# Abstract


## Applying Protein-Peptide Interactions to Fluorescently Label Proteins in *S. Cerevisiae*

Michael Hinrichsen

2018

Cells are highly complex, organized and dynamic machines, driven by intricate protein-based machinery. Understanding how this protein machinery produces the cellular processes we observe requires more than simply characterizing the individual parts in isolation. Complete understanding requires knowledge of a protein's spatial distribution in the cell, how this distribution relates to that of other proteins in the cell, and how these distributions change over time. Fluorescent labeling of proteins is an unrivaled technique for visualizing protein localizations and dynamics in living cells and can provide great insight into many of the mechanisms that underlie cellular processes.

While a number of methods can be used to fluorescently label a protein of interest, there is no one method that is suitable for all imaging applications. Oftentimes, picking the appropriate labeling method requires balancing a variety of competing factors such as label size, fluorophore properties, and labeling specificity. Label size in particular can be a cause for concern, as many labeling methods require modifications

that are nearly as big as the protein itself and can affect the protein's native localization and/or function. I sought to develop new methods for fluorescently labeling proteins *in vivo* that do not require such large modifications to the protein of interest, but that can still be readily applied to fluorescently label any protein of interest. Rather than make direct fusions to a fluorescent protein, our approach is to use protein/peptide interactions to mediate protein labeling. Proteins of interest are tagged at the C-terminus with a short peptide tag, and the corresponding binding protein that recognizes the peptide tag is fused to a fluorescent protein and expressed in the same cell. Interaction between the binding protein and the peptide drives association of the fluorescent protein to the protein of interest, fluorescently labeling the target protein. I used two different protein/peptide interaction pairs to mediate protein labeling; the irreversible SpyCatcher/SpyTag pair, and the reversible Tetratricopeptide Repeat Affinity Protein (TRAP) binding domain.

SpyCatcher/SpyTag are a protein/peptide pair that bind one another and form an irreversible isopeptide bond in solution. I harnessed this interaction pair to fluorescently label proteins *in vivo* by fusing SpyTag (13 amino acids) to the protein of interest, and separately expressing SpyoIPD, a more stable derivative of SpyCatcher, fused to a fluorescent protein (FP). I have developed the molecular biology tools and labeling protocol required to fluorescently label proteins in *S. Cerevisiae*, and successfully used our strategy to fluorescently label a range of proteins in live yeast. I have also shown that labeling proteins in this manner can be less disruptive to protein function than direct fusion to a fluorescent protein, and that pulsed labeling of SpyTagged proteins with SpyoIPD-FP can be used to track the turnover rate of a plasma membrane protein in living cells. I have also investigated the potential of different SpyoIPD-FP sequestration

strategies to reduce fluorescent signal from unbound SpyoIPD-EGFP, demonstrated that SpyCatcher can be used in combination with a variety of fluorescent proteins, and that SnoopCatcher/SnoopTag, another permanent protein/peptide interaction pair, can also be used to fluorescently label proteins *in vivo*.

Additionally, I have also investigated the use of a reversible protein/peptide interaction to fluorescently label proteins in *S. cerevisiae*. TRAPs are artificial binding domains developed in the Regan lab that bind 5 amino acid long, C-terminal peptide tags with affinities as low as 300 nM. Previous work in the Regan lab has demonstrated that TRAP domains can be used to fluorescently label a protein in *E. Coli*. Here I extend upon this work, showing that TRAPs can also be used to fluorescently label a collection of proteins in *S. cerevisiae*.

Thus, I have shown that a variety of protein/peptide interactions can be used to fluorescently label proteins *in vivo* and have developed the tools and protocols necessary to fluorescently label proteins of interest in yeast. Our SpyCatcher/SpyTag based strategy in particular offers a useful tool for fluorescently labeling target proteins that are sensitive to large modifications, or for experiments that require studying protein dynamics over time.

# Applying Protein-Peptide Interactions to Fluorescently Label Proteins in *S. Cerevisiae*

A Dissertation

Presented to the Faculty of the Graduate School

Of

Yale University

In Candidacy for the Degree of

Doctor of Philosophy

By

Michael Hinrichsen

Dissertation Director: Lynne Regan, Ph. D.

May 2018

# Table of Contents

Abstract……...……...1-3

Table of Contents……..……..…….. 6-7

List of Figures and Tables……...…….…….8-9

List of commonly used abbreviations…….……10

Acknowledgements……..…….11-12

Chapter 1 - Introduction

Chapter 1.1 - Transduction of Chemically Labeled Proteins…….…….13-16

Chapter 1.2 - Intrinsically Fluorescent Proteins…….…….……16-23

Chapter 1.3 - Self-Labeling Enzymes……...……..24-27

Chapter 1.4 - Enzyme-Mediated Labeling ……..…27-29

Chapter 1.5 - Unnatural Amino Acids ……..…… 29-31

Chapter 1.6 - Protein-Protein Interactions ……...……..31-36

Chapter 1.7 - Aims of this Work ……...…….……...…….36-37

Chapter 2 - Using SpyoIPD/SpyTag to Visualize Proteins in Living S. Cerevisiae.

Chapter 2.1 - Introduction……..….38

Chapter 2.2 - Testing SpyCatcher and SpyCatcher Variants In Vivo……...…....38-39

Chapter 2.3 - Design and Characterization of SpyoIPD, a More Stable SpyCatcher Derivative……..39-41

Chapter 2.4 - Testing the Activity of SpyoIPD(IA) and SpyoIPD(IVML) Variants In Vivo……41

Chapter 2.5 - Design of a SpyoIPD/SpyTag Based Imaging System……...….41-42

Chapter 2.6 - Imaging Proteins of Interest in Live S. Cerevisiae with SpyoIPD/SpyTag……...42-47

Chapter 2.7 - Characterizing the Effects of Different Labeling Strategies on Pma1 Function…….…….47

Chapter 2.8 - Improving Signal to Noise when Labeling Pma1-SpyTag with SpyoIPD-EGFP……...……..47-51

Chapter 2.9 - Spatiotemporally Tracking Pma1 in Living Cells……...……51-53

Chapter 2.10 – Discussion……...……..53-55

Chapter 3: Improving SpyoIPD/SpyTag labeling by Nuclear Sequestration of Unreacted SpyoIPD-EGFP

Chapter 3.1 – Introduction……...……...……...56-57

Chapter 3.2 - Selecting an Inducible Promoter……..……..…….57-59

Chapter 3.3 - Creation and Characterization a Nuclear Sequestering Protein……..59-61

Chapter 3.4 - Testing SpyoIPD-EGFP reactivity with NLS-mCherry-SpyTag……61-63

Chapter 3.5 - Using HXT1s to induce NLS-mCherry-SpyTag expression…...…....…..63-65

Chapter 3.6 - Fusing NLS directly to SpyCatcher-EGFP…….65-67

Chapter 3.7 - Discussion……...68-69

Chapter 4: Testing Additional Covalent Protein-Peptide Interaction Pairs

Chapter 4.1 - Imaging Pma1-SpyTag with SpyCatcher-EGFP……..……...……70-72

Chapter 4.2 - Testing SpyCatcher Compatibility with Additional FP's……...……72-74

Chapter 4.3 - Using SnoopCatcher/SnoopTag to image Candidate Proteins in S. Cerevisiae……..74-77

Chapter 4.4 - Discussion ……...…….……77

Chapter 5: TRAP-FP Imaging

Chapter 5.1 - Introduction…….…...……...……78

Chapter 5.2 - Design of a TRAP-FP system for imaging proteins in *S. cerevisiae*..……78-79

Chapter 5.3 - Imaging proteins in Yeast……...…….....……79

Chapter 5.4 - Discussion…….……...……79-81

Chapter 6: Methods

Chapter 6.1 - General Protocols…….……...……82-83

Chapter 6.2 - Methods from Chapter 2……....…….83-90

Chapter 6.3 - Methods from Chapter 3……..90-92

Chapter 6.4 - Methods from Chapter 4……...……92

Chapter 6.5 - Methods from Chapter 5…….…….92-93

Appendix……...…….…...……94-108

References……..……..…….109-119

# List of Tables and Figures

**Figure 1.1**: Cartoon depiction cameleon, the FRET-based calcium sensor created in Miyawaki et al.[1]……...…..….19

**Figure 1.2**: Cartoon illustrating the use of the split-GFP reassembly assay to screen for protein-protein interactions……...…….21

**Figure 1.3**: Comparison of self-labeling enzymes and enzyme-mediated labeling……..24

**Figure 1.4**: Use of unnatural amino acids to fluorescently label proteins…….29

**Figure 1.5**: Labeling proteins of interest (POI) using protein-protein interactions……..32

**Figure 1.6**: Use of split-GFP to visualize target proteins *in vivo*……..34

**Figure 2.1:** Design and characterization of SpyoIPDs……..39

**Figure 2.2:** Schematic illustration of the SpyoIPD/SpyTag labeling strategy……..42

**Figure 2.3**: A comparison of proteins labeled by direct fusion to EGFP with those labeled using SpyoIPD/SpyTag labeling……...…….45

**Figure 2.4**: Effects of different labeling methods on Pma1 localization and function….47

**Figure 2.5**: Strategies for reducing cytoplasmic fluorescence when labeling Pma1-SpyTag with SpyoIPD-EGFP……..…….….49

**Figure 2.6**: Using SpyoIPD-EGFP to study Pma1 temporal dynamics in single cells….51

**Figure 3.1**: Testing the glucose responsiveness of the HXT1s, HXT1L, and HXT3 promoters in yeast……...……...57

**Figure 3.2**: Testing the effects of SpyTag (ST) and Nuclear Localization Sequence (NLS) on protein localization……...…….59

**Figure 3.3**: Testing NLS-mCherry-SpyTag nuclear sequestration of SpyoIPD-EGFP…61

**Figure 3.4**: Testing NLS-mCherry-ST driven nuclear sequestration of SpyoIPD-EGFP.63

**Figure 3.5**: Testing the effect of NLS-mCherry-ST induction on SpyoIPD/SpyTag labeling of Pma1-SpyTag…….…….65

**Figure 3.6**: Testing the effect on target labeling of directly fusing NLS to SpyCatcher-GFP……....66

**Figure 4.1**: Imaging Pma1-SpyTag with SpyCatcher-EGFP……70

**Figure 4.2**: Glucose pulse-chase of SpyCatcher-EGFP vs SpyoIPD-EGFP…….72

**Figure 4.3**: Testing SpyoIPD compatibility with additional fluorescent proteins……74

**Figure 4.4**: Testing SnoopCatcher-EGFP labeling of target proteins in *S. cerevisiae*…..75

**Figure 5.1**: Imaging Target Proteins with TRAP-mCherry in *S. cerevisiae*…….79

**Figure A1**. $^{1}H$-$^{15}N$ HSQC NMR spectra of SpyCatcher, SpyoIPD(IA) and SpyoIPD(IVML)……93

**Figure A2:** Layout of a single chamber of a microfluidic device and cell detection……...…..94

**Figure A3:** SpyoIPD-EGFP pulse-chase imaging of Vma1-SpyTag.……95

**Figure A3**: Tracking the ratio of membrane to cytosol signal over time for cells expressing SpyoIPD-GFP and either PMA1-SpyTag or untagged PMA1……96

**Figure A4:** Investigating the effect of photobleaching in PMA1 turnover experiment…97

# Commonly Used Abbreviations

FP – Fluorescent Protein

POI – Protein of Interest

GFP – Green Fluorescent Protein

EGFP – Enhanced Green Fluorescent Protein

Kd – Dissociation Constant

FRET – Forster Resonance Energy Transfer

CFP – Cyan Fluorescent Protein

UAA – Unnatural Amino Acid

TRAP – Tetratricopeptide Repeat Affinity Protein

HXT – Hexose Transfporter

HXT1s – Hexose Transporter 1 (short promoter)

HXT1L – Hexose Transporter 1 (long promoter)

## Acknowledgements

First and foremost, I would like to thank my thesis adviser, Lynne Regan, both for giving me the material resources I needed to perform research and, more importantly, for guiding me into becoming the scientist I am today. I've particularly appreciated the freedom Lynne has always given me in my research projects, which has been critical in me becoming a competent and independent scientist.

I'd also like to thank my committee members Tom Pollard and Mark Hochstrasser. Tom's advice was critical in keeping me on track in my research and also ensuring that I finished my degree in a timely manner. Mark has also provided crucial scientific advice, particularly with regards to working with yeast, a field that does not fall within the realm of Regan lab expertise and yet constituted the bulk of my thesis work.

I'd also like to thank all the other scientists I've encountered during the course of my research. In particular, I would like to thank members of the Hochstrasser lab for all the much-needed yeast help, members of the Pollard lab for the microscope training and trusting me with their (very expensive) instrument, and the many other scientists who have helped in countless ways for no other reason than to help a fellow scientist out.

I'd also like to thank my classmates and anyone else I've encountered on a day to day basis while at Yale. Interacting with a diverse set of super smart, passionate people has been my favorite part of graduate school and definitely the part I'll miss most.

I would also need to thank my lab mates. You guys kept grad school fun even when the science wasn't going so great. In particular I would like to thank Nick for the GOT conversations and showing me it's possible to be both hard working and easy

going. Betsy for showing me what grit looks like. Ashley for the many laughs and teaching me in the art of the hustle. Curran for the many science/sports/ Sci-Fi conversations and being a role model scientist despite being two years below me. And my bay-buddy and classmate Danielle, for the many off-topic conversations and keeping me company through all the highs and lows of grad school.

Lastly, I'd like to thank my brothers Erik and Niall, and most importantly my parents, Bill and Nancy. The older I get, the more I realize how fortunate I've been to have a stable, loving, and supportive family. Who I am today is in large part a reflection of my family, which means a big chunk of this thesis is owed to you guys.

# Chapter 1 - Introduction

A key requirement for gaining a complete understanding of a protein's function is knowledge of its native localization(s). Cells are highly organized and dynamic environments, so while protein visualization in fixed cells can be helpful, tracking in living cells is essential to fully understand cellular processes. Fluorescently labeling a protein of interest (POI) allows protein localization but is no easy task. The average protein is only several nanometers in diameter, surrounded by a dense solution containing millions of other biomolecules, and separated from bulk solution by a plasma membrane and sometimes cell wall that is not permeable to all labeling reagents. A variety of solutions have been developed to overcome these technical hurdles, the most common of which are discussed in detail below.

## Chapter 1.1 - Transduction of Chemically Labeled Proteins

Labeling a POI specifically within the intracellular environment is difficult. Before the advent of the *in vivo* labeling strategies discussed below, protein labeling could only be performed *in vitro*, by chemically reacting small molecule fluorophores with exposed chemical groups on the surface of the POI. To study protein dynamics in living cells, scientists purified a POI, chemically labeled it with a fluorophore, and then transduced the labelled protein into a cell, typically by either microinjection or electroporation. Much of the initial *in vivo* characterization of the cytoskeleton was performed in this manner, using POIs purified from animal muscle, labeled on accessible amine groups with a rhodamine derivative, and microinjected into mammalian cells[2-4].

Tracking protein dynamics using chemically labeled POIs presents several advantages that still make it a useful tool for some applications. Many measurements benefit from using small molecule fluorophores, which can be brighter, more photostable, and fluoresce at longer wavelengths than fluorescent proteins (FPs). Excitation at longer wavelengths reduces the autofluroescence generated from cellular components and generates fewer cytotoxic effects. The modification introduced during the labeling process is also very small, and therefore unlikely to interfere with native protein function. Fluorophore labeling is achieved through reaction with either surface-exposed primary amine groups (present on lysine side-chains and a protein's N-terminus) or thiol groups (present on cysteine side-chains). The typical protein contains many surface exposed lysine side-chains, making it impossible to label target proteins at one designated position. While the N-terminal amine has a lower pKa than the ε-amino group of lysine, and can theoretically be selectively labeled by careful control of the pH of the labeling reaction, in practice doing so is difficult and success varies depending on both the fluorophore of choice and the POI[5]. Thiol groups are much less common than amine groups, and therefore more amenable to site-specific labeling. For recombinantly expressed proteins, unwanted Cys residues can often be mutated to Ser with minimal effect on protein function, allowing proteins to be labeled with a single fluorophore at a user-designated site. Protein transduction is typically achieved either through electroporation or microinjection. Microinjection, while less high-throughput than electroporation, allows for more precise control over the amount of protein delivered.

Transduction of chemically labeled proteins continues to be a commonly used strategy to study the dynamics of cytoskeletal proteins. Aroush et al. recently used

chemically labeled proteins introduced into fish epithelial keratocytes by electroporation to gain unprecedented insight into the *in vivo* actin dynamics that drive lamellipodial motility. By transducing fluorescently labeled actin into cells with and without co-transduction of capping protein, Aroush et al. were able to selectively label and quantify the growth rates of the barbed and pointed ends of actin filaments. Additionally, by introducing fluorescently labeled phalloidin, they were able to distinguish between oligomeric and monomeric actin (phalloidin binds only actin oligomers)[6]. From these experiments, it was discovered that a large fraction of freely diffusing actin is oligomeric, indicating that actin filament severing and debranching are important steps in the disassembly of actin networks. Moreover, the measured concentration of unincorporated actin was much greater than that required to support the observed actin polymerization rates, suggesting that a large fraction of free actin is held in a non-polymerizable state.

Chemical labeling followed by microinjection was used to collect single-molecule FRET measurements of ProTα, an intrinsically disordered protein, in mammalian cells[7]. In addition to the availability of fluorescent dyes with superior fluorescent properties, attaching small molecule fluorophores using chemical labeling methods also allows fluorophores to be placed with high precision. To fluorescently label ProTα, point mutations were made to introduce two surface-exposed Cys residues. To obtain homogenous samples of protein that contain two different fluorophores, each at a specific position, purified ProTα was first labeled with one fluorophore using subsaturating amounts of fluorescent label. The products of this reaction were then purified by reverse-phase chromatography in order to isolate a singly-labeled permutant, and then this singly-labeled protein was labeled again with the second fluorophore and purified again so as to

remove unreacted fluorophore. This doubly-labeled protein was then microinjected into mammalian cells for single-molecule FRET measurements. From these experiments, the dimensions, temperature-dependence of protein stability, cold denaturation and folding dynamics of ProTα were quantified *in vivo*.

Compared to other methods for fluorescent labeling proteins, transduction of chemically labeled proteins is a labor-intensive process. POIs must be purified, fluorescently labeled, purified again to remove unincorporated fluorescent label, and then transduced into the target cell. Because of these limitations, transduction of chemically labeled proteins is not the method of choice for most experiments that require protein localization. Below are discussed options that are more broadly applicable.

**Chapter 1.2 - Intrinsically Fluorescent Proteins**

The discovery of Green Fluorescent Protein (GFP), the first known intrinsically fluorescent protein, revolutionized the biological sciences because it offered a fast, easy and highly specific method for fluorescently labeling any POI. Since GFP can be genetically encoded and its fluorescence does not depend on any cofactor, any protein in any organism can be fluorescently tagged by making a simple genetic fusion to GFP[8]. Years of protein engineering, selection, and searching for additional FPs from new organisms (most commonly originating from jellyfish and coral, but FP's have also been isolated from bacteria[9] and vertebrates[10]) led to FP markers that are less disruptive to native protein localization, allow multiprotein tracking in the same cell, and capable of sensing and reporting information about the local environment surrounding the FP.

**Improvements in FPs for Protein Localization:**

Much design effort has been directed towards improving the ability of FPs to faithfully and accurately report native protein localizations. The development of brighter FP's, such as enhanced GFP (EGFP), allow proteins to be visualized at lower expression level[11]. At the time of this writing, the current leader in brightness is mNeonGreen, which is approximately 6.5 times brighter than wild-type GFP through improvements in both quantum yield and extinction coefficient[12].

Many FP's are naturally oligomers and can therefore affect the oligomerization state of the protein to which it is fused. Even the weakly dimeric GFP ($K_d$ = 110 μM) can affect oligomeric state if the target protein is present at high local concentrations. The development of truly monomeric FPs was therefore essential in order to provide more accurate reporters of native protein localization. The design of FPs that are monomeric at high concentrations and in the crowded environment of the cell has been difficult. Monomeric versions of FPs can often be designed by examining the crystal structure of the oligomeric protein, and rationally introducing mutations at interaction surfaces so as to disrupt binding interfaces and introduce charge-charge repulsions[13]. Some FPs retain their spectral properties during this process, but for reasons that are not well understood[14, 15], monomer variants often retain the excitation and emission spectrum of the parent protein but are less bright[16].

