase CK2 induces reactive oxygen species and DNA double strand breaks." *Oncol Rep* 21 (6):1593-7.

Schneider, M. D., N. Najand, S. Chaker, J. M. Pare, J. Haskins, S. C. Hughes, T. C. Hobman, J. Locke, and A. J. Simmonds. 2006. "Gawky is a component of

cytoplasmic mRNA processing bodies required for early Drosophila development." *J Cell Biol* 174 (3):349-58. doi: 10.1083/jcb.200512103.
Seaman, M. N. 2012. "The retromer complex - endosomal protein recycling and beyond." *J Cell Sci* 125 (Pt 20):4693-702. doi: 10.1242/jcs.103440.
Serpell, L. C. 2000. "Alzheimer's amyloid fibrils: structure and assembly." *Biochim Biophys Acta* 1502 (1):16-30.
Sevecka, M., and G. MacBeath. 2006. "State-based discovery: a multidimensional screen for small-molecule modulators of EGF signaling." *Nat Methods* 3 (10):825-31. doi: 10.1038/nmeth931.
Shaner, N. C., G. G. Lambert, A. Chammas, Y. Ni, P. J. Cranfill, M. A. Baird, B. R. Sell, J. R. Allen, R. N. Day, M. Israelsson, M. W. Davidson, and J. Wang. 2013. "A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum." *Nat Methods* 10 (5):407-9. doi: 10.1038/nmeth.2413.
Shen, Qing-Ji, Hakimi Kassim, Yong Huang, Hui Li, Jing Zhang, Guang Li, Peng-Ye Wang, Jun Yan, Fangfu Ye, and Ji-Long Liu. 2016. "Filamentation of Metabolic Enzymes in Saccharomyces cerevisiae." *Journal of Genetics and Genomics*. doi: http://dx.doi.org/10.1016/j.jgg.2016.03.008.
Shinder, G. A., M. C. Lacourse, S. Minotti, and H. D. Durham. 2001. "Mutant Cu/Zn-superoxide dismutase proteins have altered solubility and interact with heat shock/stress proteins in models of amyotrophic lateral sclerosis." *J Biol Chem* 276 (16):12791-6. doi: 10.1074/jbc.M010759200.
Shively, J. M. 1974. "Inclusion bodies of prokaryotes." *Annu Rev Microbiol* 28 (0):167-87. doi: 10.1146/annurev.mi.28.100174.001123.
Shively, J. M., F. Ball, D. H. Brown, and R. E. Saunders. 1973. "Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of Thiobacillus neapolitanus." *Science* 182 (4112):584-6.
Skube, S. B., J. M. Chaverri, and H. V. Goodson. 2010. "Effect of GFP tags on the localization of EB1 and EB1 fragments in vivo." *Cytoskeleton (Hoboken)* 67 (1):1-12. doi: 10.1002/cm.20409.
Slaton, J. W., G. M. Unger, D. T. Sloper, A. T. Davis, and K. Ahmed. 2004. "Induction of apoptosis by antisense CK2 in human prostate cancer xenograft model." *Mol Cancer Res* 2 (12):712-21.
Smith, G. K., W. T. Mueller, P. A. Benkovic, and S. J. Benkovic. 1981. "On the cofactor specificity of glycinamide ribonucleotide and 5-aminoimidazole-4-carboxamide ribonucleotide transformylase from chicken liver." *Biochemistry* 20 (5):1241-5.
Smith, G. K., W. T. Mueller, G. F. Wasserman, W. D. Taylor, and S. J. Benkovic. 1980. "Characterization of the enzyme complex involving the folate-requiring enzymes of de novo purine biosynthesis." *Biochemistry* 19 (18):4313-21.
Snapp, E. 2005. "Design and use of fluorescent fusion proteins in cell biology." *Curr Protoc Cell Biol* Chapter 21:Unit 21 4. doi: 10.1002/0471143030.cb2104s27.
Snapp, E. L. 2009. "Fluorescent proteins: a cell biologist's user guide." *Trends Cell Biol* 19 (11):649-55. doi: 10.1016/j.tcb.2009.08.002.

So, A. K., G. S. Espie, E. B. Williams, J. M. Shively, S. Heinhorst, and G. C. Cannon. 2004. "A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell." *J Bacteriol* 186 (3):623-30.