For proteins that are naturally poorly soluble, attaching an FP tag can increase misfolding and aggregation *in vivo*, making localization difficult or impossible. This limitation led to the development of superfolder GFP (sfGFP), a GFP variant that folds with greater efficiency and increases the solubility and expression of proteins to which it

is fused. Fluorophore maturation is also faster in sfGFP, making it a more accurate reporter of protein translation rates[17].

Through protein engineering and searching for new FPs, a great variety of FP's have been developed that fluoresce at different wavelengths, enabling multiple proteins to be tracked in a single cell[18-20]. More photostable FP variants have also been developed, allowing proteins to be tracked for longer periods of time[21]. The development of photoactivatable FP's (i.e. FP's that switch fluorescent states when illuminated with a particular wavelength of light), also allows for super-resolution imaging[22]. For a comprehensive review of the many fluorescent proteins that have been developed, see Rodriguez et al.[23] and fpvis.org.

**Fluorescent Proteins for Sensing Applications**

Intrinsically fluorescent proteins are the foundation for many fluorescent sensors that function *in vivo*. For sensing applications, one needs to couple binding of an analyte to a change in FP fluorescence, where the change in fluorescence can either be 'intensiometric' or 'ratiometric.' In 'intensiometric sensing,' the intensity of FP fluorescence changes in response to analyte binding (either increasing or decreasing). Signal from intensiometric sensors depends in part on the amount of sensor present, meaning that signal interpretation requires either that the concentration of sensor be known or that it be held constant throughout the experiment. Ratiometric sensors track the change in the ratio of intensities at two wavelengths (either excitation or emission), and require no correction for sensor concentration. Employing pairs of FP in a FRET (Forster resonance energy transfer) setting is a convenient way to achieve ratiometric

sensing, although FRET-based sensing is limited by relatively small changes in signal in response to analyte binding.

**Calcium Sensors**

The first effective synthetic protein calcium sensor was the 'cameleon sensor,' developed by Tsien and colleagues (see Figure 1.1)[1]. The basis of the design was to connect a donor FP (BFP or CFP) to an acceptor FP (GFP or YFP) through a calmodulin protein and 'the M13 peptide,' a calmodulin-binding peptide originating from myosin light chain kinase (MLCK). In the presence of calcium, calmodulin binds calcium, becomes more structured, and binds the M13 peptide. The Ca/calmodulin/M13 complex is more compact than the unbound linker, bringing donor and acceptor fluorophores closer together and increasing FRET. While elegant, this design is limited by poor sensitivity, as the maximum change in FRET from no calcium to saturating amounts of calcium is only 70%. Nevertheless, the sensor allows determination of calcium concentrations in individual cells and in different sub-cellular compartments. Since this initial report, many new synthetic proteins have been created to act as intracellular calcium sensors, often using the same principles as the original cameleon design. Currently, the best performing of these are the so called 'Twitch' sensors, which use a domain of Tropinin C as the calcium-binding unit, and achieve a maximum ratio change of 400%.[24]

**PH Sensors**

GFP-based pH sensors have been developed by taking advantage of the natural pH sensitivity of wild-type GFP. Wild-type GFP has two excitation peaks, at 395 and 475 nm, that correspond to the protonated and deprotonated state of Tyr66, one of the amino acids that forms the chromophore. The ratio between these peaks is relatively

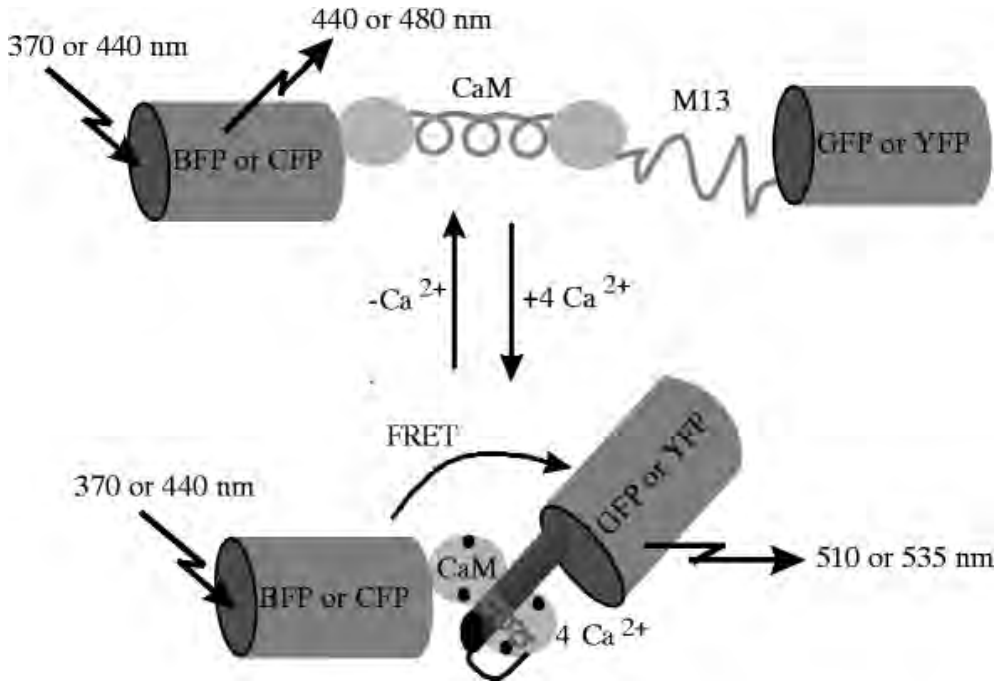



**Figure 1.1**: Cartoon depiction cameleon, the FRET-based calcium sensor created in Miyawaki et al.[1] A donor fluorescent protein (BFP or CFP) is fused to an acceptor fluorescent protein (GFP or YFP) via a calmodulin protein (CaM) and the M13 peptide from the myosin light chain kinase (MLCK). In the absence of calcium, the CaM-M13 linker is unstructured and the fluorophores far apart, resulting in low levels of FRET. When CaM binds calcium, CaM changes conformation and binds the M13 peptide. The Ca/CaM/M13 complex is more compact, bringing the donor and acceptor fluorophores closer, and increasing FRET. Reproduced with permission from Miyawaki et al.[1]

unaffected by pH in wild-type GFP, staying constant between pH 5.5 and 10. By randomly mutating residues near GFP's chromophore, both ratiometric and intensiometric GFP-based pH sensors were created (called 'pHlourins'). In the ratiometric variant, the ratio of 395 nm to 475 nm excitation peaks decreases as pH decreases, while in the intensiometric sensor (called the 'ecliptic sensor') fluorescence from 475 nm excitation is quenched as pH decreases. 'Phlourin2,' a brighter version of the ratiometric pH sensor was later created by adding mutations from EGFP into the pHourin construct[25]. Continued engineering of the intensiometric pH sensor yielded the 'super ecliptic sensor,' which is ~9 times more fluorescent than the original ecliptic sensor[26]. A red intensiometric sensor has also been developed, allowing pH changes to be tracked in multiple subcellular compartments[27]. Intensiometric pH sensors have become a powerful tool for detecting synaptic vesicle fusion events, as the fluorescent sensors are effectively nonfluorescent while in the acidic lumen of vesicles (pH ~ 5.5), and strongly fluorescent in the neutral extracellular environment (ph~7.4). For determining the pH of an intracellular compartment, however, ratiometric sensors are superior, as they require no correction for sensor concentration. For a complete review of pH sensor FP's, see Benčina et al.[28]

**Split fluorescent proteins**

Many FPs can be split into two peptide chains and still fold and form a functional fluorescent protein when the two halves are brought into close proximity (see Figure 1.2)[29]. These split-FPs can be used to sense protein-protein interactions *in vivo* by fusing one half of the split-FP to each potential interaction partner. Because the reaction is essentially irreversible, given enough time even the weakest interacting pairs will become

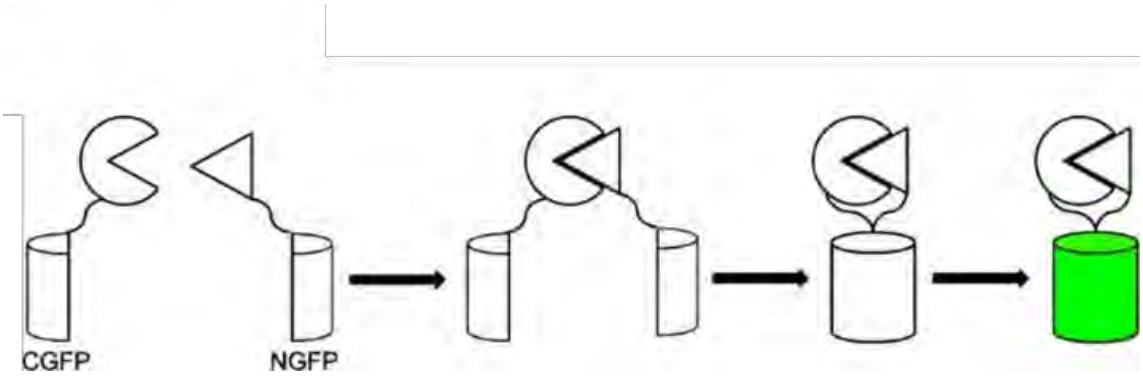



**Figure 1.2**: Cartoon illustrating the use of the split-GFP reassembly assay to screen for protein-protein interactions. GFP is split into two fragments, creating C-terminal GFP (CGFP) and N-terminal GFP (NGFP). To each fragment, one half of a potential interaction pair is fused. CGFP and NGFP do not self-associate, and if the attached proteins do not interact then split-GFP will be nonfluorescent. Only if the CGFP and NGFP are brought in close proximity with one another through association driven by the interaction pair, can split-GFP assemble and form a fluorescent protein (green cylinder).

fluorescent[30]. Split-FPs are a useful tool both for testing whether two proteins interact *in vivo*, as well as an aid in designing new protein-protein interactions, by serving as a method for fluorescently screening libraries of potential interaction partners[31-33]. For a comprehensive review of the split-FP method see Kodama et al.[34]

**Limitations of Fluorescent Proteins**

For most applications that require fluorescently labeling a protein *in vivo*, FP's are the method of choice because of the ease with which proteins can be labeled. FP's are not without their limitations however. As mentioned previously, the photophysical properties of FP's are still inferior to those of small molecule dyes, which can be limiting in some applications. While FP's have been developed that cover a broad spectrum of excitation and emission wavelengths, ranging from blue ($\lambda_{ex}$ = 400 nm/ $\lambda_{em}$ = 450 nm) to red ($\lambda_{ex}$ = 600 nm/ $\lambda_{em}$ = 630 nm), no bright far-red (>700 nm) fluorescent protein has yet been created. IR excitation and emission is necessary for deep-tissue imaging as shorter wavelengths of light can be absorbed by natural components of the cell. Small molecule dyes also provide fluorophores that are brighter, more photostable, and have switching properties that make them superior for super-resolution imaging. The relatively large size of FP's (GFP= 238 amino acids) can also interfere with the proper assembly, localization, and function of the protein to which it is fused. Proteome-wide studies in mammalian cells[35] and yeast cells[36] indicate that the localization of approximately 20 to 25% of proteins is affected by an FP fusion. Plasma membrane transporter proteins appear particularly sensitive, with only 46 of the 139 putative yeast transporter proteins exhibiting any plasma membrane fluorescence when fused to GFP, and only 20 of 139 exhibiting plasma membrane fluorescence with no mislocalization[36, 37].

**Chapter 1.3 - Self-Labeling Enzymes**

Because of the photophysical limitations associated with FP's, various methods have been developed to enable small molecule labeling of proteins inside living cells. Of these, the most successful and widely used tools are self-labeling enzymes. Self-labeling enzymes, such as SNAP-tag[38], CLIP-tag[39], HaloTag[40] and TMP-Tag[41, 42] allow proteins of interest to be labeled with small molecule fluorophores and can serve as a useful tool for applications that require fluorophores with fluorescent properties that cannot be provided by FPs. Self-labeling enzymes bind to and react with a specific small molecule 'handle' that is covalently linked to a fluorophore. Importantly, the reaction between the 'handle' and enzyme is unaffected by the fluorophore conjugated to the 'handle.' Thus, so long as the fluorescent labeling molecule is cell permeable and can reach the self-labeling enzyme, proteins can be labeled with a wide variety of small molecule fluorophores *in vivo* (see Figure 1.3)[39]. Protein labeling is achieved by genetically fusing the self-labeling enzyme to the protein of interest, and then separately adding the fluorescent labeling molecule to media. Reaction between the enzyme and the fluorescent molecule labels the protein of interest.

The strength of self-labeling enzymes lies in their compatibility with a wide range of small molecule fluorophores, many of which are commercially available. Self-labeling enzymes can also enable protein tracking experiments that would be impossible using FP tags. For example, for proteins that self-associate and form large complexes in the cell, possessing many copies of the POI, tracking the dynamics of individual proteins in that complex is impossible if every target protein is labeled. Deliberately adding

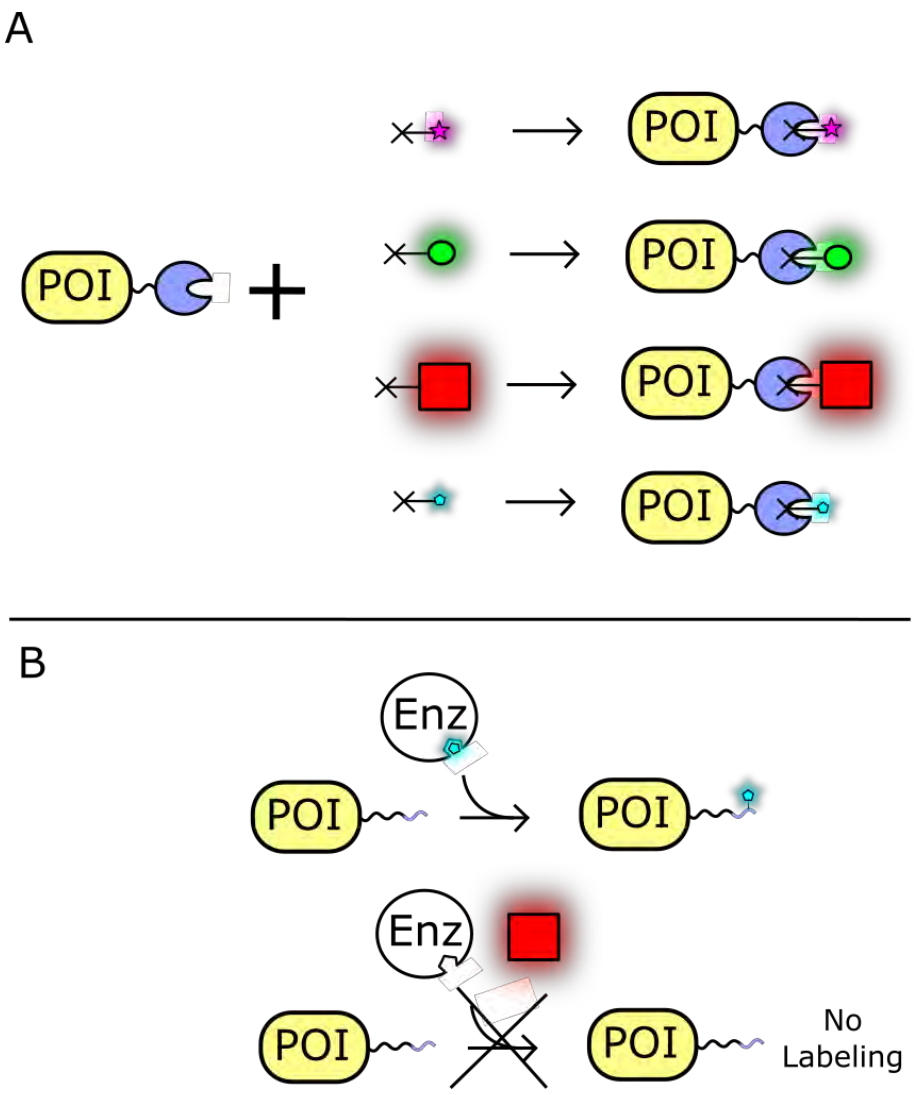



**Figure 1.3**: Comparison of self-labeling enzymes and enzyme-mediated labeling. A) Self-labeling enzymes such SNAP-tag, CLIP-tag, HaloTag, and TMP-Tag (blue circle), are fused to the protein of interest (POI). Fluorescent labeling is accomplished by reaction of the enzyme with a small-molecule that consists of a reactive handle (x) linked to a fluorophore. Reaction between the handle and enzyme is independent of the attached fluorophore, allowing proteins of interest to be labeled with a wide variety of fluorophores with different fluorescent properties (represented by the different shapes fused to the reactive handle). B) In enzyme-mediated labeling, a protein of interest (POI) is tagged with a short peptide tag (blue squiggle), targeting it for modification by an enzyme (Enz). If the fluorophore is bound within the active site of the enzyme (as in PRIME labeling) the size of the fluorophore is limited, and a new enzyme must be engineered for every new fluorophore. Fluorophores that cannot bind within the active site of the enzyme cannot be used to label the protein of interest (bottom).

subsaturating amounts of label, however, labels only a portion of the protein in the complex, allowing the behavior of individual target proteins to be tracked[43]. Pulse-labeling a target protein with different colored fluorophores at different times also allows the temporal dynamics of a protein to be studied[44].

The main disadvantage of self-labeling enzymes comes from their use of an exogenous label. Wash steps are required to remove unreacted label, which complicates the imaging procedure, although versions of SNAP-tag[45] and HaloTag[46] have recently been developed that require no wash step. Once excess label is removed, newly synthesized target proteins will also no longer be labeled. Target labeling is difficult in some model organisms such as *S. Pombe* and *S. Cerevisiae*, which possess cell walls and drug efflux pumps that can prevent small-molecule labels from reaching their intracellular targets. Some workarounds to these permeability issues have been developed, but for the most part they are cumbersome and come with the possibility of interfering with native processes[44, 47]. Self-labeling enzymes are also large (19.4, 20.6, and 33 kDa for SNAP-tag, CLIP-tag, and HaloTag respectively), and therefore have the same potential to interfere with native protein function as fluorescent proteins. , although versions of have been recently developed that require no wash steps.

Related to self-labeling enzymes is the FlAsH/ReAsH labeling method, which uses biarsenical dyes to label 6 amino acid long tetracysteine tags[48, 49]. The great advantage of FlAsH/ReAsH labeling strategy is the small size of the labeling tag. Theoretically, the method should also provide strong signal to noise and require no wash steps, as the dye does not become fluorescent until binding the tetracysteine peptide. Concerns exist, however, regarding labeling specificity, as the biarsenical dyes can bind

to both membranes and other cysteine-rich proteins. Cytotoxicity is a concern as well[50]. Labeling in the presence of thiol-rich compounds, such as 1,2-ethanediol reduces some of these effects, but off-target labeling can still be an issue. The tetracysteine tag can also disrupt normal disulphide-bond formation. RhoBo, a related labeling strategy, uses a bisboronic rather than bisarsenic dye to label a tetraserine tags, which solves the cytotoxicity issues associated with FlAsH/ReAsH. Nonspecific binding is still an issue however, as the tetraserine sequence occurs naturally in many proteins[51].

**Chapter 1.4 - Enzyme-Mediated Labeling**

Another strategy to label proteins in living cells is to fuse a short peptide tag to the protein of interest, which acts as a substrate recognition sequence to which an enzyme will covalently attach a fluorophore. Various enzyme/tag combinations have been used, including biotin ligase[52], bacterial phosphopantetheinyl-transferases (PPTases)[53], sortases[54, 55], and lipoic acid ligase[56]. Theoretically, the small size of the peptide tag and the high-specificity of the enzymes should make enzyme-mediated labeling a powerful approach. Labeling a target protein in the complex intracellular environment is difficult, however, and most enzyme-mediated labeling strategies cannot be used to label proteins inside cells. Enzyme-mediated labeling can still serve as a useful tool for labeling proteins in the extracellular environment, however, such as at the cell-surface. Biotin ligase offers a particularly convenient method for labeling extracellular portions of cell surface proteins, and is compatible with a wide range of fluorophores, requires only a small modification to the POI, and labels only proteins that are at the cell-surface[52]. POIs are fused to a 15 amino acid long acceptor peptide (AP). To label proteins, BirA, the *E. Coli* biotin ligase, and biotin are first added to the growth medium to initiate biotinylation

of the AP tag. Cells are then washed, incubated with fluorescently labeled monovalent streptavidin, and washed again. Binding between biotin and streptavidin is tight and essentially irreversible, allowing protein tracking over long periods of time.

The PRIME method (PRobe Incorporation Mediated by Enzymes), which uses a derivative of lipoic acid ligase (LplA), is one of the few examples of enzyme-mediated labeling that has been used successfully to label intracellular proteins[56]. LplA naturally ligates lipoic acid onto substrate proteins involved in oxidative metabolism in *E. Coli*. Proteins are recognized for ligation through a short sequence of 13 amino acids, which is transferable to other proteins. To create a fluorophore ligating version of LplA, the enzyme's active site was mutated and enlarged so as to accommodate the fluorophore 7-hydroxycoumarin. In contrast to the self-labeling enzymes discussed above, in PRIME the fluorophore is bound within the active site of the conjugating enzyme. While innovative, having the enzyme bind the fluorophore directly limits both the maximum size of the fluorophore that can be conjugated, and requires that a new enzyme be created for every new fluorophore (see Figure 1.3). The fluorophore conjugated in first demonstration of PRIME (7-hydroxycoumarin) fluoresces at shorter wavelengths (excitation 387–405 nm; emission 448 nm), which produce cellular autofluorescence that can be limiting for some applications[57].

As with any method that relies on an exogenous fluorophore, enzyme-mediated labeling requires wash steps to remove unincorporated fluorescent label. Until versions of the enzyme-mediated labeling method can be developed that are functional in the intracellular environment and can accommodate a wider range of fluorophores, enzyme-mediated labeling is best suited for studies that require labeling outside the cell, in either

the *in vitro*[55] or extracellular[52] environment, where it can provide highly specific protein labeling with minimal modifications to the protein of interest.

**Chapter 1.5 - Unnatural Amino Acids**

Unnatural amino acids (UAAs) offer another route for fluorescently labeling proteins with small molecule fluorophores[58]. UAAs are amino acids that have been created by scientists for the purpose of decorating proteins with novel functional groups that do not naturally exist within proteins. Of relevance to fluorescent labeling are UAAs that have fluorescent side-chains, or side-chains with reactive functional groups that can be labeled with a fluorophore in a second step (see Figure 1.4). Although direct incorporation of a fluorescent amino acid is simpler than two-step labeling, the number of fluorophores that have been introduced in this manner is limited, and those that have been incorporated fluoresce at short wavelengths (excitation below 360 nm, and emission below 500 nm)[59, 60]. Labeling a UAA in a second step allows a greater variety of fluorophores to be used, but requires a labeling reaction that can proceed under intracellular conditions and react with only the target protein. While difficult, intracellular labeling of intracellular target proteins such as the nuclear protein Jun[61] and the cytosolic protein actin[62] has been achieved with both high efficiency and minimal off-target labeling using a two-step strategy. The advantage of using UAAs to fluorescently label protiens is the ability to place fluorophores with high precision within a target protein, as well as the extremely small modification that is introduced in the process. The fluorescence of some UAAs is sensitive to the environment, being affected by both solvent and nearby amino acids. Through careful placement, these properties allow UAAs to serve as sensitive reporters of protein-DNA interactions[63], folding dynamics[64],

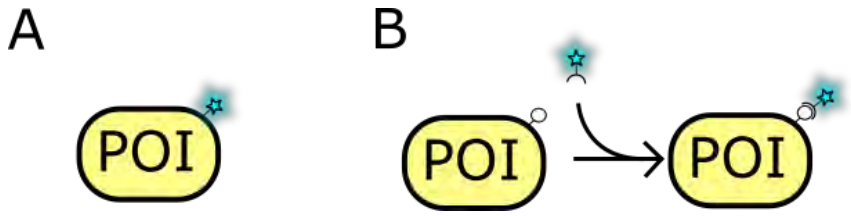



**Figure 1.4**: Use of unnatural amino acids to fluorescently label proteins. Proteins of interest can be labeled either directly through incorporation of an intrinsically fluorescent amino acid (A) or indirectly by incorporating a reactive functional group that is labeled with a fluorophore in a second reaction (B).

protein surface hydration[65] and protein conformation changes induced by ligand binding[66].

A limitation of using UAAs to label proteins in vivo is the requirement of extensive wash steps to remove any UAA and/or fluorescent label. Additionally, the efficiency of UAA incorporation varies depending on the surrounding mRNA sequence. Off-target incorporation is also possible, although strategies are being developed to address this issue[67].

**Chapter 1.6 - Protein-Protein Interactions**

Protein-protein interactions provide another route to achieving intracellular protein labelling. In this strategy, target proteins are labeled indirectly, using a fluorescently labeled protein binding domain that recognizes the target protein. Typically, DNA encoding the binding protein fused to a fluorescent protein is introduced into the cell, and the fusion protein is expressed *in situ*. Although less common, another option is to use physical or chemical methods such as microinjection or mild detergents to introduce a chemically labelled binding protein into the intracellular environment. Interaction between the fluorescently labeled binding protein and the POI localizes the fluorophore to the POI, allowing target visualization.

Theoretically any binding protein that is functional in the intracellular environment and binds the target protein specifically and with sufficient tightness can be used to visualize a POI. Binding proteins that have been used to visualize target proteins include single chain Fragment variable (scFV's)[68], nanobodies[69], fibronectin[70], Designed Ankyrin Repeat Proteins (DARPins)[71] and Tetratricopeptide Repeat Affinity Proteins

(TRAPs)[72]. To the best of our knowledge, the weakest binding interaction that has been successfully used to label a protein *in vivo* is TRAP binding domain which binds its cognate peptide tag with a $K_d$ of ~ $10^{-6}$ [72]. Pratt et al. systematically tested the effect of decreasing the strength of the TRAP-peptide binding interaction, and found that for $K_d$'s above 1.7 μM, protein visualization became impossible. For a comprehensive review of protein labeling mediated by protein-protein interactions, see Kaiser et al[73]. Full-length antibodies cannot be used to recognize target proteins because of their dependence on internal thioester bonds, which cannot form in the reducing intracellular environment. In addition, protein-protein interactions can disrupt normal protein function depending on where the binding protein binds the POI, meaning that even when binding proteins are being used to visualize completely unmodified proteins, it is still possible for native protein function to be perturbed[71, 74].

An important advantage of using protein-protein interactions to fluorescently label proteins is the minimal modification that is required to visualize the POI. If a binding protein is available that binds directly to the POI, it is even possible to track proteins without making any modification to the protein of interest (see Figure 1.5A)[69, 75-77]. Binding proteins can also be developed that recognize specific post-translational modifications, protein conformations, or oligomeric states, allowing the visualization and tracking of particular subpopulations of a POI. Using an scFv fused to EGFP, this strategy was used to track acetylated histones, and detected enrichment in euchromatin that changed during embryogenesis and in response to deactylase inhibitors[68].

Although effective, using binding proteins to fluorescently label proteins requires generating a new binding domain for every POI. A more scalable approach is to attach a

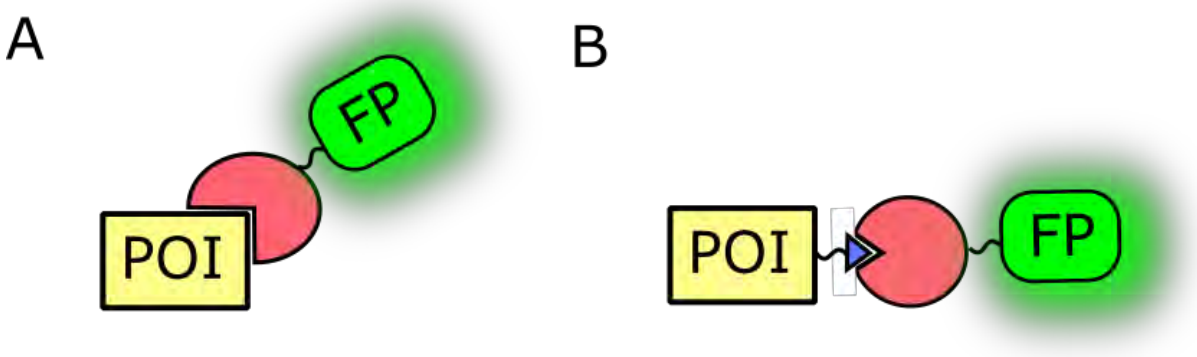



**Figure 1.5**: Labeling proteins of interest (POI) using protein-protein interactions. A) POIs can be fluorescently labeled by direct interaction with a binding protein, if a binding protein (light red circle with cutout) is available that binds the POI directly. B) POI's can also be recognized by attaching a short recognition tag to the protein of interest (blue triangle) for which a binding protein already exists (light red circle with cutout). In both cases, binding proteins are typically fluorescently labeled by fusion to a fluorescent protein (FP).

short recognition tag to the POI for which there already exists a well characterized protein binding partner (see Figure 1.5B). The SunTag method uses this strategy to fluorescently label target proteins, using an scFv fused to EGFP to recognize copies of a 19 amino acid long peptide[78]. By fusing many peptide tags in tandem, the authors were able to recruit up to 24 copies of scFv-GFP to a POI, greatly amplifying target signal. Such amplification is an asset for single-molecule studies and imaging proteins over long periods of time, but the large size of the tag (64 kDa per copy when not bound to scFv-GFPs) reduces the usefulness of this strategy.

The great limitation of using protein-protein interactions to label POIs is the high amount of background signal that is produced by unbound binding domain. Using protein-protein interactions with increased affinity reduces the amount of unbound binding protein, but only so long as the concentration of binding protein is kept below that of the POI. The recently developed split-FP labeling system solves these background issues by using a binding protein that does not become fluorescent until binding its cognate recognition tag[79]. This imaging strategy takes advantage of the split-FP modules mentioned previously that are used for sensing protein-protein interactions. The original split-FP's cannot self-associate, and do not become fluorescent unless fused to a protein-protein interaction partner. By splitting sfGFP, the Huang lab was able to develop a split-FP pair that can self-associate and form a fluorescent protein independent of fusion partners. Target proteins are visualized using this split-FP pair by fusing one half to the protein of interest, and separately expressing the other half. Because the binding protein does not become fluorescent until binding the other half of split-FP, POIs can be labeled with none of the background fluorescence observed using other methods (see Figure 1.6).

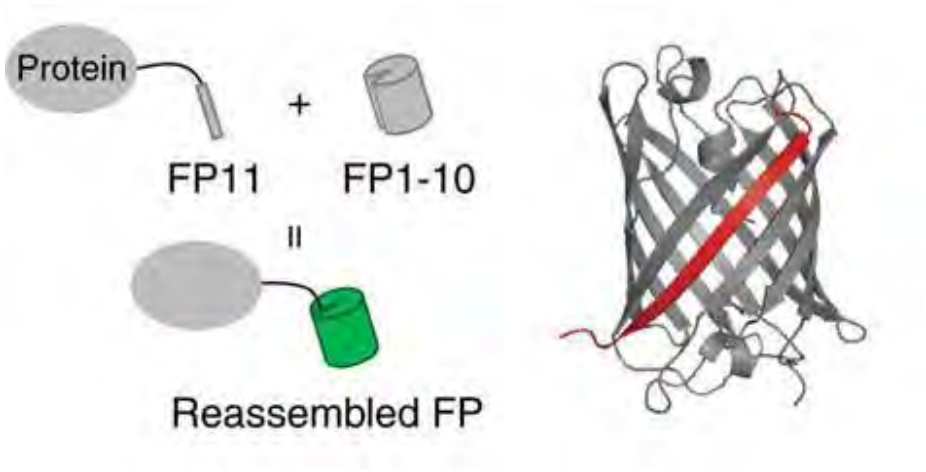



**Figure 1.6**: Use of split-GFP to visualize target proteins *in vivo*. By splitting sfGFP, Kamiyama et al. were able to generate a split-GFP pair capable of self-associating and forming a fluorescent fold without the help of any attached interaction partners[79]. Proteins of interest are tagged with one half of split-GFP (FP11), and the other half (FP1-10) is expressed separately. FP 11 and FP1-10 associate and form the fluorescent GFP fold, fluorescently labeling the protein of interest. Shown to the right is a ribbon structure of the sfGFP protein, with FP11 colored in red. Reproduced with permission from Kamiyama et al.[79]

Split-FP labeling has since been extended to allow mCherry labeling[80], but still offers a limited selection of fluorophores. Split-FP labeling is also reversible (albeit slowly), and therefore cannot be used to study temporal dynamic processes such as the accumulation of post-translational modifications over time or the exchange of proteins between large complexes, as the labeling protein will exchange between newly synthesized and already existing proteins.

**Chapter 1.7 - Aims of this work**

Protein-protein interactions show great promise for labeling proteins *in vivo*, but are currently limited by high amounts of background fluorescence and/or limited fluorophore selection. I hypothesized that SpyCatcher/SpyTag, a permanent protein/peptide interaction pair, could be used to fluorescently label proteins *in vivo*, and that the irreversible nature of this interaction would generate less background fluorescence than previous methods that use reversible interactions. Additionally, the permanent nature of the interaction could potentially allow the study of protein temporal dynamics. In Chapter 2, I outline a general approach for using SpyoIPD/SpyTag, a more stable derivative of SpyCatcher/SpyTag, to fluorescently label target proteins in *S. cerevisiae*. I demonstrated that SpyoIPD/SpyTag can be used to label a range of proteins in yeast, and that labeling in this manner can be less deleterious to protein function than a direct fluorescent protein fusion. In collaboration with Peter Swain's lab at the University of Edenborough, we have also shown that pulsed labeling with SpyoIPD/SpyTag can be used to measure the *in vivo* half-life of Pma1, a plasma membrane protein. In Chapter 3, I outline the development and characterization of a nuclear sequestration strategy for reducing background fluorescence, and in Chapter 4, I

describe the use of additional irreversible protein-peptide interaction pairs to fluorescently label proteins in yeast.

Previous work in the Regan lab has shown that TRAP domains can be used to visualize FtsZ, a bacterial homologue of Tubulin, in live *E. coli*[72]. TRAP-based imaging requires attaching on a very short, 5 amino-acid long peptide to the C-terminus of POIs, and should be readily compatible with any POI. One of the goals of this work was to test whether this TRAP-based imaging could also be used to fluorescently label proteins in *S. cerevisiae*. Chapter 5 presents the results of this work and shows that while TRAP-based imaging can be used to visualize a handful of proteins *in vivo*, in its current form its high background fluorescence precludes its usefulness as a general method for fluorescently labeling proteins in *S. cerevisiae*.

# Chapter 2 - Using SpyoIPD/SpyTag to Visualize Proteins in Living *S. Cerevisiae*.

[Text and Figures adapted from Hinrichsen, M. *et al.* (2017). A new method for post-translationally labeling proteins in live cells for fluorescence imaging and tracking. *Protein Engineering, Design and Selection*. **30**, 771-780][81].

## Chapter 2.1 - Introduction

One of the goals of this work is to develop an improved protein-peptide interaction-based method for fluorescently labeling proteins in living cells. A major limitation of previous methods is the large amount of background fluorescence that is produced by unbound binding domain. Using a tighter binding protein-peptide interaction pair should reduce the fraction of binding domain that is unbound and reduce background fluorescence. I hypothesized that SpyCatcher/SpyTag, a protein-peptide pair that bind one another permanently *via* an isopeptide bond, would provide excellent signal to noise for fluorescently labeling target proteins, as the interaction is essentially irreversible. I therefore set out to test the ability of SpyCatcher/SpyTag to fluorescently label a collection of target proteins in *S. cerevisiae*.

## Chapter 2.2 - Testing SpyCatcher and SpyCatcher Variants *In Vivo*

I first tested whether the original SpyCatcher/SpyTag pair is active in yeast, as to the best of our knowledge this has not be previously reported. I created a yeast strain coexpressing SpyCatcher and EGFP tagged at the C-terminus with SpyTag. For the purpose of this experiment, EGFP serves simply as a convenient handle to increase the

mass of SpyTag, so as to allow visualization by SDS-PAGE. Because SpyCatcher forms a covalent bond with SpyTag, the conjugate species is resistant to SDS denaturation and can be detected as a higher molecular weight species in a Western Blot. Under steady-state expression conditions, only a small fraction of EGFP-SpyTag was found to be covalently labeled (Figure 2.1D). Unconjugated SpyCatcher was never detected *in vivo*, indicating that EGFP-SpyTag labeling does not go to completion because the amount of SpyCatcher is limiting. This was surprising given the strong promoter and high-copy number plasmid used to drive SpyCatcher expression (p424 GAL1).