Soen, Y., A. Mori, T. D. Palmer, and P. O. Brown. 2006. "Exploring the regulation of human neural precursor cell differentiation using arrays of signaling microenvironments." *Mol Syst Biol* 2:37. doi: 10.1038/msb4100076.

Squirrell, J. M., Z. T. Eggers, N. Luedke, B. Saari, A. Grimson, G. E. Lyons, P. Anderson, and J. G. White. 2006. "CAR-1, a protein that localizes with the mRNA decapping component DCAP-1, is required for cytokinesis and ER organization in Caenorhabditis elegans embryos." *Mol Biol Cell* 17 (1):336-44. doi: 10.1091/mbc.E05-09-0874.

Stadler, C., E. Rexhepaj, V. R. Singan, R. F. Murphy, R. Pepperkok, M. Uhlen, J. C. Simpson, and E. Lundberg. 2013. "Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells." *Nat Methods* 10 (4):315-23. doi: 10.1038/nmeth.2377.

Starkuviene, V., and R. Pepperkok. 2007. "The potential of high-content high-throughput microscopy in drug discovery." *Br J Pharmacol* 152 (1):62-71. doi: 10.1038/sj.bjp.0707346.

Stefani, M., and C. M. Dobson. 2003. "Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution." *J Mol Med (Berl)* 81 (11):678-99. doi: 10.1007/s00109-003-0464-5.

Stillman, T. J., P. J. Baker, K. L. Britton, and D. W. Rice. 1993. "Conformational flexibility in glutamate dehydrogenase. Role of water in substrate recognition and catalysis." *J Mol Biol* 234 (4):1131-9. doi: 10.1006/jmbi.1993.1665.

Stojiljkovic, I., A. J. Baumler, and F. Heffron. 1995. "Ethanolamine utilization in Salmonella typhimurium: nucleotide sequence, protein expression, and mutational analysis of the cchA cchB eutE eutJ eutG eutH gene cluster." *J Bacteriol* 177 (5):1357-66.

Sue, M., K. Yamazaki, S. Yajima, T. Nomura, T. Matsukawa, H. Iwamura, and T. Miyamoto. 2006. "Molecular and structural characterization of hexameric beta-D-glucosidases in wheat and rye." *Plant Physiol* 141 (4):1237-47. doi: 10.1104/pp.106.077693.

Suresh, H. G., A. X. da Silveira Dos Santos, W. Kukulski, J. Tyedmers, H. Riezman, B. Bukau, and A. Mogk. 2015. "Prolonged starvation drives reversible sequestration of lipid biosynthetic enzymes and organelle reorganization in Saccharomyces cerevisiae." *Mol Biol Cell* 26 (9):1601-15. doi: 10.1091/mbc.E14-11-1559.

Suzuki, Y., M. Minami, M. Suzuki, K. Abe, S. Zenno, M. Tsujimoto, K. Matsumoto, and Y. Minami. 2009. "The Hsp90 inhibitor geldanamycin abrogates colocalization of eIF4E and eIF4E-transporter into stress granules and association of eIF4E with eIF4G." *J Biol Chem* 284 (51):35597-604. doi: 10.1074/jbc.M109.036285.

Swartz, J. R. 2001. "Advances in Escherichia coli production of therapeutic proteins." *Curr Opin Biotechnol* 12 (2):195-201.

Swulius, M. T., and G. J. Jensen. 2012. "The helical MreB cytoskeleton in Escherichia coli MC1000/pLE7 is an artifact of the N-Terminal yellow fluorescent protein tag." *J Bacteriol* 194 (23):6382-6. doi: 10.1128/JB.00505-12.

Szklarczyk, D., A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen, and C. von Mering. 2015. "STRING v10: protein-protein interaction networks, integrated over the tree of life." *Nucleic Acids Res* 43 (Database issue):D447-52. doi: 10.1093/nar/gku1003.

Tartaglia, G. G., S. Pechmann, C. M. Dobson, and M. Vendruscolo. 2007. "Life on the edge: a link between gene expression levels and aggregation rates of human proteins." *Trends Biochem Sci* 32 (5):204-6. doi: 10.1016/j.tibs.2007.03.005.

Teixeira, D., U. Sheth, M. A. Valencia-Sanchez, M. Brengues, and R. Parker. 2005. "Processing bodies require RNA for assembly and contain nontranslating mRNAs." *RNA* 11 (4):371-82. doi: 10.1261/rna.7258505.