## Chapter 2.3 - Design and Characterization of SpyoIPD, a More Stable SpyCatcher Derivative

The second CnaB domain of streptococcal surface protein FbaB contains a covalent isopeptide bond between a Lys on the N-terminal β-strand and an Asp on the C-terminal β-strand[82]. Splitting this domain gave rise to the SpyCatcher/SpyTag system, where the β-strand containing the Asp residue (SpyTag) is expressed separately from the remainder of the protein (SpyCatcher). SpyTag associates and reacts with SpyCatcher to form the isopeptide bond between the Asp and Lys sidechains in *trans*, so that each β-strand now comes from a separate protein[83]. Independently, the Schwarz-Linek lab at the University of St. Andrews developed more structurally stable derivatives of SpyCatcher by designing "open" isopeptide domains (SpyoIPDs) that retain mutated portions of the C-terminal β-strand that was removed to create SpyTag. The reintroduced β-strand was mutated to remove the reactive Asp (Asp556Ala) involved in isopeptide bond formation. Additional mutations were also introduced into the β-strand so as to weaken its interaction with the rest of the protein. This was necessary, as the reintroduced strand

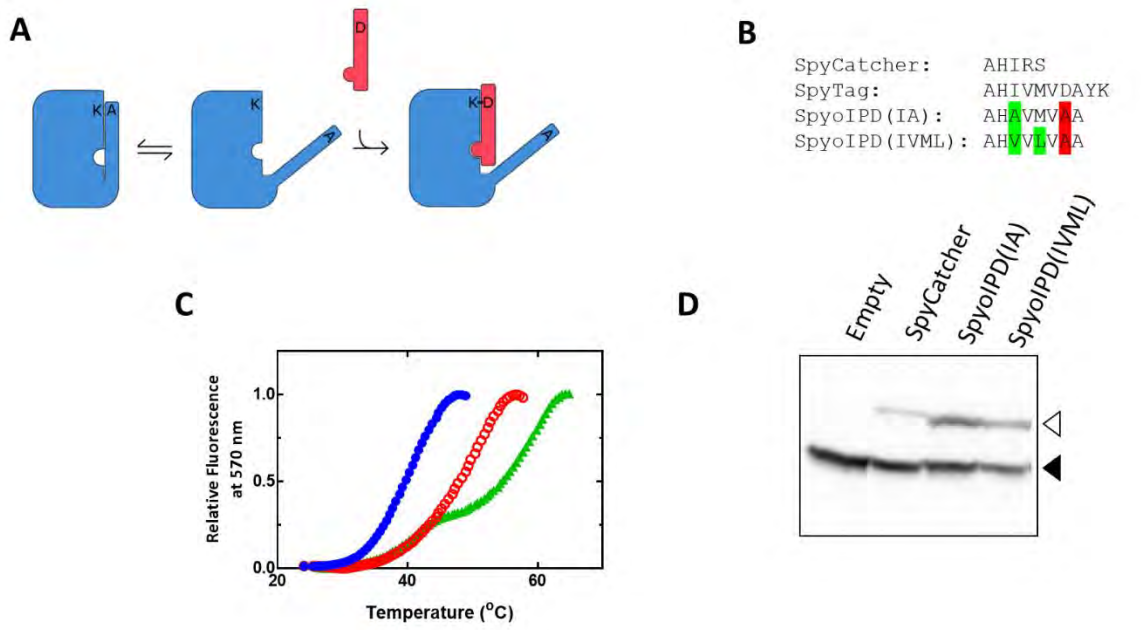



**Figure 2.1:** Design and characterization of SpyoIPDs. A) To improve the stability of the SpyCatcher protein (blue rectangle), portions of the C-terminal β-strand (blue thin rectangular 'overhang') that was originally removed to make SpyTag were reintroduced back onto SpyCatcher. The reactive Asp on this extension was mutated to Ala (D556A) to prevent reaction with the Lys in the SpyCatcher region, and the appended sequence was also mutated to weaken its interaction with the rest of the domain, allowing SpyTag (red thin rectangle) to displace the reintroduced β-strand and react with the SpyCatcher domain. B) Comparison of the C-terminal sequences of SpyTag, SpyCatcher, SpyoIPD(IA) and SpyoIPD(IVML). The Ala that replaces the isopeptide bond-forming Asp is highlighted in red. The mutations made to weaken the interaction between SpyCatcher and the reintroduced sequence are highlighted in green. C) Differential scanning fluorimetry traces of SpyCatcher (blue circles), SpyoIPD(IA) (red hollow circles) and SpyoIPD(IVML) (green triangles). D) Comparison of the *in vivo* activity of SpyCatcher, SpyoIPD(IA), and SpyoIPD(IVML). SpyTag was expressed as a fusion to EGFP from a medium strength promoter on a low copy number plasmid. An N-terminal V5 epitope was also fused to EGFP to facilitate detection. Yeast lysate separated by SDS-PAGE, and GFP-SpyTag detected by Western blot via chemiluminescence, using an anti-V5 mouse antibody, and HRP conjugated secondary. Lanes and bands are as labeled. The lower molecular weight band corresponds to un-reacted EGFP-ST (filled triangle), and the higher molecular weight band to the covalent EGFP-ST-SC or EGFP-ST-SpyoIPD conjugate (hollow triangle).

binds at the same site as SpyTag, and if bound tightly, would inhibit the reaction between SpyoIPD and SpyTag (see Figures 2.1A and 2.1B). Two SpyoIPD variants were created, containing different point mutations at the SpyCatcher/ β-strand interface. SpyoIPD(IA) contains the mutation (Ile552Ala), and SpyoIPD (IVML) contains the mutations Ile552Val and Met554Leu. $^{1}H$-$^{15}N$ HSQC NMR spectra collected in the Schwarz-Linek lab show that SpyCatcher, SpyoIPD(IA), and SpyoIPD(IVML) are folded in solution, even in the absence of SpyTag (Appendix Figure A1). Differential Scanning Fluorimetry (DSF) analysis indicate that both SpyoIPD(IA) and SpyoIPD(IVML) are more thermally stable than SpyCatcher (Figure 2.1C). Unexpectedly, SpyoIPD(IVML) exhibited atypical DSF traces that do not fit standard protein melt curves. It is unclear what is producing these abnormal results.

**Chapter 2.4 - Testing the Activity of SpyoIPD(IA) and SpyoIPD(IVML) Variants *In Vivo***

To compare the *in vivo* activities of SpyoIPD(IA) and SpyoIPD(IVML) with SpyCatcher, I used the western blot assay used previously to test SpyCatcher activity, and compared yeast strains expressing SpyoIPD(IA), SpyoIPD(IVML), or SpyCatcher together with EGFP tagged at the C-terminus with SpyTag. Both SpyoIPD(IA) and SpyoIPD(IVML) labeled greater fractions of EGFP-SpyTag than SpyCatcher, with the greatest fraction labeled observed using SpyoIPD(IA) (Figure 2.1D). SpyoIPD(IA) was therefore used for subsequent imaging applications, and will be referred to henceforth as SpyoIPD.

**Chapter 2.5 - Design of a SpyoIPD/SpyTag Based Imaging System**

Our strategy for fluorescently labeling proteins in living cells is illustrated schematically in Figure 2.2. The genomic DNA sequence encoding the target protein is tagged at the 3' end with DNA encoding SpyTag. Although in principle SpyTag can be placed anywhere in the target protein as long as it is accessible, for consistency and convenience I typically place SpyTag at the C-terminus. Note that the gene encoding the target protein fused to SpyTag replaces the wild-type copy and is expressed from the target protein's endogenous promoter. SpyoIPD fused to an FP is expressed from the strong, galactose-inducible promoter, GAL1. To gain more precise control over SpyoIPD-FP expression levels, SpyoIPD-FP was inserted at the GAL2 locus so as to simultaneously delete the GAL2 galactose permease. Whereas expression from the GAL1 promoter is sigmoidal in wild-type cells, deleting GAL2 has previously been shown to make expression from GAL1 linear with respect to galactose concentration[84]. Precise control of SpyoIPD-FP expression levels is critical in order to keep the concentration of SpyoIPD-FP below that of target protein and minimize the amount of background fluorescence that is produced from unbound SpyoIPD-FP. Target proteins are labeled by adding galactose to induce expression of SpyoIPD-FP, which then reacts with SpyTag, and covalently labels the target protein with an FP. For initial imaging experiments SpyoIPD was fused to EGFP.

### Chapter 2.6 - Imaging Proteins of Interest in Live *S. Cerevisiae* with SpyoIPD/SpyTag

To test the ability of SpyoIPD-EGFP to fluorescently label SpyTagged protein *in vivo*, initial imaging candidates were chosen that are abundant, not known to have an

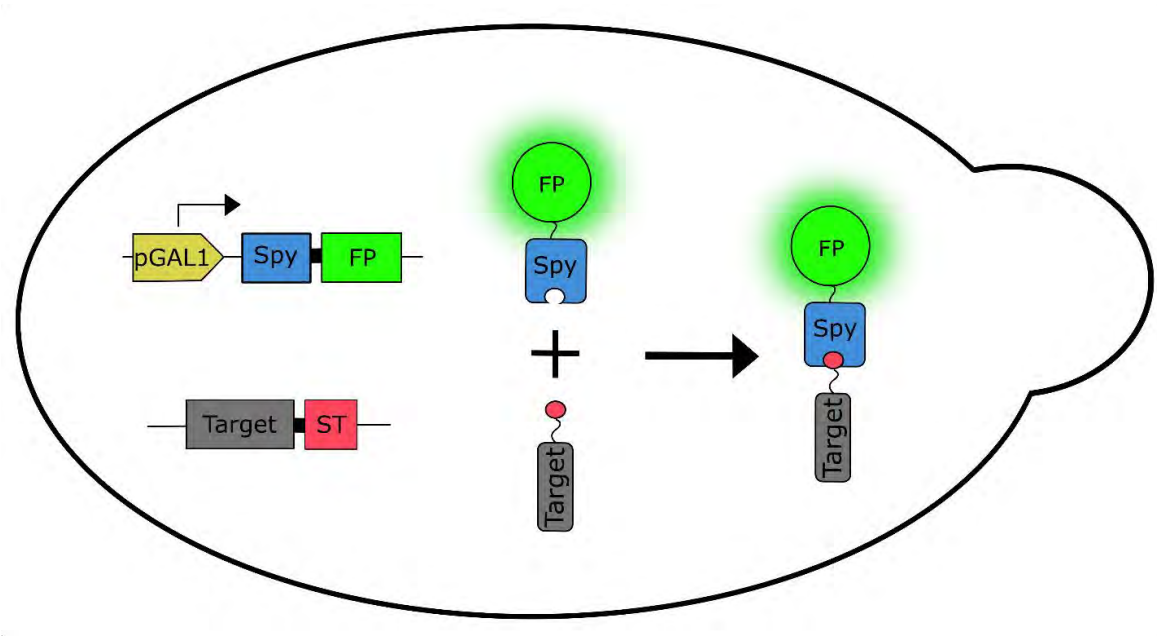



**Figure 2.2:** Schematic illustration of the SpyoIPD/SpyTag labeling strategy. The genomic copy of the gene encoding the target protein (Target, gray) is fused at the 3' end to a sequence encoding the SpyTag (ST, red), replacing the chromosomal copy of the target's gene. Expression of the SpyTagged target protein is from the target's endogenous promoter. DNA encoding a fusion of SpyoIPD (Spy, blue) and FP (FP, green) is integrated at the GAL2 locus, simultaneously deleting GAL2 in the process. Expression of SpyoIPD-FP is from the GAL1 promoter (pGAL1, gold). Once expressed, the SpyoIPD-FP reacts with the SpyTagged target protein, fluorescently labeling the target protein.

inaccessible C-terminus, and which localize to a distinct region of the cell. Here I present data on four proteins from different subcellular compartments: the Plasma Membrane ATPase (Pma1), the histone H2B (Htb2), the cell division cycle 12 protein (Cdc12, a component of the septin ring that localizes to the bud neck), and the Vacuolar Membrane ATPase (Vma1). In each imaging experiment, cells were grown overnight in non-inducing medium, diluted the next day into galactose-containing medium, and grown an additional 8 hours before imaging.

HTB2 encodes the core histone protein H2B, which is required chromatin assembly and naturally localizes exclusively to the nucleus. When HTB2 is fused directly to EGFP, sharp rings of fluorescence are observed near the center of the cell, corresponding to the nucleus. A similar localization pattern is observed when HBT2 is tagged with SpyTag and co-expressed with SpyoIPD-EGFP. Minimal cytoplasmic fluorescence is observed, indicating that the majority of SpyoIPD-EGFP is bound to Htb2-SpyTag.

CDC12 localizes to the bud neck[85], and is a component of the septin ring, a large complex of proteins that serves as a scaffold for recruiting cell division factors that drive cytokinesis. When Cdc12 is visualized by immunofluorescence in fixed and permeabilized cells, tight rings of fluorescence are observed around the bud necks of dividing cells[86]. A similar pattern is observed in live cells expressing Cdc12-SpyTag and SpyoIPD-EGFP. In the majority of cells expressing Cdc12 directly fused to EGFP, fluorescent rings are also observed. Approximately 5% of cells expressing Cdc12-EGFP, however, display an elongated cellular morphology, indicating that cytokinesis is not progressing as normal and that Cdc12 function is being affected. This phenotype is not

when Cdc12 is imaged using SpyoIPD/SpyTag. This suggests that SpyoIPD/SpyTag mediated labeling may be less disruptive to Cdc12 function than a direct GFP fusion.

PMA1 is an ATPase that is the primary plasma membrane proton pump in *S. cerevisiae*. PMA1 plays an essential role in maintaining cytoplasmic pH and membrane voltage potential (pumping protons out of the cytoplasm) and is one of the most highly expressed proteins in yeast. Cells expressing Pma1-SpyTag and SpyoIPD-EGFP display a ring of fluorescence around the cellular periphery, consistent with immunofluorescence data that shows localization of Pma1 to the plasma membrane[87]. By contrast, cells expressing Pma1 directly fused to EGFP display both plasma membrane fluorescence and strong vacuolar fluorescence (Figure 2.3).

Vma1 is a subunit of the V-ATPase, the primary proton pump of the vacuole that is essential for acidifying the vacuole. While Vma1 is not an integral membrane protein, it naturally localizes strongly to the vacuolar membrane. Of the four proteins tested, only Vma1 failed to produce a clear localization signal when labeled with SpyoIPD/SpyTag. Cells expressing Vma1 directly fused to EGFP display tight rings of fluorescence that correspond to the vacuolar membrane. Most cells expressing Vma1-SpyTag and SpyoIPD-EGFP, in contrast, display diffuse cytoplasmic fluorescence, with a subset of cells displaying what may be a weak ring of fluorescence around the vacuole (Figure 2.3 shows an example of one of the more strongly labeled cells).

Also included in Figure 2.3 is a fluorescent image of cells expressing SpyoIPD-EGFP and no SpyTagged protein. Diffuse cellular fluorescence is for the most part observed, with some accumulation in what appears to be the nucleus. It is unclear what is

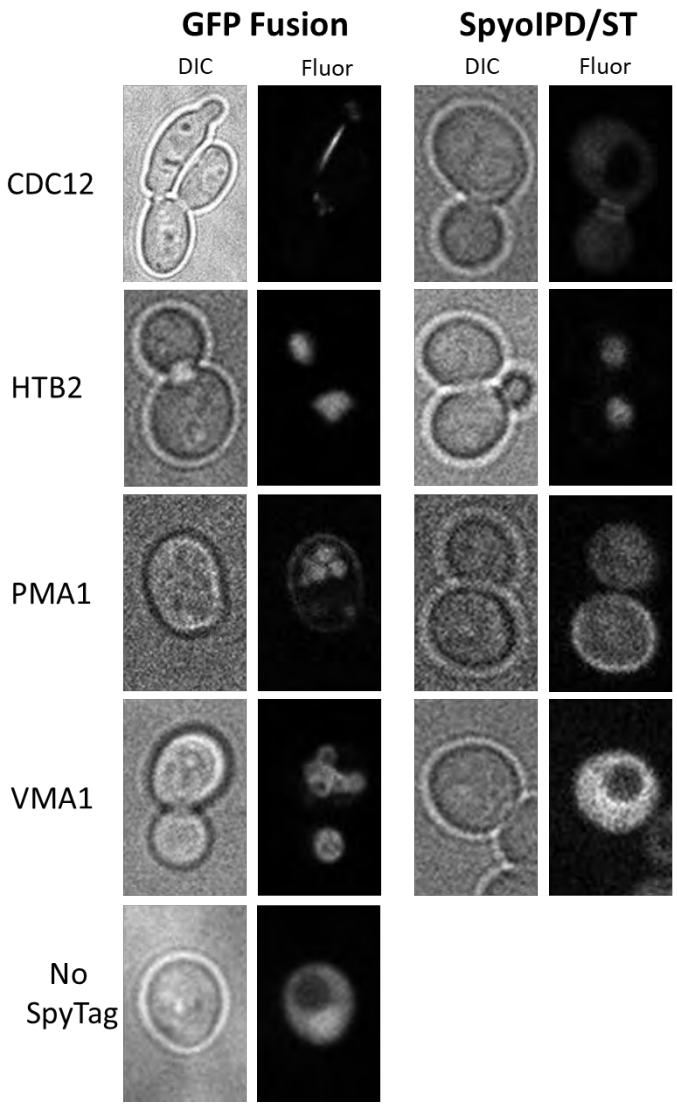



**Figure 2.3**: A comparison of proteins labeled by direct fusion to EGFP with those labeled using SpyoIPD/SpyTag labeling. Brightfield (DIC) and fluorescence (Fluor) images are shown for target proteins fused directly to EGFP (GFP Fusion) and the same proteins labeled using SpyoIPD/SpyTag (SpyoIPD/ST). The identity of the target protein is given to the left. The 'No SpyTag' strain expresses SpyoIPD-EGFP, but no SpyTagged protein.

driving this nuclear accumulation, although nuclear accumulation of free EGFP has been previously reported[88].

## Chapter 2.7 - Characterizing the Effects of Different Labeling Strategies on Pma1 Function

Pma1 is an essential plasma membrane proton pump in yeast that is responsible for maintaining cytosolic pH and plasma membrane potential[89]. Pma1 has been proposed to play a role in cell aging[90] and has been used as a model protein to study protein quality control pathways in the secretory system[91]. Immunofluorescent labeling of epitope-tagged Pma1 in fixed cells shows Pma1 normally localizes exclusively to the plasma membrane (Figure 2.4)[87]. By contrast, when Pma1 is fused directly to a fluorescent protein, fluorescence is observed both at the plasma membrane and the vacuole, indicating that directly attaching a fluorescent protein to Pma1 is somehow interfering with normal protein localization (Figure 2.4). Yeast expressing Pma1 directly fused to an FP also exhibit compromised cell growth (Figure 2.4). Cells expressing Pma1-SpyTag and SpyoIPD-EGFP exhibit neither vacuolar mislocalization nor a growth defect (Figures 2.3 and 2.4).

## Chapter 2.8 - Improving Signal to Noise when Labeling Pma1-SpyTag with SpyoIPD-EGFP

As can be seen in Figure 2.3, when Pma1-SpyTag is labeled with SpyoIPD-EGFP, significant amounts of cytoplasmic fluorescence is observed due to unbound SpyoIPD-EGFP. One potential explanation for this high amount of background fluorescence is that the concentration of SpyoIPD-EGFP is in excess of Pma1-SpyTag. I

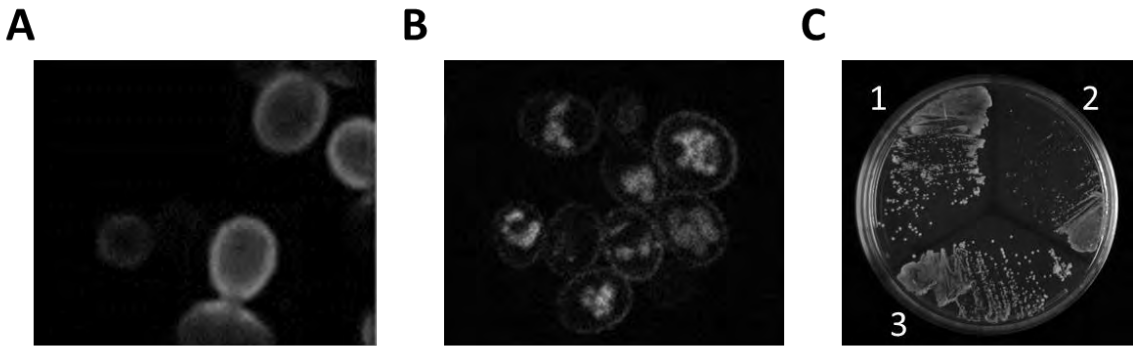



**Figure 2.4**: Effects of different labeling methods on Pma1 localization and function. A) Immuno-staining of fixed yeast cells (using anti-HA antibodies) in a strain expressing Pma1 fused to the HA peptide. This 'native' Pma1 localizes exclusively to the plasma membrane, with none evident in the vacuole. Reproduced with permission from Mason et al. B) Live cell imaging of yeast expressing a Pma1-EGFP fusion protein, expressed from the endogenous Pma1 promoter. A significant amount of fluorescence is observed in the vacuole in addition to that present at the plasma membrane. C) Comparison of the growth of yeast expressing untagged Pma1 (1), Pma1 C-terminally tagged with mCherry (2), or Pma1 C-terminally tagged with SpyTag (3). Strains are streaked on media containing 2% galactose, so that yeast expressing Pma1-SpyTag are also expressing SpyoIPD-EGFP. In both tagged strains Pma1 is expressed under control of its native promoter, and the tagged copy of the strain is the only copy of the protein present.

therefore investigated whether reducing the concentration of galactose used to induce SpyoIPD-EGFP expression would reduce cytoplasmic fluorescence when imaging Pma1-SpyTag. At low concentrations of galactose, fluorescent signal from both the cytoplasm and the plasma membrane is weak, making resolution of plasma membrane signal difficult (Figure 2.5A). Increasing the concentration of galactose increases the intensity of plasma membrane signal, but also increases the diffuse cytosolic background so that the improvement in membrane to cytosol signal is negligible.

A distinctive characteristic of Pma1's spatial distribution is that it is retained by mother cells during cell division, so that little to no Pma1 is inherited by daughter cells. Because of this asymmetric division, the irreversible nature of the SpyoIPD-SpyTag interaction and the long half-life of Pma1 (*vide infra*), I predicted that unbound cytosolic signal could be cleared and plasma membrane signal retained if new SpyoIPD-EGFP expression were turned off following labeling. Since GAL1 is inhibited by glucose, I performed experiments in which SpyoIPD-EGFP expression was first induced with galactose and then shut off with glucose, and samples were imaged at different time point following the addition of glucose (Figure 2.5B). Using this strategy significantly increased the ratio of membrane to cytosolic fluorescence in Pma1-SpyTag expressing yeast. The amount of time required to clear cytosolic signal depended on the concentration of galactose used to induce SpyoIPD-EGFP expression, with higher concentrations taking longer to clear (data not shown). A pulse-chase labeling strategy was also used to visualize Vma1-SpyTag, and found to significantly improved fluorescent labeling (see Appendix Figure A3). I attempted to determine the rate at

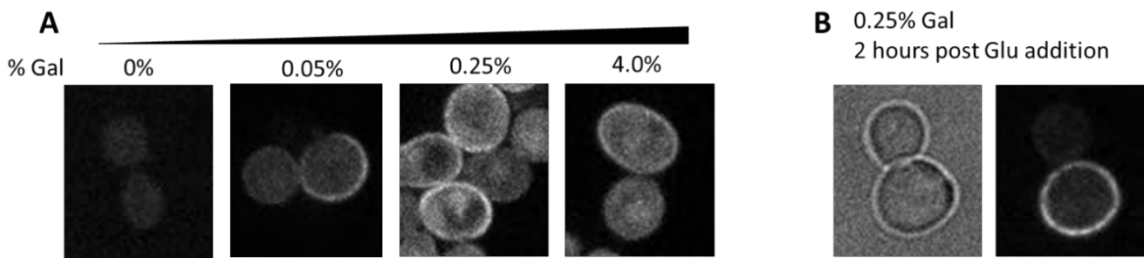



**Figure 2.5**: Strategies for reducing cytoplasmic fluorescence when labeling Pma1-SpyTag with SpyoIPD-EGFP: A. Fluorescent images of yeast cells expressing Pma1-SpyTag, in which SpyoIPD-EGFP expression has been induced by the indicated concentration of galactose (units in % w/v). Due to the large differences in SpyoIPD-EGFP expression levels, it was necessary to image samples with different exposure times (1000 ms for [Gal] = 0%; 200 ms for [Gal] = 0.05%; 200 ms for [Gal] = 0.25%, and 20 ms for [Gal] = 4.0%). B. Yeast expressing Pma1-SpyTag and SpyoIPD-EGFP were induced for 8 hours in media containing 0.25% galactose, switched into glucose containing media, and grown for an additional 2 hours before imaging. Glucose inhibits the GAL1 promoter, turning off new synthesis of SpyoIPD-EGFP (100 ms exposure).

which SpyoIPD-EGFP reacts with EGFP-SpyTag *in vivo*, but were never able to simultaneously detect both unreacted EGFP-SpyTag and unreacted SpyoIPD-EGFP in the same cell, indicating the reaction proceeds quickly in yeast (data not shown).

**Chapter 2.9 - Spatiotemporally Tracking Pma1 in Living Cells**

The labeling strategy that I present can also be used to follow a protein's spatiotemporal dynamics in cells. Because the protein of interest is labeled post-translationally, only protein that is present when SpyoIPD-FP is expressed will be labeled. Thus, turning off expression of SpyoIPD-FP allows one to follow the fate of only the protein that was present during the labeling phase. In collaboration with the Swain Lab at the University of Edinburgh, we used this strategy to transiently label Pma1 and determine the half-life of Pma1 at the cell membrane. To follow individual cells over many hours, we used a microfluidic device that holds individual mother cells in place, but allows daughter cells to be washed away by media flow (see Appendix Figure 2)[92]. Cells were first grown overnight in the presence of galactose to induce expression of SpyoIPD-EGFP and label Pma1-SpyTag. Cells were then loaded into the microfluidic device and, after a short equilibration period, switched to glucose containing media to inhibit expression of SpyoIPD-EGFP. Cells were tracked for many hours following the switch to glucose containing medium.

We estimated the half-life of in-membrane PMA1 in two complementary ways. In the first method, we assumed that the total cellular fluorescence of an individual cell expressing Pma1-SpyTag and SpyoIPD-EGFP is comprised of i) fluorescence from SpyoIPD-EGFP covalently bound to Pma1-SpyTag at the membrane; (ii) fluorescence from unreacted cytosolic SpyoIPD-EGFP and (iii) cellular auto-fluorescence. We

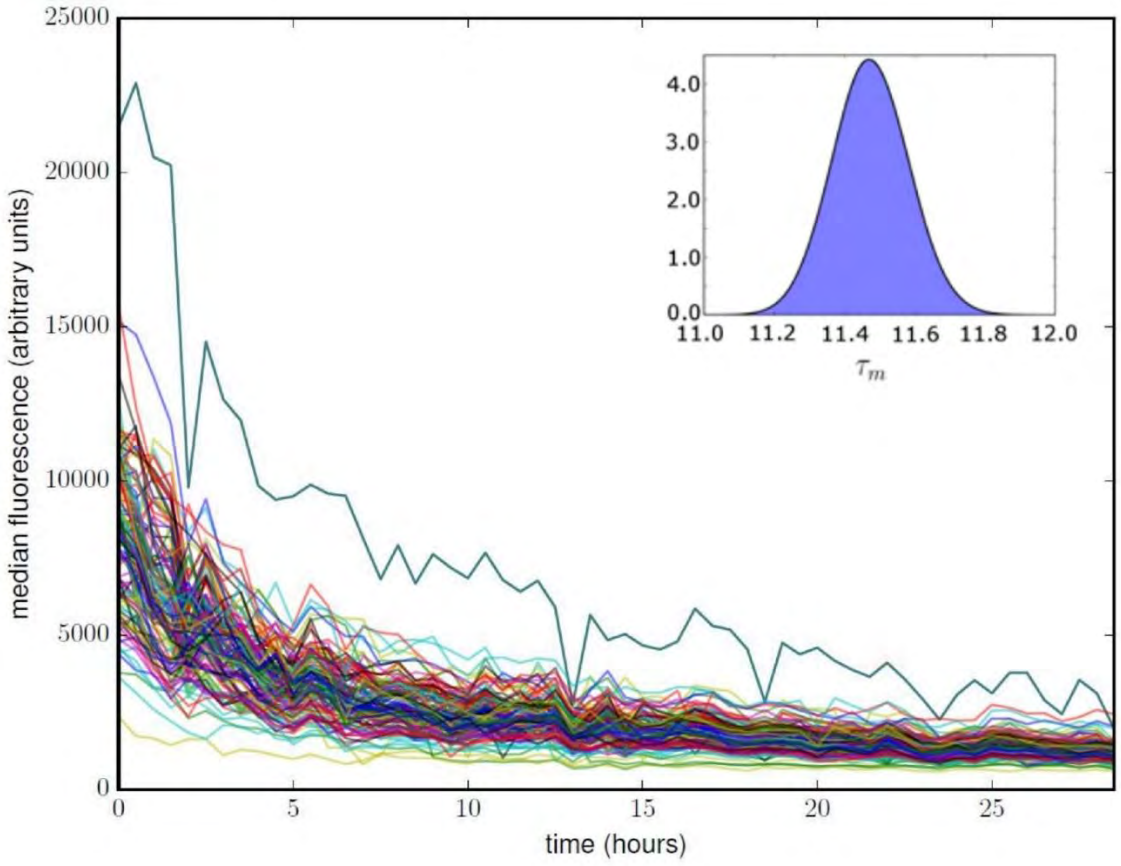



**Figure 2.6**: Using SpyoIPD-EGFP to study Pma1 temporal dynamics in single cells. Each trace in the plot corresponds to the total cellular fluorescence versus time for a single yeast cell, with t = 0 hours corresponding to the time of glucose addition. See the appendix for a detailed description of how these data are analyzed to calculate the in-membrane half-life of fluorescently labeled Pma1. Inset: The posterior probability for the half-life of Pma1, $\tau_m$, found by integrating the probability corresponding to the surface in Appendix Figure A5.

developed a novel Bayesian analysis that integrates the data across all cells to infer a half-life for Pma1 in the plasma membrane. Details of this analysis are given in the Appendix. We verified our Bayesian approach with a second *ad hoc* method that uses data from cells expressing Pma1-SpyTag and SpyoIPD-EGFP and also from cells expressing untagged Pma1 and SpyoIPD-EGFP. Comparing the change in fluorescence intensity over time in the membrane region of these two strains allowed us to extract the signal from Pma1 at the plasma membrane (see Appendix Figure A3). The two approaches give similar estimates of the in-membrane half-life of Pma1 - 11.5 and 10.2 hours respectively, agreeing with each other and the previously reported half-life of Pma1 of 11 hours, determined by cycloheximide inhibition of translation[93].

**Chapter 2.10 - Discussion**

Permanent labeling *via* SpyoIPD/SpyTag presents a minimally invasive method for fluorescently labeling proteins *in vivo*. I have shown that SpyoIPD/SpyTag can be used to label a variety of proteins in living yeast, and that labeling in this manner can be less disruptive to protein function than direct fusion to a fluorescent protein, presumably because of the smaller size of SpyTag. At first blush, this result may be surprising, as the labeled form of the target protein actually contains a larger modification than a fluorescent protein alone (ST-SC-FP vs FP). I believe SpyoIPD/SpyTag is less detrimental to protein function because of two features of the SpyoIPD/SpyTag imaging system. First, labeling proteins post-translationally using SpyoIPD allows proteins to properly fold and reach their final cellular destination before being labeled. If a FP tag is interfering with normal protein function and/or localization by interfering with a maturation step, SpyoIPD/SpyTag mediated labeling may therefore offer a less disruptive

method to achieve fluorescent labeling. Second, not every target protein in the cell is necessarily being labeled with SpyoIPD/SpyTag. Some POIs may be able to tolerate a portion of their population being labeled with no apparent effect on protein function, but be unable to function properly when all copies of the POI are labeled. This second explanation is supported by the fact that defects in the septin ring and in cell division are only observed when Cdc12-SpyTag is coexpressed with high amounts of SpyoIPD-EGFP (data not shown).

In its current form, SpyoIPD/SpyTag labeling presents a useful tool for researchers interested in labeling and tracking proteins of interest that poorly tolerate direct fusions to fluorescent proteins. Labs that study membrane proteins in particular, may find SpyoIPD/SpyTag labeling useful, as membrane proteins are notoriously difficult to label and undergo an extensive maturation process before reaching their final subcellular localization. Other POIS that undergo extensive maturation steps, or that are present in dense complexes may also benefit from SpyIPD/SpyTag labeling.

It is still unclear why SpyoIPD/SpyTag is more effective at labeling some target proteins than others. Protein concentration plays a significant role but cannot account for all the variation in labeling efficiency that is observed. For example, Vma1 is more than 25 times more abundant than Cdc12, and yet Cdc12 is much more readily visible when labeled with SpyoIPD-EGFP. Local concentrations most likely play some role in explaining this variation, as proteins with high local concentrations will produce a more intense signal that is more readily visible against background fluorescence. Variation in the accessibility of the C-terminal SpyTag may play some role as well. More extensive characterization of SpyoIPD-EGFP labeling of different target proteins needs to be

performed to gain a better understanding of any additional factors that may play a role in determining labeling effectiveness.

Because SpyoIPD and SpyTag bind irreversibly, our method presents a useful tool for studying a variety of time-dependent changes. Here we used SpyoIPD-EGFP to follow the turnover of Pma1, a plasma membrane protein, in individual cells. SpyoIPD/SpyTag could also be used to track a variety of other dynamic processes that occur in living cells, such as the accumulation of post-translational modifications, exchange of protein interaction partners, and the selective labeling of organelles, sub-cellular membraneless compartments and even entire cells in an age-dependent manner.

# Chapter 3: Improving SpyoIPD/SpyTag labeling by Nuclear Sequestration of Unreacted SpyoIPD-EGFP

## Chapter 3.1 - Introduction

The presence of unbound SpyoIPD-EGFP is a major limitation for SpyoIPD/SpyTag mediated labeling, as it both produces background fluorescence that obscures signal from the labeled target protein, and requires that SpyoIPD-EGFP expression levels be optimized for every target protein. Using glucose to inhibit new synthesis of SpyoIPD-EGFP solves some of these issues, presumably by allowing any remaining unbound SpyoIPD-EGFP to be cleared from the cytoplasm through a combination of degradation at the proteasome, reaction with unlabeled target protein, and dilution from cell growth. Clearance of unreacted SpyoIPD-EGFP can still take many hours however, especially if the concentration of SpyoIPD-EGFP is much greater than that of the target protein. I hypothesized that the rate at which unreacted SpyoIPD-EGFP is cleared from the cytoplasm could be accelerated if expression of a second SpyTagged protein was induced that localizes to some subcellular localization away from the POI. Specifically, I tested whether inducing the expression of a nuclear localized SpyTagged protein would accelerate the rate at which unreacted SpyoIPD-EGFP is cleared from the cell when labeling Pma1-SpyTag. In short, our strategy would be to first label Pma1-SpyTag with Pma1-SpyTag using the protocol outlined in Chapter 2. Following Pma1-SpyTag labeling, glucose would be added to the media to turn off expression of SpyoIPD-EGFP, and at the same time expression of a SpyTagged nuclear protein would

be turned on. Unreacted SpyoIPD-EGFP should then react with the nuclear SpyTagged protein, clearing SpyoIPD-EGFP from Pma1.

**Chapter 3.2 - Selecting an Inducible Promoter**

To be effective, the inducible promoter I use to drive expression of the nuclear SpyTagged protein needs to be minimally active in the absence of inducer and strongly active in its presence. The promoters that drive the expression of some HXT proteins (glucose transporter proteins) have been shown to be sensitive to glucose[94-96]. In particular the HXT3 promoter and HXT1 promoter have both been shown to be strongly induced by glucose. Using glucose to induce expression would have the added benefit that the same signal could be used to both turn off new expression of SpyoIPD-EGFP and turn on expression of the nuclear sequestering protein.

To test the glucose-responsiveness of the HXT promoters, yeast expression vectors were constructed expressing EGFP under control of the HXT3 promoter, a short fragment of the HXT1 promoter (HXT1s, 818 bp upstream of HXT1) and a longer fragment of the HXT1 promoter (HXT1L, corresponding to the 1312 bp upstream of HXT1). Yeast expressing these vectors were grown under the same media conditions used previously to label target proteins with SpyoIPD-EGFP; first overnight in noninducing media and then 8 hours in inducing media (containing galactose) with and without glucose. Induction was assessed based on bulk culture fluorescence measurements collected using a fluorimeter (Figure 3.1).

From these bulk fluorescence measurements, HXT1s and HXT1L are both strong candidates for inducing expression of the nuclear sequestering protein. Both promoters

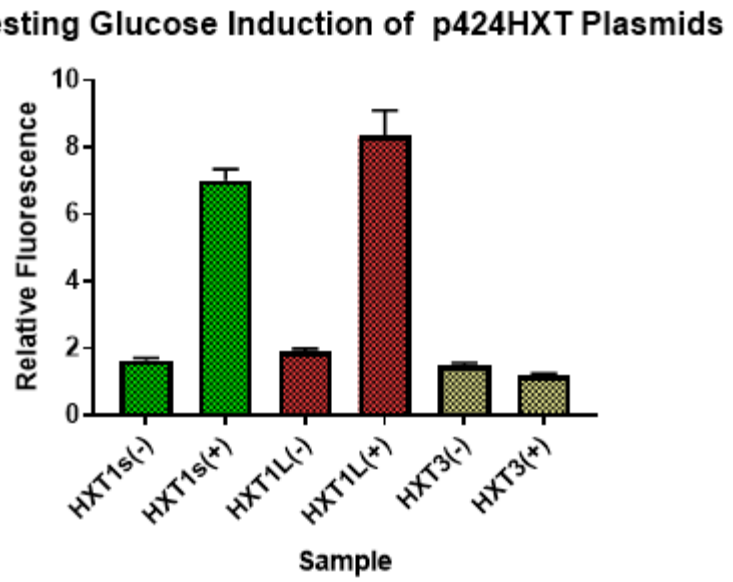



**Figure 3.1**: Testing the glucose responsiveness of the HXT1s, HXT1L, and HXT3 promoters in yeast. Yeast expressing EGFP from p424 HXT1s, p424 HXT1L, or p424 HXT3 were grown overnight in noninducing media (2% sucrose/1% raffinose), and diluted the next day into galactose-containing media with (+) and without glucose (-), and grown for an additional 8 hours, at which point bulk fluorescence measurements were collected using a fluorimeter. Shown above is the relative fluorescence of each strain, where fluorescence values are defined relative to a YPH499 yeast parent strain expressing no fluorescent proteins. Average fluorescence of 3 cultures shown, with error bars corresponding to the standard deviation of these 3 measurements.

are strongly induced by glucose, exhibiting 4.3 (HXT1s) and 4.4 (HXT1L) fold more fluorescence in the presence of glucose than its absence. Some promoter leakiness is observed for both promoters (1.6x and 1.9x background fluorescence in the absence of glucose), but is relatively low. Yeast expressing EGFP under control of the HXT3 promoter, in contrast, are actually less fluorescent in the presence of glucose than its absence (1.5x background without glucose vs 1.2x with). This behavior can be explained if expression from the HXT3 promoter is being induced by one of the other sugars present (raffinose, sucrose, or galactose). Previous experiments have shown that high glucose concentrations can decrease expression from the HXT3 promoter[94], which may explain the decreased fluorescence in the presence of glucose. Of the 3 promoters tested, HXT1s showed the best combination of strong glucose induction and minimal promoter leakiness, and was therefore selected to drive expression of the nuclear sequestering protein.

**Chapter 3.3 - Creation and Characterization a Nuclear Sequestering Protein**

To create a nuclear sequestering protein, I fused the Nuclear Localizing Sequence (NLS) from the SV40 virus onto the N-terminus of mCherry. To test the effects of the different tags on mCherry localization, NLS-mCherry, mCherry-SpyTag, NLS-mCherry-SpyTag, and untagged mCherry were expressed in yeast (see Figure 3.2). Cells expressing untagged mCherry display homogenous, evenly distributed fluorescence, indicating that mCherry does not naturally accumulate in any subcellular compartment. Yeast expressing NLS-mCherry, in contrast, display spheres of fluorescence roughly the size and shape of the nucleus, supporting nuclear localization. Weak cytoplasmic fluorescence is also observed, indicating that nuclear localization is not 100%. Yeast

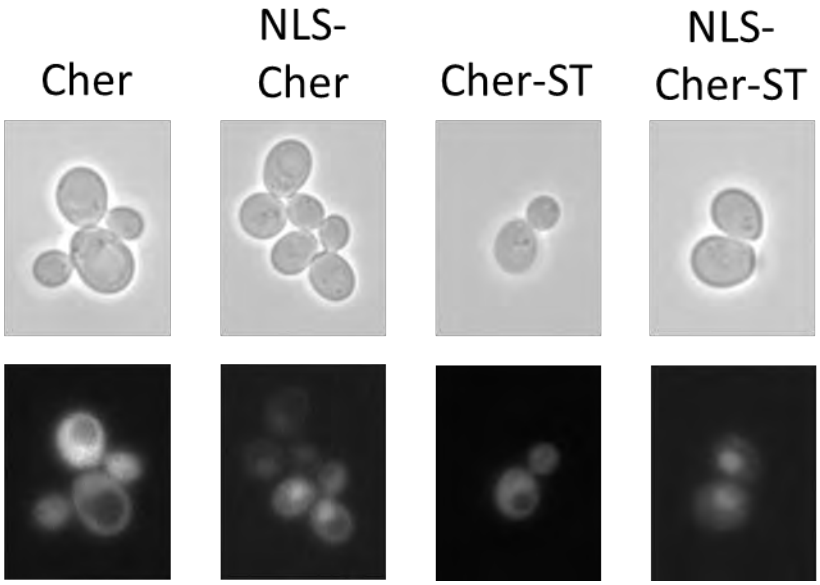



**Figure 3.2**: Testing the effects of SpyTag (ST) and Nuclear Localization Sequence (NLS) on protein localization. Yeast expressing untagged mCherry (Cher), mCherry tagged at the N-terminus with NLS (NLS-Cher), mCherry tagged at the C-terminus with SpyTag (Cher-ST), and mCherry tagged with both NLS and SpyTag (NLS-Cher-ST) were imaged. Top row: brightfield. Bottom row: fluorescence.

expressing mCherry-SpyTag and NLS-mCherry-SpyTag produce fluorescent images resembling those of untagged mCherry and NLS-mCherry respectively, indicating that the C-terminal SpyTag has no effect on mCherry localization.