Tenberge, K. B., and R. Eising. 1995. "Immunogold labelling indicates high catalase concentrations in amorphous and crystalline inclusions of sunflower (Helianthus annuus L.) peroxisomes." *Histochem J* 27 (3):184-95.

Thomas, Elaine C, Jennifer H Gunter, Julie A Webster, Nicole L Schieber, Viola Oorschot, Robert G Parton, and Jonathan P Whitehead. 2012. "Different characteristics and nucleotide binding properties of inosine monophosphate dehydrogenase (IMPDH) isoforms." *PloS one* 7 (12):e51096.

Tkach, J. M., A. Yimit, A. Y. Lee, M. Riffle, M. Costanzo, D. Jaschob, J. A. Hendry, J. Ou, J. Moffat, C. Boone, T. N. Davis, C. Nislow, and G. W. Brown. 2012. "Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress." *Nat Cell Biol* 14 (9):966-76. doi: 10.1038/ncb2549.

Trembley, J. H., G. Wang, G. Unger, J. Slaton, and K. Ahmed. 2009. "Protein kinase CK2 in health and disease: CK2: a key player in cancer biology." *Cell Mol Life Sci* 66 (11-12):1858-67. doi: 10.1007/s00018-009-9154-y.

Tyedmers, J., A. Mogk, and B. Bukau. 2010. "Cellular strategies for controlling protein aggregation." *Nat Rev Mol Cell Biol* 11 (11):777-88. doi: 10.1038/nrm2993.

Uhlen, M., E. Bjorling, C. Agaton, C. A. Szigyarto, B. Amini, E. Andersen, A. C. Andersson, P. Angelidou, A. Asplund, C. Asplund, L. Berglund, K. Bergstrom, H. Brumer, D. Cerjan, M. Ekstrom, A. Elobeid, C. Eriksson, L. Fagerberg, R. Falk, J. Fall, M. Forsberg, M. G. Bjorklund, K. Gumbel, A. Halimi, I. Hallin, C. Hamsten, M. Hansson, M. Hedhammar, G. Hercules, C. Kampf, K. Larsson, M. Lindskog, W. Lodewyckx, J. Lund, J. Lundeberg, K. Magnusson, E. Malm, P. Nilsson, J. Odling, P. Oksvold, I. Olsson, E. Oster, J. Ottosson, L. Paavilainen, A. Persson, R. Rimini, J. Rockberg, M. Runeson, A. Sivertsson, A. Skollermo, J. Steen, M. Stenvall, F. Sterky, S. Stromberg, M. Sundberg, H. Tegel, S. Tourle, E. Wahlund, A. Walden, J. Wan, H. Wernerus, J. Westberg, K. Wester, U. Wrethagen, L. L. Xu, S. Hober, and F. Ponten. 2005. "A human protein atlas for normal and cancer tissues

based on antibody proteomics." *Mol Cell Proteomics* 4 (12):1920-32. doi: 10.1074/mcp.M500279-MCP200.

Uryu, K., C. Richter-Landsberg, W. Welch, E. Sun, O. Goldbaum, E. H. Norris, C. T. Pham, I. Yazawa, K. Hilburger, M. Micsenyi, B. I. Giasson, N. M. Bonini, V. M. Lee, and J. Q. Trojanowski. 2006. "Convergence of heat shock protein 90 with ubiquitin in filamentous alpha-synuclein inclusions of alpha-synucleinopathies." *Am J Pathol* 168 (3):947-61.

Valentine, R. C., B. M. Shapiro, and E. R. Stadtman. 1968. "Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from Escherichia coli." *Biochemistry* 7 (6):2143-2152. doi: 10.1021/bi00846a017.

Verrier, F., S. An, A. M. Ferrie, H. Sun, M. Kyoung, H. Deng, Y. Fang, and S. J. Benkovic. 2011. "GPCRs regulate the assembly of a multienzyme complex for purine biosynthesis." *Nat Chem Biol* 7 (12):909-15. doi: 10.1038/nchembio.690.

Vilar, M., H. T. Chou, T. Luhrs, S. K. Maji, D. Riek-Loher, R. Verel, G. Manning, H. Stahlberg, and R. Riek. 2008. "The fold of alpha-synuclein fibrils." *Proc Natl Acad Sci U S A* 105 (25):8637-42. doi: 10.1073/pnas.0712179105.

von Mering, C., R. Krause, B. Snel, M. Cornell, S. G. Oliver, S. Fields, and P. Bork. 2002. "Comparative assessment of large-scale data sets of protein-protein interactions." *Nature* 417 (6887):399-403. doi: 10.1038/nature750.