**Chapter 3.4 - Testing SpyoIPD-EGFP reactivity with NLS-mCherry-SpyTag**

As an initial goal, I first sought to determine whether NLS-mCherry-SpyTag could react with SpyoIPD-EGFP and drive localization of SpyoIPD-EGFP into the nucleus. Yeast strains were constructed expressing SpyoIPD-EGFP, PMA1 with and without SpyTag, and NLS-mCherry with or without SpyTag. For these experiments, NLS-mCherry was expressed from the CUP1 promoter and both mCherry and SpyoIPD-EGFP were expressed at the same time. Yeast were grown overnight in non-inducing media, and diluted the next day into media containing galactose and $Cu^{2+}$ to induce expression of both mCherry and SpyoIPD-EGFP. Cells were imaged following 8 hours of induction (see Figure 3.3).

Cells expressing NLS-mCherry-SpyTag, SpyoIPD-EGFP and untagged Pma1 display GFP and mCherry fluorescence primarily in the nucleus, with some cytoplasmic mCherry fluorescence. Cells expressing Pma1-SpyTag, SpyoIPD-EGFP, and NLS-mCherry show GFP fluorescence primarily at the plasma membrane and cytoplasm, and mCherry signal primarily in the nucleus and cytoplasm, with little overlap between the two. Only in cells expressing SpyoIPD-EGFP with NLS-mCherry-SpyTag and Pma1-SpyTag display GFP fluorescence at both the nucleus and at the plasma membrane. From these experiments, SpyoIPD-EGFP is capable of labeling both NLS-mCherry-SpyTag and Pma1-SpyTag in the same cell, and reaction with NLS-mCherry-SpyTag

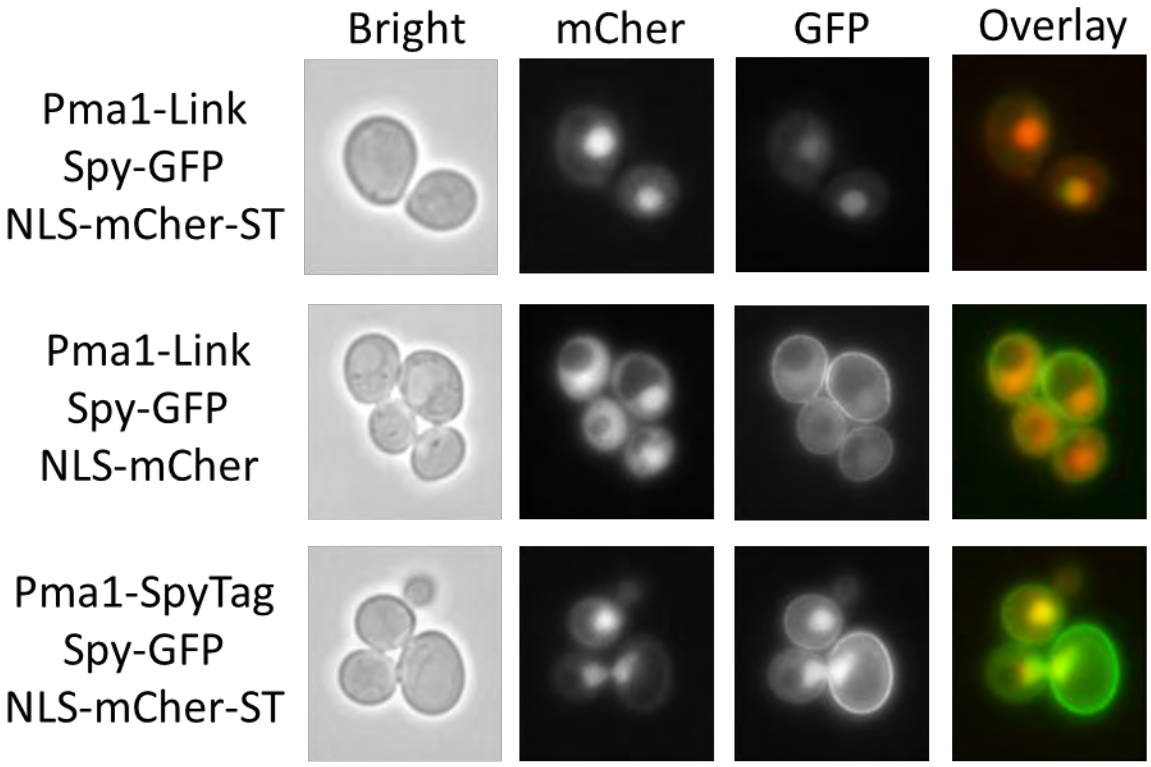



**Figure 3.3**: Testing NLS-mCherry-SpyTag nuclear sequestration of SpyoIPD-EGFP. Yeast expressing SpyoIPD-EGFP (Spy-GFP), Pma1 with SpyTag (Pma1-SpyTag) or without SpyTag (Pma1-Link), and NLS-mCherry with SpyTag (NLS-mCherry-ST) and without SpyTag (NLS-mCherry) expressed from pCu415CUP1 were imaged after 8 hours growth in media containing galactose and $Cu^{2+}$. Cells imaged under bright-field (Bright), with mCherry excitation and emission filters (mCher), and with GFP excitation and emission filters. Also shown is the overlay of GFP and mCherry fluorescent images (Overlay).

drives localization of SpyoIPD-EGFP to the nucleus. Interestingly, NLS-mCherry-SpyTag localizes more exclusively to the nucleus when labeled with SpyoIPD-EGFP. This could potentially be explained if reacting with SpyoIPD-EGFP increases the size of NLS-mCherry-SpyTag so that it can no longer passively diffuse through the nuclear pore complex.

**Chapter 3.5 - Using HXT1s to induce NLS-mCherry-SpyTag expression**

A high-copy number yeast plasmid was constructed containing NLS-mCherry-SpyTag under control of the HXT1s promoter. As an initial test to see whether NLS-mCherry-SpyTag can clear unreacted SpyoIPD-EGFP from cells faster, yeast expressing Pma1-SpyTag and SpyoIPD-EGFP with or without the above plasmid were grown overnight and diluted the next day into galactose containing media. After 8 hours of SpyoIPD-EGFP induction, glucose was added to the media, and samples were taken at regular time intervals for imaging (see Figure 3.4).

In the absence of glucose, expression of NLS-mCherry-SpyTag from the HXT1s promoter is minimal, and SpyoIPD-EGFP expression can be observed both at the plasma membrane and distributed evenly throughout the cytoplasm. Two hours after the addition of glucose, mCherry fluorescence is faintly visible in the nucleus, and some nuclear GFP signal is observed as well. Four hours after adding glucose, mCherry fluorescence is strong, and SpyoIPD-EGFP nuclear localization is significant.

From these images, qualitatively it appears that inducing expression of NLS-mCherry-SpyTag may increase the rate at which SpyoIPD-EGFP is cleared from the cytoplasm. To gain a more quantitative assessment of SpyoIPD-EGFP clearance rates,

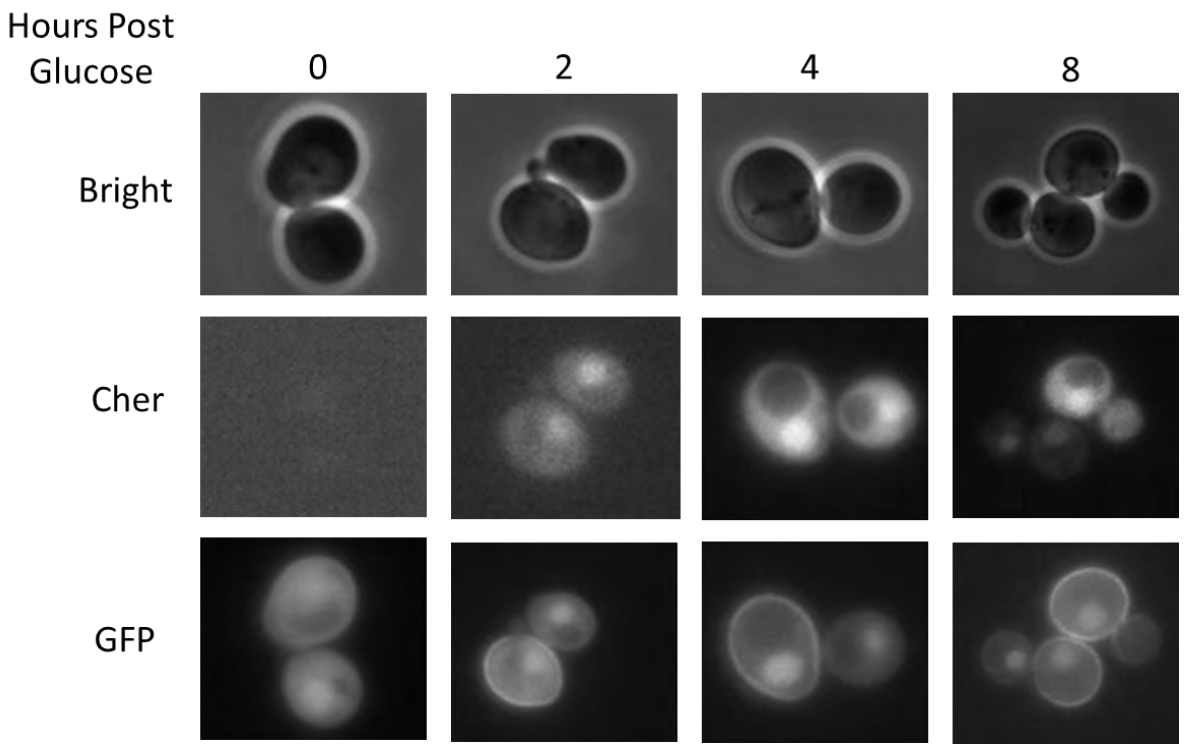



**Figure 3.4**: Testing NLS-mCherry-ST driven nuclear sequestration of SpyoIPD-EGFP. Yeast cells expressing SpyoIPD-EGFP and Pma1-SpyTag with HXT1s driven NLS-mCherry-SpyTag were grown overnight in noninducing media, diluted the next day into galactose containing media, and grown an additional 8 hours before adding glucose. Shown above are images following addition of glucose. GFP and Cher refer to GFP fluoresence and mCherry fluorescence channels.

we tracked cells using the yeast microdevice used previously to determine Pma1 half-life (see chapter 2.7). The ratio of plasma membrane to cytosol fluorescence was tracked in yeast expressing Pma1-SpyTag and SpyoIPD-EGFP with and without NLS-mChery-SpyTag (see Figure 3.5). Unexpectedly, inducing NLS-mCherry-SpyTag showed no effect on the time required to clear cytoplasmic SpyoIPD-EGFP, with the ratio of membrane/cytosol signal plateauing at approximately 9 hours in both strains.

**Chapter 3.6 - Fusing NLS directly to SpyCatcher-EGFP**

In one last effort to use nuclear sequestration to improve SpyCatcher-EGFP labeling, NLS was fused directly to the N-terminus of SpyCatcher-EGFP (note that SpyCatcher was used instead of SpyoIPD for these experiments. Results from imaging experiments described in chapter 4 found the two proteins to be equally effective in labeling target proteins). Theoretically, unbound NLS-SpyCatcher-EGFP should be at equilibrium between the nucleus and the cytoplasm, with the bulk of unreacted NLS-SpyCatcher-EGFP in the nucleus. Once bound, however, SpyCatcher-EGFP would be trapped in the subcellular location of the target protein. To test this idea, NLS-SpyCatcher-EGFP was expressed in cells also expressing a collection of SpyTagged target proteins, and cells were imaged after 8 hours of expressing NLS-SpyCatcher-EGFP. As can be seen in Figure 3.6, the NLS tag may improve target labeling slightly, but for the most part is negligible. For Cdc12-SpyTag and Pma1-SpyTag, the ratio of target signal to cytoplasmic background may be slightly improved. Htb2-SpyTag is labeled regardless of the NLS tag, and Vma1-SpyTag, while weakly labeled by SpyCatcher-EGFP, is not labeled at all by NLS-SpyCatcher-EGFP.

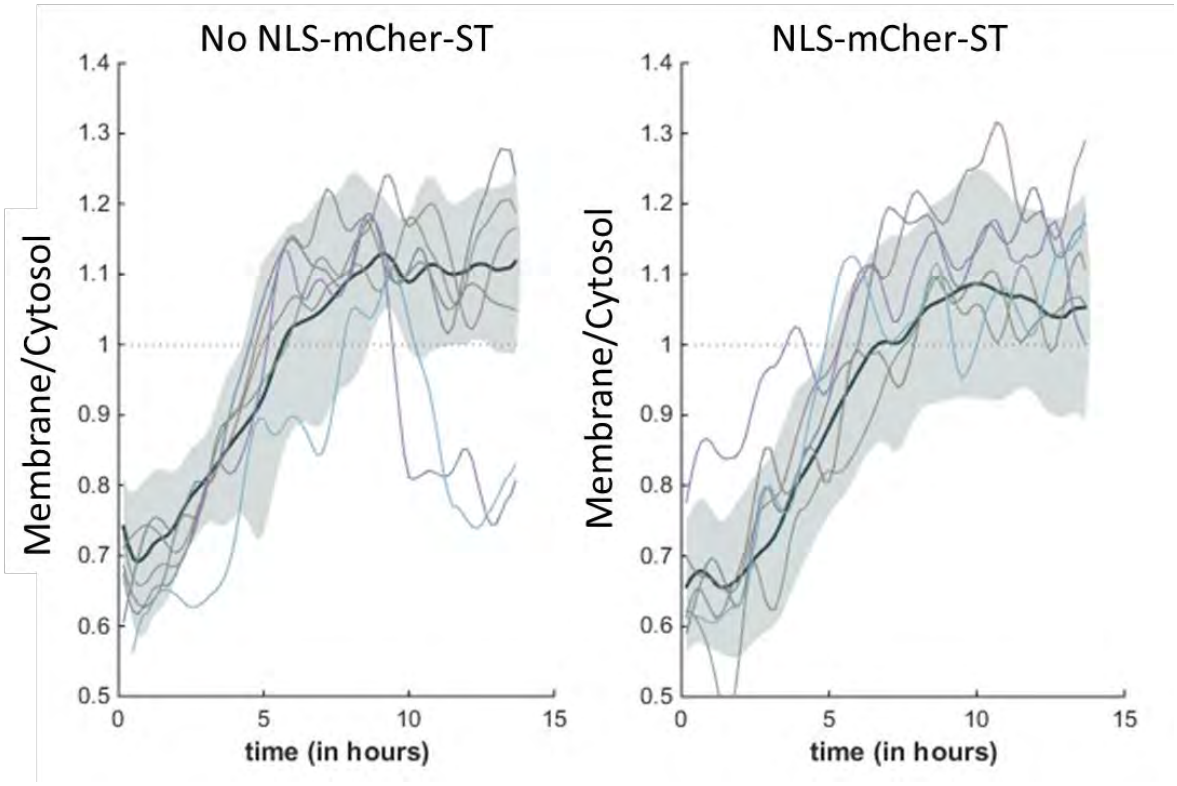



**Figure 3.5**: Testing the effect of NLS-mCherry-ST induction on SpyoIPD/SpyTag labeling of Pma1-SpyTag. Yeast expressing Pma1-SpyTag and SpyoIPD-EGFP without (No NLS-mCher-ST) and with NLS-mCherry-SpyTag (NLS-mCher-ST) were grown overnight in the presence of galactose, loaded into a yeast microdevice, and switched into glucose containing medium. Membrane/cytosol fluorescence tracked following switch to glucose.

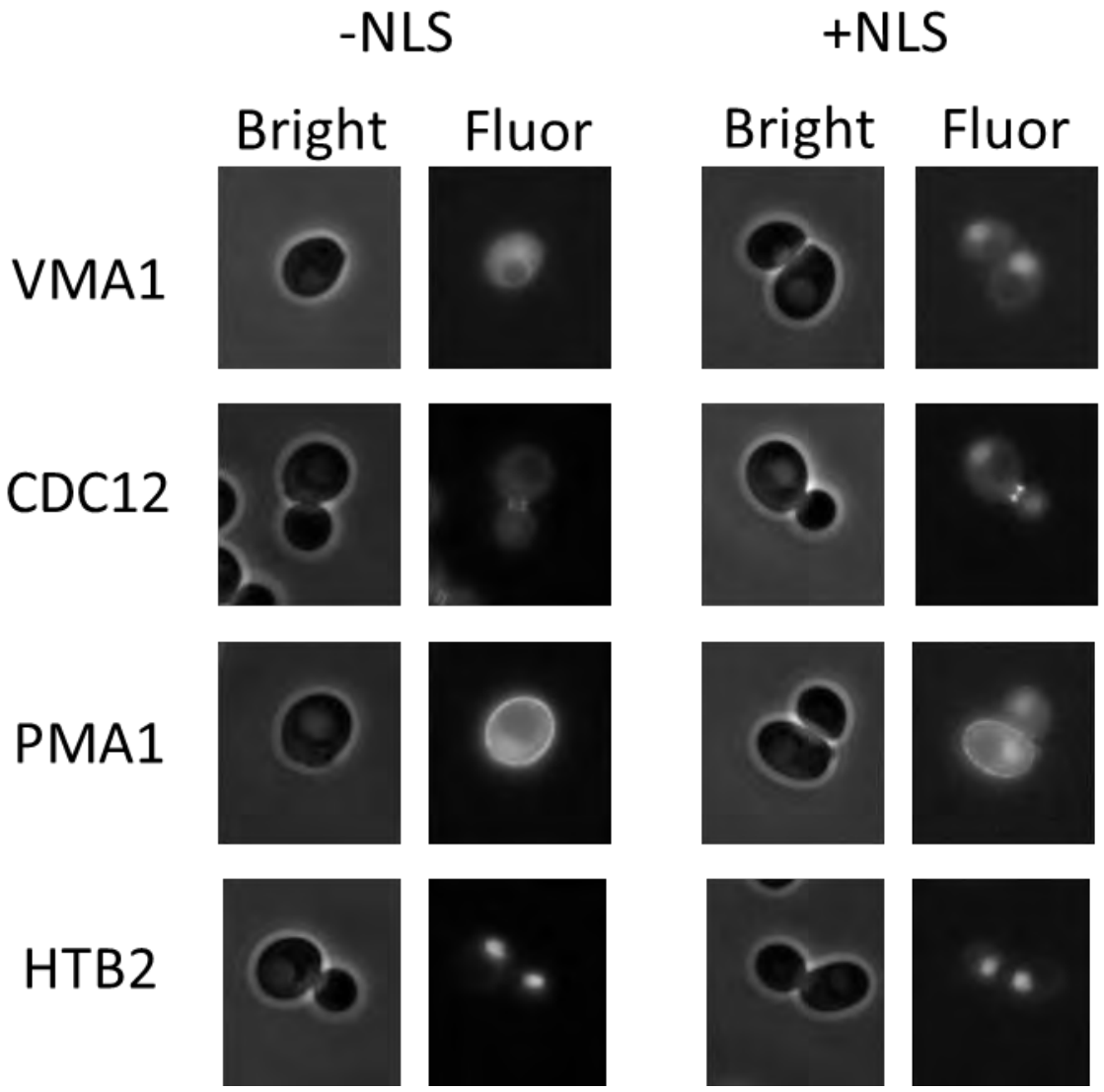



**Figure 3.6**: Testing the effect on target labeling of directly fusing NLS to SpyCatcher-GFP. Yeast expressing SpyCatcher-GFP without an N-terminal NLS (-NLS) and with an N-terminal NLS (+NLS) were grown overnight in the presence of a variety of target proteins tagged with SpyTag (target identities shown to the left). Images collected after 8 hours of growth.

## Chapter 3.7 Discussion

It's possible that the apparent lack of an effect of NLS-mCherry-SpyTag induction on SpyoIPD-EGFP clearance rates is somehow an artefact of image analysis, or that diffraction from nuclear fluorescence is artificially increasing the cytoplasmic fluorescence that is measured. Regardless, it appears that any effect of nuclear sequestration, if real, is most likely minimal. One potential explanation for this lack of an effect is if inducing NLS-mCherry-SpyTag from a high-copy number plasmid slows cell growth, slowing the rate at which SpyoIPD-EGFP is diluted due to cell division and negating any benefit from nuclear sequestration. It is also possible that expressing NLS-mCherry-SpyTag decreases the fraction of Pma1-SpyTag that is labeled, so that the net result on membrane/cytosol signal is negligible. While performing these experiments, further optimization of the glucose-pulse chase protocol was being performed in parallel (using lower concentrations of galactose, and diluting cells with fresh media to keep cultures in mid-log phase. The results presented in Chapter 2 were obtained using this optimized procedure). Using this optimized procedure, I was able to obtain strongly labeled Pma1-SpyTag within two hours of adding glucose, which is less time than is typically required to express significant amounts of NLS-mChery-SpyTag from the HXT1s promoter (~ 4 hours). I therefore took a break from our efforts to use nuclear sequestration to improve SpyoIPD-EGFP labeling. HXT1s induction of NLS-mCherry-SpyTag could still potentially be of some value, and some benefit may be observed in situations where SpyoIPD-EGFP concentrations are greatly in excess over those of target proteins.

Nuclear sequestration would be more effective at clearing unreacted SpyoIPD-EGFP if sequestration could be achieved faster. The induction strategy described above is fundamentally limited by the long amount of time required to express NLS-mCherry-SpyTag. A faster, and more powerful strategy would be to have nuclear SpyTagged protein already present in the cell during POI labeling, but unable to react with SpyoIPD-EGFP until signaled to do so. This could be achieved by creating caged versions of SpyTag that would only be released and free to react with SpyoIPD-EGFP in response to some light or small molecule stimulus. The LOV2 domain in particular offers an attractive tool for creating a light-releasable SpyTag, and could likely be achieved using a strategy similar to that used previously to create a light-inducible degron[97]. A light-inducible SpyTag could also be useful tool for studying protein temporal dynamics.

# Chapter 4: Testing Additional Covalent Protein-Peptide Interaction Pairs

## Chapter 4.1 - Imaging Pma1-SpyTag with SpyCatcher-EGFP

SpyoIPD(IA) was initially chosen for labeling target proteins because it produced a higher fraction of covalently labeled EGFP-SpyTag than SpyCatcher, presumably because it is more stably expressed in *S. cerevisiae*. If SpyCatcher is unstable *in vivo*, however, this could actually be an asset for protein labeling purposes, as unreacted SpyCatcher-GFP would theoretically be cleared from the cytoplasm faster than SpyoIPD-EGFP through proteasomal degradation. The SpyCatcher/SpyTag labeling reaction also may occur faster than SpyoIPD/SpyTag labeling, as SpyCatcher does not contain portions of the C-terminal β-strand, which has the potential to inhibit reaction with SpyTag.

To test the ability of SpyCatcher-EGFP to label target proteins *in vivo*, I created a strain of yeast expressing Pma1-SpyTag and SpyCatcher-EGFP under control of the GAL1 promoter and integrated at the GAL2 locus. SpyCatcher-EGFP was then induced with varying concentrations of galactose using the same induction procedure used previously to image target proteins with SpyoIPD-EGFP. As can be seen in Figure 4.1, SpyCatcher-EGFP is stably expressed in yeast, and produces fluorescence at the cell periphery when coexpressed with Pma1-SpyTag. The amount of fluorescent signal and degree of Pma1 labeling observed is surprisingly high given the weak SpyCatcher expression and activity that was observed previously in western blots (Figure 2.1). These seemingly incongruous results can be potentially reconciled however, if microscope detection of SpyCatcher-EGFP fluorescence is more sensitive than the western blot

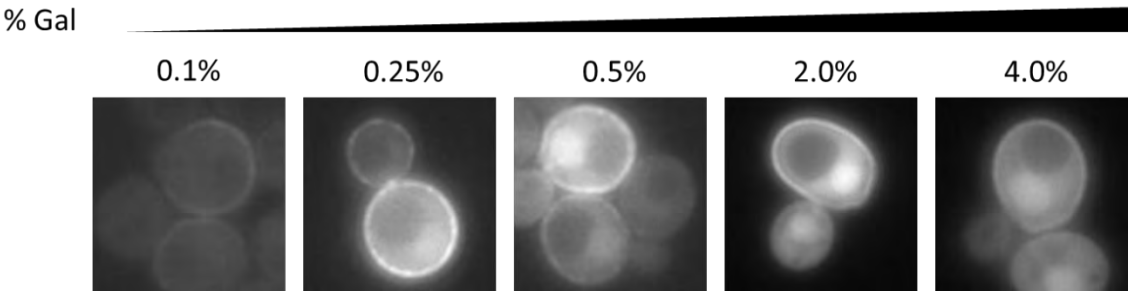



**Figure 4.1**: Imaging Pma1-SpyTag with SpyCatcher-EGFP. Yeast expressing SpyCatcher-EGFP under control of the GAL1 promoter, and integrated at the GAL2 genomic locus were induced with increasing concentrations of galactose (% w/v indicated above image). Yeast imaged following 8 hours of induction.

detection of SpyCatcher, or if fusion to EGFP stabilizes SpyCatcher expression. At the same concentration of galactose, fluorescence from cells expressing SpyCatcher-EGFP is significantly lower than that of cells expressing SpyoIPD-EGFP, but expression levels can be increased to the desired level by simply increasing the concentration of galactose. The ratio of membrane to cytosol fluorescence appears roughly similar between cells expressing SpyCatcher-EGFP and those expressing SpyoIPD-EGFP, and no significant difference between the two proteins is observed when both proteins are being expressed continuously (i.e. expression from GAL1 is not turned off). Some nuclear fluorescence is observed in cells expressing SpyCatcher-EGFP, as it was in cells expressing SpyoIPD-EGFP (see No Tag sample in Figure 2.3). It is unclear what drives this nuclear accumulation, although nuclear localization of free EGFP has been reported observed[88].

A glucose pulse-chase labeling experiment was also performed using SpyCatcher-EGFP to label Pma1-SpyTag and compared to cells labeled with SpyoIPD-EGFP induced with the same concentration of galactose (Figure 4.2). From these experiments, SpyCatcher-EGFP appear to be cleared faster from the cytoplasm of yeast cells, being almost cleared by t = 1.5 hours and gone by 3 hours, while cytoplasmic SpyoIPD-EGFP does not disappear until t = 5 hours. It is important to note that while the same concentration of galactose was used to induce both SpyoIPD-EGFP and SpyCatcher-EGFP, less SpyCatcher-EGFP is expressed at a given concentration of galactose, so it is possible that SpyCatcher-EGFP only appears to be cleared faster than SpyoIPD-EGFP because less of the protein is present at initial time points.

**Chapter 4.2 - Testing SpyCatcher Compatibility with Additional FP's**

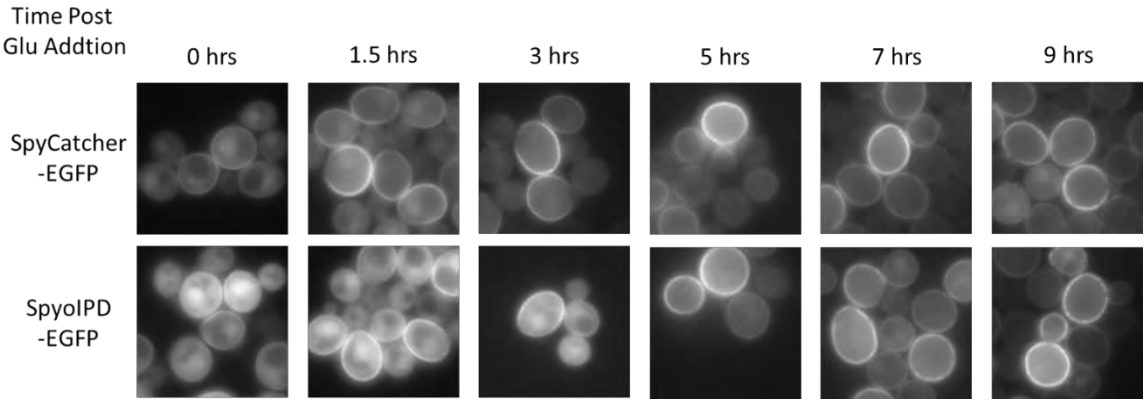



**Figure 4.2**: Glucose pulse-chase of SpyCatcher-EGFP vs SpyoIPD-EGFP. Yeast expressing Pma1-SpyTag and either SpyCatcher-EGFP (SpyCatcher-EGFP) or SpyoIPD-EGFP (SpyoIPD-EGFP) were grown for 8 hours in media containing galactose. Glucose was then added to 2% final concentration, and samples were taken for imaging at the above time points following addition of glucose.

One of the advantages of SpyoIPD/SpyTag based labeling over split-FP labeling is that SpyoIPD/SpyTag is readily compatible with any fluorescent protein. Creating split-FP pairs that efficiently bind one another and reconstitute a fluorescent protein requires extensive engineering and optimization, and split-mCherry and split-GFP are currently the only split-FP pairs that can be used to efficiently label target proteins *in vitro*[79, 98]. SpyoIPD/SpyTag labeling, in contrast, can theoretically be used in combination with any fluorescent protein without affecting labeling efficiency.

To test SpyoIPD compatibility with a range of fluorescent proteins, SpyoIPD was fused to CFP, Azami-Green, and mEOS3.2, and used to fluorescently label Htb2-SpyTag and Cdc12-SpyTag. Yeast were imaged under steady-state expression conditions, inducing with several low concentrations of galactose. As can be seen in Figure 4.3, both Htb2-SpyTag and Cdc12-SpyTag can be readily visualized using all 3 fluorescent proteins, and the identity of the fluorescent protein does not appear to affect the specificity with which target proteins are labeled. It is important to note that the identity of the fluorescent proteins did affect the amount of fluorescent signal observed, with some fluorescent proteins requiring higher concentrations of galactose to produce the same amount of labeling.

**Chapter 4.3 - Using SnoopCatcher/SnoopTag to image Candidate Proteins in *S. Cerevisiae***

SnoopCatcher/SnoopTag is protein/peptide isopeptide bond forming pair that is orthogonal to SpyCatcher/SpyTag, and offers a means to achieve multiprotein tracking *in vivo*[99]. To test SnoopCatcher's ability to label proteins in *S. cerevisiae*, I fused SnoopCatcher to EGFP and expressed the fusion protein from the GAL1 promoter in

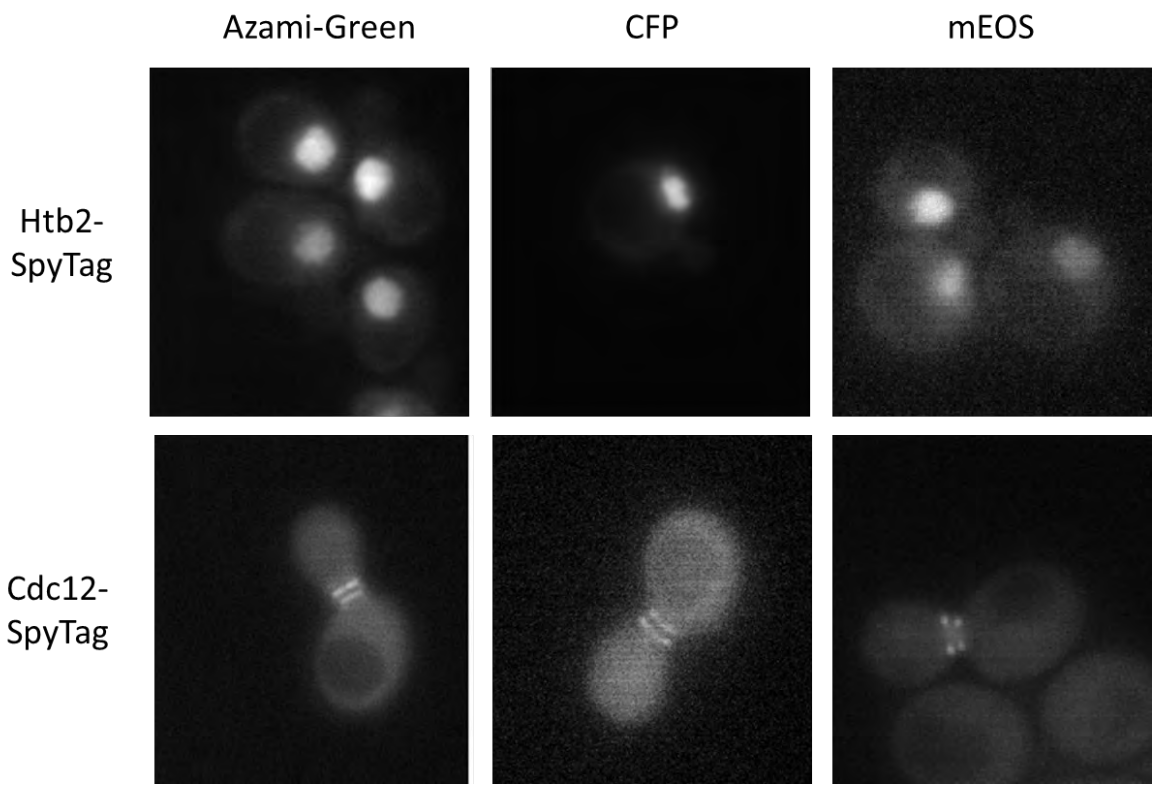



**Figure 4.3**: Testing SpyoIPD compatibility with additional fluorescent proteins. SpyoIPD fused to Azami-Green, CFP, or mEOS3.2 was expressed in cells also expressing Htb2-SpyTag or Cdc12-SpyTag. Cells were imaged after 8 hours of induction.

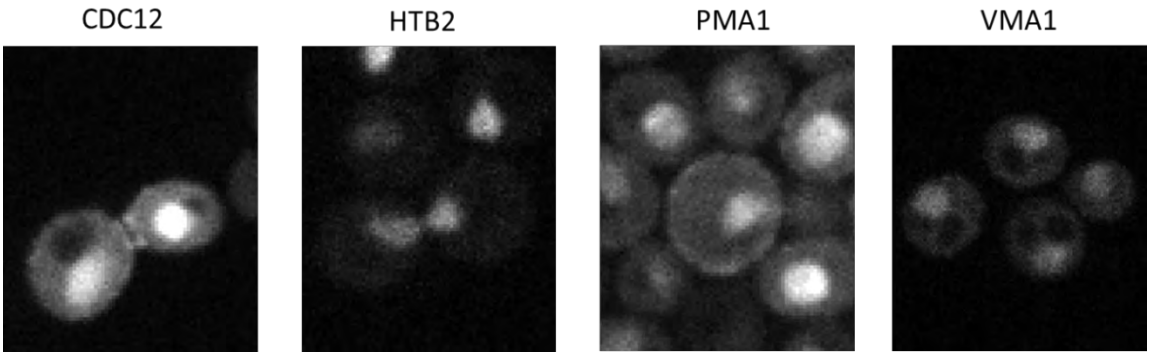



**Figure 4.4**: Testing SnoopCatcher-EGFP labeling of target proteins in *S. cerevisiae*. SnoopCatcher-EGFP was expressed from the GAL1 promoter in live yeast cells also expressing SnoopTagged versions of Cdc12, Htb2, Pma1, and Vma1. Cells imaged after 8 hours growth in galactose-containing medium.

cells also expressing a collection of SnoopTagged target proteins. When expressed continuously, SnoopCatcher-EGFP showed a similar ability to label target proteins as SpyoIPD-EGFP (Figure 4.4). Cdc12, Pma1, and Htb2 were all readily visualized by SnoopCatcher-EGFP, while Vma1 failed to be visualized, similar to the results observed previously using SpyoIPD-EGFP. Unexpectedly, SnoopCatcher-EGFP also exhibited significant nuclear fluorescence, regardless of the target protein being labeled. NLS search algorithms do not identify any potential NLS sequences within SnoopCatcher-EGFP, suggesting that SnoopCatcher-EGFP may be interacting with an endogenous nuclear protein.

**Chapter 4.4 - Discussion**

SpyCatcher-EGFP appears roughly similar to SpyoIPD-EGFP in ability to label proteins *in vivo*, with SpyCatcher-EGFP maybe being cleared faster from cells in glucose pulse-chase experiments. Regardless, any difference between the two, if real, is slight, and SpyCatcher-EGFP and SpyoIPD –EGFP can both be used to image proteins *in vivo* with similar results.

The demonstration that SpyCatcher is compatible with multiple FP's is a significant result, and a significant advantage over other protein-peptide interaction based labeling methods. SpyoIPD/SpyTag based imaging would also be greatly improved by the development of a second protein/peptide pair, which would enable multiprotein tracking. It is therefore unfortunate that SnoopCatcher/SnoopTag displays such strong nuclear fluorescence. It may be possible to decrease this nuclear accumulation by either increasing the size of SnoopCatcher-EGFP so that passive diffusion into the nucleus is no longer possible, or by attaching a nuclear export signal to SnoopCatcher-EGFP.

# Chapter 5: TRAP-FP Imaging

## Chapter 5.1 - Introduction

Tetratricopeptide Repeat Affinity Proteins (TRAPs) are designed proteins created in the Regan lab that bind 5 amino acid long C-terminal peptides with dissociation constants as low as 300 nM. Previous work in the Regan Lab has demonstrated that TRAPs can be used to fluorescently label a protein in *E. Coli*[72]. I sought to test whether TRAP-based imaging could be extended to *S. Cerevisiae*, which offers a wider selection of localized proteins for imaging.

## Chapter 5.2 - Design of a TRAP-FP system for imaging proteins in *S. cerevisiae*

I used the same basic strategy to image proteins in yeast with TRAP-FP that was used previously to visualize target proteins with SpyoIPD-EGFP (see Figure 2.2). Work in *E. Coli* has shown that TRAP-peptide pairs that bind with tighter affinities reduces the amount of TRAP-FP that is unbound, and produces clearer target labeling[72]. I therefore used the TRAP/peptide pair with the tightest binding affinity available, the TRAP 4/MEEVF binding interaction ($K_d$ = 300 nM, TRAP 4 referred to henceforth as simply TRAP). Using the same strategy as used previously for SpyoIPD/SpyTag based imaging, TRAP 4 was fused to mCherry and expressed using the GAL1 promoter from the GAL2 genomic locus. Target proteins were tagged at their genomic loci with a –MEEVF C-terminal peptide, connected to the target protein by a flexible GS linker. I initially used mCherry based on the logic that the longer excitation wavelengths should produce less cellular autofluroescence, although follow-up experiments using TRAP-GFP and TRAP-mEOS found that autofluorescence at shorter excitation wavelengths is not an issue. Target proteins were chosen using the same criteria used previously to select SpyoIPD

imaging candidates, selecting proteins that are highly abundant, localize to distinct regions of the cell, and which are not know to have an inaccessible C-terminal tags.

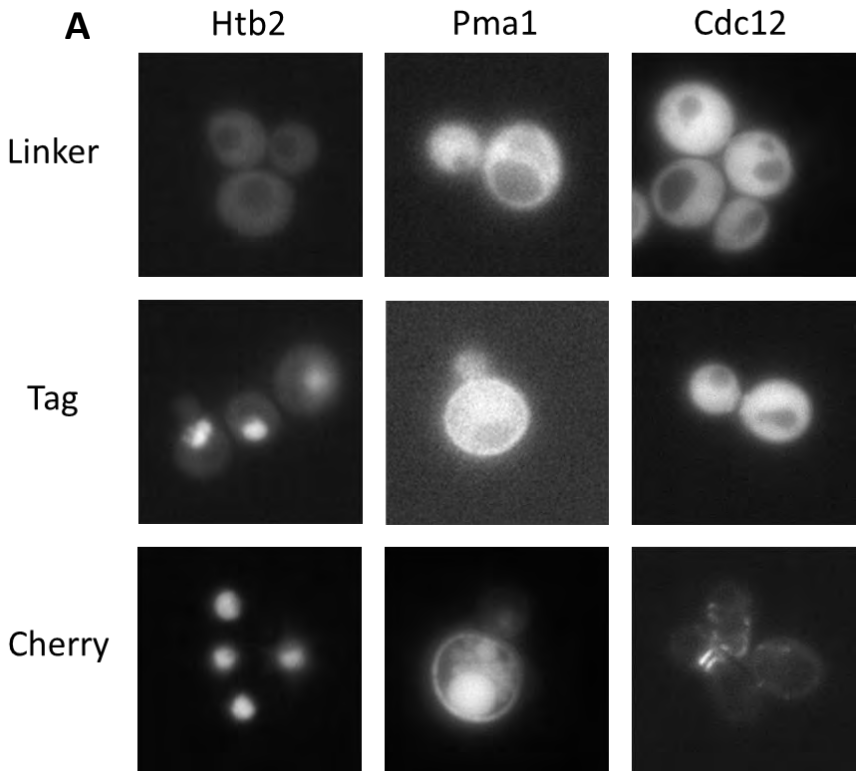
**Chapter 5.3 - Imaging proteins in Yeast**

TRAP-mCherry was coexpressed with a range of –MEEVF tagged proteins in *S. cerevisiae*. As can be seen in the Figure 5.1a, TRAP-mCherry is capable of visualizing several target proteins in yeast, producing localization patterns that are similar to those observed using SpyoIPD-EGFP for Pma1, Htb2, and Cdc12. Higher amounts of cytoplasmic signal are observed using TRAP-mCherry however, presumably because of the reversible and relatively weak binding interaction between the TRAP domain and its cognate peptide.