Wainscoat, J. S., J. I. Bell, S. L. Thein, D. R. Higgs, G. R. Sarjeant, T. E. Peto, and D. J. Weatherall. 1983. "Multiple origins of the sickle mutation: evidence from beta S globin gene cluster polymorphisms." *Mol Biol Med* 1 (2):191-7.

Wan, C., B. Borgeson, S. Phanse, F. Tu, K. Drew, G. Clark, X. Xiong, O. Kagan, J. Kwan, A. Bezginov, K. Chessman, S. Pal, G. Cromar, O. Papoulas, Z. Ni, D. R. Boutz, S. Stoilova, P. C. Havugimana, X. Guo, R. H. Malty, M. Sarov, J. Greenblatt, M. Babu, W. B. Derry, E. R. Tillier, J. B. Wallingford, J. Parkinson, E. M. Marcotte, and A. Emili. 2015. "Panorama of ancient metazoan macromolecular complexes." *Nature* 525 (7569):339-44. doi: 10.1038/nature14877.

Wang, K., C. Cheng, L. Li, H. Liu, Q. Huang, C. H. Xia, K. Yao, P. Sun, J. Horwitz, and X. Gong. 2007. "GammaD-crystallin associated protein aggregation and lens fiber cell denucleation." *Invest Ophthalmol Vis Sci* 48 (8):3719-28. doi: 10.1167/iovs.06-1487.

Wassmer, T., N. Attar, M. V. Bujny, J. Oakley, C. J. Traer, and P. J. Cullen. 2007. "A loss-of-function screen reveals SNX5 and SNX6 as potential components of the mammalian retromer." *J Cell Sci* 120 (Pt 1):45-54. doi: 10.1242/jcs.03302.

Watabe, M., and T. Nakaki. 2011. "Protein kinase CK2 regulates the formation and clearance of aggresomes in response to stress." *J Cell Sci* 124 (Pt 9):1519-32. doi: 10.1242/jcs.081778.

Watabe, M., and T. Nakaki. 2012. "CK2 as anti-stress factor." *Commun Integr Biol* 5 (3):278-80. doi: 10.4161/cib.19473.

Werner, J. N., E. Y. Chen, J. M. Guberman, A. R. Zippilli, J. J. Irgon, and Z. Gitai. 2009. "Quantitative genome-scale analysis of protein localization in an asymmetric

bacterium." *Proc Natl Acad Sci U S A* 106 (19):7858-63. doi: 10.1073/pnas.0901781106.

Wheeler, D. B., A. E. Carpenter, and D. M. Sabatini. 2005. "Cell microarrays and RNA interference chip away at gene function." *Nat Genet* 37 Suppl:S25-30. doi: 10.1038/ng1560.

Wickner, R. B. 1994. "[URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae." *Science* 264 (5158):566-9.

Woolfolk, C. A., and E. R. Stadtman. 1967. "Regulation of glutamine synthetase. 3. Cumulative feedback inhibition of glutamine synthetase from Escherichia coli." *Arch Biochem Biophys* 118 (3):736-55.

Wu, R. Z., S. N. Bailey, and D. M. Sabatini. 2002. "Cell-biological applications of transfected-cell microarrays." *Trends Cell Biol* 12 (10):485-8.

Xu, Y., J. Cao, S. Huang, D. Feng, W. Zhang, X. Zhu, and X. Yan. 2015. "Characterization of tetratricopeptide repeat-containing proteins critical for cilia formation and function." *PLoS One* 10 (4):e0124378. doi: 10.1371/journal.pone.0124378.

Yamaoka, T., M. Kondo, S. Honda, H. Iwahana, M. Moritani, S. Ii, K. Yoshimoto, and M. Itakura. 1997. "Amidophosphoribosyltransferase limits the rate of cell growth-linked de novo purine biosynthesis in the presence of constant capacity of salvage purine biosynthesis." *J Biol Chem* 272 (28):17719-25.

Yamaoka, T., M. Yano, M. Kondo, H. Sasaki, S. Hino, R. Katashima, M. Moritani, and M. Itakura. 2001. "Feedback inhibition of amidophosphoribosyltransferase regulates the rate of cell growth via purine nucleotide, DNA, and protein syntheses." *J Biol Chem* 276 (24):21285-91. doi: 10.1074/jbc.M011103200.

Yeates, T. O., C. A. Kerfeld, S. Heinhorst, G. C. Cannon, and J. M. Shively. 2008. "Protein-based organelles in bacteria: carboxysomes and related microcompartments." *Nat Rev Microbiol* 6 (9):681-91. doi: 10.1038/nrmicro1913.

Yeates, T. O., Y. Tsai, S. Tanaka, M. R. Sawaya, and C. A. Kerfeld. 2007. "Self-assembly in the carboxysome: a viral capsid-like protein shell in bacterial cells." *Biochem Soc Trans* 35 (Pt 3):508-11. doi: 10.1042/BST0350508.

Yokota, S., K. Kamijo, and T. Oda. 2000. "Aggregate formation and degradation of overexpressed wild-type and mutant urate oxidase proteins. Quality control of organelle-destined proteins by the endoplasmic reticulum." *Histochem Cell Biol* 114 (6):433-46.

Yuan, P., G. Jedd, D. Kumaran, S. Swaminathan, H. Shio, D. Hewitt, N. H. Chua, and K. Swaminathan. 2003. "A HEX-1 crystal lattice required for Woronin body function in Neurospora crassa." *Nat Struct Biol* 10 (4):264-70. doi: 10.1038/nsb910.

Zacharias, D. A., J. D. Violin, A. C. Newton, and R. Y. Tsien. 2002. "Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells." *Science* 296 (5569):913-6. doi: 10.1126/science.1068539.

Zatloukal, K., C. Stumptner, A. Fuchsbichler, H. Heid, M. Schnoelzer, L. Kenner, R. Kleinert, M. Prinz, A. Aguzzi, and H. Denk. 2002. "p62 Is a common component

of cytoplasmic inclusions in protein aggregation diseases." *Am J Pathol* 160 (1):255-63. doi: 10.1016/S0002-9440(10)64369-6.

Zeiri, L., and E. Reisler. 1978. "Uncoupling of the catalytic activity and the polymerization of beef liver glutamate dehydrogenase." *J Mol Biol* 124 (1):291-5.

Zeytuni, N., and R. Zarivach. 2012. "Structural and functional discussion of the tetra-trico-peptide repeat, a protein interaction module." *Structure* 20 (3):397-405. doi: 10.1016/j.str.2012.01.006.

Zhang, J., L. Hulme, and J. L. Liu. 2014. "Asymmetric inheritance of cytoophidia in Schizosaccharomyces pombe." *Biol Open* 3 (11):1092-7. doi: 10.1242/bio.20149613.

Zhang, Y. B., J. Howitt, S. McCorkle, P. Lawrence, K. Springer, and P. Freimuth. 2004. "Protein aggregation during overexpression limited by peptide extensions with large net negative charge." *Protein Expr Purif* 36 (2):207-16. doi: 10.1016/j.pep.2004.04.020.

Zhao, A., M. Tsechansky, J. Swaminathan, L. Cook, A. D. Ellington, and E. M. Marcotte. 2013. "Transiently transfected purine biosynthetic enzymes form stress bodies." *PLoS One* 8 (2):e56203. doi: 10.1371/journal.pone.0056203.

Zhao, H., C. R. Chiaro, L. Zhang, P. B. Smith, C. Y. Chan, A. M. Pedley, R. J. Pugh, J. B. French, A. D. Patterson, and S. J. Benkovic. 2015. "Quantitative analysis of purine nucleotides indicates that purinosomes increase de novo purine biosynthesis." *J Biol Chem* 290 (11):6705-13. doi: 10.1074/jbc.M114.628701.

Zhao, J., W. Niu, J. Yao, S. Mohr, E. M. Marcotte, and A. M. Lambowitz. 2008. "Group II intron protein localization and insertion sites are affected by polyphosphate." *PLoS Biol* 6 (6):e150. doi: 10.1371/journal.pbio.0060150.

Zheng, D., C. Y. Chen, and A. B. Shyu. 2011. "Unraveling regulation and new components of human P-bodies through a protein interaction framework and experimental validation." *RNA* 17 (9):1619-34. doi: 10.1261/rna.2789611.

Ziauddin, J., and D. M. Sabatini. 2001. "Microarrays of cells expressing defined cDNAs." *Nature* 411 (6833):107-10. doi: 10.1038/35075114.