**Chapter 5.4 - Discussion**

I have shown that TRAP domains can be used to fluorescently label proteins in *S. cerevisiae*, although background fluorescence presents a major limitation, and greatly hinders its potential usefulness. Despite the handful of proteins that were successfully visualized using the TRAP domain, the majority of proteins tested in *S. cerevisiae* did not produce observable target labeling (see Figure 5.1b). As with SpoIPD/SpyTag labeling, it is unclear exactly why some proteins can be visualized with TRAP-FP, while others cannot. Likely the same issues discussed in Chapter 2.8 play a critical role here. TRAP based imaging has also been previously tested using TPR(MMY) and the –MEEVF tag ($K_d$ = 1.2 μM) in *S. pombe*, but no target proteins were successfully visualized[100].

TRAP-based imaging would be greatly aided by the development of a tighter binding TRAP/peptide pair, which would reduce background fluorescence from unbound




**B**

| Protein | Concentration (μM) | Localization | TRAP Signal |
|---|---|---|---|
| PMA1 | 22.60 | PM | Faint |
| VMA1 | 4.80 | VM | None |
| TUB4 | 0.15 | SP | None |
| HTB2 | 7.85 | N | Strong |
| CDC12 | 0.17 | BN | Faint |
| MRH1 | 2.96 | PM | None |
| VPH1 | 1.06 | VM | None |
| VMA2 | 3.03 | VM | None |

**Figure 5.1**: Imaging Target Proteins with TRAP-mCherry in *S. cerevisiae*. A) Fluorescent images of target proteins tagged at the C-terminus with a flexible GS linker (Linker) or the –MEEVF affinity tag (Tag), were coexpressed with TRAP-mCherry in live *S. cerevisiae*. Also shown are yeast expressing target proteins tagged directly with mCherry (Cherry). Identity of target proteins shown above images. Yeast imaged after 8 hours growth in galactose. B) Table summarizing the results of imaging a range of target proteins with TRAP-mCherry.

TRAP-FP. If issues of background fluorescence were resolved, the reversible nature of the TRAP/peptide interaction would provide a fluorescent labeling tool with several useful properties, such as an increased resistance to photobleaching (relative to a direct FP fusion), and possibly a novel route to achieving super-resolution imaging.

# Chapter 6: Methods

## Chapter 6.1 - General Protocols

### Molecular Biology

Insertion of linear DNA fragments into circular plasmids was performed either using a traditional restriction enzyme digest/DNA ligation strategy or using Circular Polymerase Extension Cloning (CPEC) following published procedures[101]. Unless stated otherwise, tags were attached to inserts through incorporation within PCR primers. Restriction enzymes, T4 DNA ligase, Phusion polymerase, and DpnI were purchased from New England Biolabs (NEB), and used with NEB supplied buffers. Digests were performed either for 3 hours or overnight at the temperature recommended by NEB. Ligations were performed for 2 hours or overnight at 16$^{o}$ C. To confirm gene insertion, 1 uL reaction was transformed into electrocompetent E. Coli and plated on selection media. All oligonucleotide synthesis and DNA sequencing was performed by the Keck Facility at Yale University. DNA purification was performed using Qiagen kits.

### Yeast Strain Manipulation

Unless noted otherwise, standard techniques and growth media were used for cultivating and genetically manipulating yeast strains[102]. Dropout media was prepared using purchased amino acid dropout mixes (Clontech). All experiments were performed in the parent yeast strain MHY2587 (an Ade+ variant of YPH499). To insert PCR products into the yeast genome, oligos were designed to amplify the desired sequence and attach 45 bp homology arms to either end. Genomic transformations were performed using a LiAc high-efficiency transformation (a modified version of the protocol found in

Geitz et al[103]. Following 2-3 days of growth on selection media, individual transformant colonies were picked into selection media and grown overnight at 30° C. The next day, these overnight cultures were lysed, and their genomic DNA purified using a glass bead lysis adapted from[104]. The resulting purified genomic DNA was screened for proper incorporation of the desired construct using oligonucleotides that anneal outside the targeted insertion area. The resulting PCR products were checked by size, and sequenced.

**Chapter 6.2 - Methods from Chapter 2**

**Molecular Biology**

*Constructs for testing the in vivo activity of SpyCatcher, SpyoIPD(IA), and SpyoIPD(IVML):* A plasmid containing the original SpyCatcher construct[83] was purchased from Addgene (Addgene plasmid #35044), monomeric EGFP[13] was amplified from the Regan lab vector pPROEX HTa M EGFP-MEEVD (pPROEX HTa M is a modified version of the pPROEX HTa vector (Invitrogen)) and mCherry was amplified from pNAS1b (Addgene plasmid # 61968[32]). SpyoIPD(IA) and SpyoIPD(IVML) were generated in the Schwarz-Linek lab, by site-directed mutagenesis of FbaB-CnaB2-Asp556Ala[82]. SpyoIPD(IA) was created by introducing the Ile552Ala mutation and SpyoIPD (IVML) was created by introducing Ile552Val and Met554Leu point mutations (introduced point mutations are highlighted in red in the below sequence overlays). The SpyCatcher protein contains the point mutations Ile473Glu and Met508Tyr, which were omitted from the SpyoIPD designs (highlighted in red below). The SpyCatcher protein also contains an additional 20 residues N-terminal residues from FbaB that were omitted from SpyoIPD, and 2 residues at the C-terminus (Arg-Ser) that are not present in the

native FbaB sequence, and which were also omitted from the SpyoIPD constructs (residues highlighted in yellow below).

```
SpyCatcher
MRLSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIK
SpyoIPD(IA)      --MSYYHHHHHHDCDIPTTENLYFQGAMV----------------------
DSATHIK
SpyoIPD(IVML)    --MSYYHHHHHHDCDIPTTENLYFQGAMV----------------------
DSATHIK

SpyCatcher
FSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVAT
SpyoIPD(IA)
FSKRDIDGKELAGATMELRDSSGKTISTWISDGQVKDFYLMPGKYTFVETAAPDGYEVAT
SpyoIPD(IVML)
FSKRDIDGKELAGATMELRDSSGKTISTWISDGQVKDFYLMPGKYTFVETAAPDGYEVAT

SpyCatcher       AITFTVNEQGQVTVNGKATKGDAHIRS---
SpyoIPD(IA)      AITFTVNEQGQVTVNGKATKGDAHAVMVAA
SpyoIPD(IVML)    AITFTVNEQGQVTVNGKATKGDAHVVLVAA
```

Sequence overlay of $His_6$ tagged SpyCatcher, SpyoIPD(IA) and SpyoIPD(IVML).

$His_6$ tagged SpyCatcher and SpyoIPD constructs were amplified from their bacterial expression vectors and inserted into p424 GAL1 by CPEC. P424 GAL1 is a yeast shuttle vector containing the strong galactose-inducible promoter GAL1, the high copy number 2μ replication origin, and the TRP1 selection marker[105]. DNA encoding EGFP was tagged at the 5' end with a sequence coding for the V5 epitope, and the 3' end with DNA coding for SpyTag, and inserted into pCu415CUP1 by CPEC[106]. pCu415CUP1 is a yeast shuttle vector that contains the intermediate strength, copper-inducible CUP1 promoter[107], a low copy number CEN replication origin, and a LEU2 selectable marker.

*Constructs for testing SpyoIPD/ST imaging in* S. cerevisiae: SpyoIPD was attached to the N-terminus of EGFP *via* an 8-residue linker (GGSGSGLQ), and inserted into p424 GAL1 by CPEC. The GAL1 promoter, SpyoIPD-EGFP, and CYC1 terminator ($CYC1_T$)

were then amplified from this vector and inserted using CPEC into pFA6a-HIS3MX6[108], a yeast insertion vector that contains a HIS3 selectable marker. To create a template plasmid that could be used to fuse a flexible linker peptide onto the 3'end of the genomic sequences of target proteins, oligonucleotides were used to amplify $CYC1_T$ from p424 GAL1 and attach a linker (GGSGSGLQ) upstream of $CYC1_T$. This fragment was inserted using CPEC into pFA6a-KANMX6[108], a yeast insertion vector with a kanamycin selectable marker (KanR). This construct was used as the basis for a template plasmid that could be used to tag the C-terminus of target proteins with linker-SpyTag (GGSGSGLQAHIVMVDAYKPTK). A sequence encoding the flexible linker and a portion of the pFA6a-KANMX6 vector backbone upstream of the flexible linker was amplified. In the process of amplification, SpyTag was fused to the flexible linker. This PCR product was then inserted back into pFA6a-Link-KANMX6 to create pFA6a-SpyTag-KANMX6 SpyTag. mCherry was inserted into pFA6a-KanMX6 using the same strategy as for pFA6a-Link-KanMX6.

**SpyoIPD-EGFP Sequence:**

MSYYHHHHHHDCDIPTTENLYFQGAMVDSATHIKFSKRDIDGKELAGATMELR
DSSGKTISTWISDGQVKDFYLMPGKYTFVETAAPDGYEVATAITFTVNEQGQVT
VNGKATKGDAHAVMVAAGGSGSGLQSKGEELFTGVVPILVELDGDVNGHKFSV
SGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDF
FKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG
HKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGP
VLSPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

**Linker:** GGSGSGLQ

**Linker-SpyTag:** GGSGSGLQAHIVMVDAYKPTK

Sequences of the SpyoIPD-EGFP labeling protein and Linker and Linker-SpyTag C-terminal fusions attached to target proteins for visualization. Highlighted in magenta is SpyoIPD, in yellow is the flexible linker, and in green is EGFP.

**Characterization of SpyCatcher and the SpyoIPD variants**

NMR spectra and DSF traces were collected in the Schwarz-Linek lab using the below protocols.

*NMR*: Uniformly $^{15}$N-labelled samples of SpyCatcher and SpyoIPD were produced and purified using established protocols[82]. NMR samples typically contained 0.1 mM protein in phosphate buffered saline, pH 7.4, 2% (v/v) $D_2O$. $^{1}$H-$^{15}$N HSQC spectra were recorded on a Bruker Ascend 700 MHz spectrometer equipped with a Prodigy TCI probe at 22 °C. Spectra were processed with NMRPipe[109] and analyzed with CCPN Analysis 2[110].

*Differential Scanning Fluorimetry*: Fluorescence at 570 nm (excitation 480 nm) of SYPRO© Orange in the presence of 5 μM protein was recorded using a real-time PCR instrument. The samples were heated at a rate of 1 °C per minute, between 25-95 °C. These denaturation transitions are irreversible, so it is inappropriate to calculate a Tm. We show the raw data.

**Constructing Yeast Strains**

Target proteins were tagged at the C-terminus with DNA encoding linker-SpyTag or linker alone in the yeast strain MHY2587 (an Ade+ variant of YPH499) by amplifying the desired tag and the KanR selectable marker from the appropriate template vector. During amplification, 45 bp overhangs were attached on either end corresponding to the final 45 bp of the target protein (attached to the 5' end) and the 45 bp immediately following the stop codon (to the 3' end). To insert SpyoIPD-EGFP, the GAL1 promoter, SpyoIPD-EGFP, $CYC1_T$, and HIS3 marker were amplified from pFA6His3MX6 – SpyoIPD-EGFP, attaching 45 homology arms in the process. Homology arms were

designed to match a 45 bp sequence 700 bp upstream of the GAL2 gene, and 45 downstream of the GAL2 stop codon.

Yeast strains expressing target proteins fused at the C-terminus to EGFP were obtained from the Yeast EGFP Clone Collection (Thermo Fisher), originally created and described in Huh et al[111]. Note that there are some differences between the EGFP directly fused to the target proteins in this library, and the EGFP fused to SpyoIPD, the most notable being the lack of the monomerizing A206K mutation in the yeast GFP library.

**Western Blot Analysis:**

To assess the *in vivo* activity of SpyCatcher and the SpyoIPD variants, yeast colonies were picked and grown overnight in synthetic defined media ($Leu^-$/ $Trp^-$) containing 0.1% glucose, 2% galactose, and 100 μM $CuSO_4$. The next day, cultures were diluted to an $OD_{600}$ of 0.2 into fresh selection media containing 2% galactose and 100 μM $CuSO_4$, and grown to an $OD_{600}$ between 1.0 and 2.0 (usually about 20 hours at 30 ºC). At this point, 10 $OD_{600}$ equivalents were pelleted, washed once with $H_2O$, and stored at -80 ºC for later analysis.

Yeast pellets (from 10 $OD_{600}$ equivalents) were lysed using the alkali lysis procedure[112] and final pellets were resuspended in 50 μL of 1xSDS-PAGE buffer. Lysate (10 μL) was loaded onto 10% or 15% SDS-PAGE gels, transferred to nitrocellulose, and probed using appropriate primary antibodies. Mouse anti-$His_6$ (GenScript (Piscataway, NJ), Cat. # A00186-100), and mouse antiV5 (Invitrogen (Carlsbad, CA), Cat. # 46-0705) primary antibodies were each diluted for use 1:1,000 in Tris Buffered Saline with 0.1%

Tween (TBST) and 5% w/v nonfat dry milk. For all immunoblots, the secondary antibody used was sheep anti-mouse IgG (diluted 1:10,000 in 5% milk/TBST, GE (Little Chalfont, UK), Product code NXA931) conjugated to horse radish peroxidase. Immunoblots were visualized by enhanced chemiluminscence, using Clarity$^{TM}$ ECL Western Blotting substrate (BioRad) and imaged using a GBox - Chemi 16 Bio Imaging System (Syngene).

**Microscopy**

For imaging experiments, single colonies were picked and grown overnight in non-inducing $His^-/G418^+$ synthetic defined media (2% sucrose/1% raffinose). The next day, overnight cultures were diluted to an $OD_{600}$ of 0.05 into fresh $His^-/G418^+$ synthetic defined media (2% sucrose/1% raffinose), supplemented with the desired concentration of galactose. Cultures were grown 8 hours before imaging.

For pulse-chase experiments, glucose was added to a final concentration of 2% w/v after 8 hours of induction to turn off protein expression from the GAL1 promoter. The $OD_{600}$ was sampled at regular intervals following glucose addition, and kept below 2.0 throughout the experiment by diluting with prewarmed media (that exactly matched the original growth media).

Fluorescent images were collected using Olympus IX-71 microscope with a 100× 1.4 NA Plan Apo lens (Olympus) and a CSU-X1 (Andor Technology) confocal spinning-disk confocal system equipped with an iXON-EMCCD camera (Andor Technology).

**Microfluidics Experiments**

All microfluidic experiments were performed in the Swain lab at the University of Edinburgh, by either Manuel Lenz, Lynne Regan, or Elco Baker. Microfluidic devices were fabricated with polydimethylsiloxane using standard techniques[92]. Single colonies were inoculated into a liquid culture of synthetic complete media (2% raffinose, 0.5% galactose) and grown overnight at 30 ºC. The following day, cells were loaded into a pre-warmed (30 ºC) microfluidic device and incubated in the synthetic complete media (2% raffinose, 0.25% galactose) for 1h before switching to glucose (0.1%). Throughout the experiment, the device was perfused with fresh media at a flow rate of 4 µl/min, controlled by syringe pumps (World Precision Instruments), and temperature was maintained at a constant environment of 30 ºC using a temperature controlled incubation chamber (Okolabs).

Time-lapse image acquisition was performed on a Nikon Eclipse Ti inverted microscope, with a 60X 1.4 NA oil immersion objective (Nikon). The experiment was controlled using a custom Matlab script (Mathworks) written for Micromanager[113]. Images were taken in bright-field and fluorescence, using a filter set appropriate for EGFP. Exposure intensities (LED lamp, 4V), exposure times (30ms) and imaging intervals ($0.5h^{-1}$) were set to avoid photobleaching. Data analysis was performed using image segmentation, cell tracking and data extraction using custom Matlab script[92]. To determine the ratio of membrane to cytosol signal, median membrane pixel intensities for each cell were extracted from images using the cell outline generated during cell identification.

**Photobleaching**

We investigated the possible contribution of photobleaching to fluorescence decay by comparing the fluorescence of cells irradiated multiple times at each time-point to cells irradiated once at each time point. We observed no significant difference in the cellular fluorescence over time between the two sets of cells, indicating that photo-bleaching does not contribute significantly to the fluorescent decay observed (See Appendix Figure 4).

**Chapter 6.3 - Methods from Chapter 3**

**Molecular Biology**

*Glucose inducible promoters:* Oligonucleotides were used to amplify the 818 bp (short HXT1 or HXT1s promoter) and 1312 bp (Long HXT1 or HXT1L immediately upstream of the HXT1 gene, and the 2033 bp immediately upstream of the HXT3 gene. Both short and long HXT1 promoters were inserted into the p424 GAL1 plasmid using SacI and BamHI (deleting the GAL1 promoter in the process), and HXT3 was inserted into the p424 GAL1 plasmid using the SacI and SmaI restriction sites (also deleting GAL1). For the purpose of testing glucose induction, EGFP was inserted into the above constructs using SmaI and SalI.

*NLS tagged proteins*: MCherry alone, NLS-mCherry, mCherry-SpyTag, and NLS-mCherry-SpyTag were made by incorporating the above tags into oligonucleotides, and inserted into the p424 HXT1s plasmid using SpeI and BamHI restriction sites. NLS sequence derived from the SV40 NLS (Biobrick part BBa_J63008). Both NLS and SpyTag sequences codon optimized for yeast expression. NLS was attached to SpyCatcher-GFP during PCR amplification from the pFA6a-SCGFP-HIS3MX6 plasmid,

and then the resulting product was inserted back into pFA6a-SCGFP-HIS3MX6 plasmid to create pFA6a-NLS-SCGFP-HIS3MX6.

**Yeast Induction**

To test the glucose responsiveness of the p424 HXT plasmids, yeast strains expressing EGFP inserted into p424 HXT1s, p424 HXT1L, and p424 HXT3 were grown overnight in noninducing selection media ($Trp^-$). The next day, overnight cultures were diluted into media supplemented with 1% galactose, with and without 4% glucose. Bulk fluorescence of cell cultures was collected following 8 hours of growth in this induction media. NLS-mCherry-SpyTag expression was also induced from HXT1s using 4% glucose. For induction from the CUP1 promoter, the same procedure was followed, but inducing with 100 μM $CuSO_4$.

**Microscopy**

Cells were analyzed by fluorescence microscopy using an inverted Nikon Eclipse Ti-S microscope equipped with a GFP filter cube set (49002-ET-EGFP [FITC/Cy2], Chroma Technology). Images were collected by using a Nikon Plan Fluor 100×, 1.3 numerical aperture oil objective, and an Andor Zyla 5.5 sCMOS camera.

**Microfluidics Experiments**

Unless noted otherwise, the same equipment and image analysis procedures were used performed as described in Chapter 2. Yeast expressing Pma1-SpyTag, SpyoIPD-EGFP, and NLS-mCherry-SpyTag from p424HXT1s were grown for 5 hours in media containing 1% galactose before injecting into the yeast microdevice. After approximately 1 hour in the microdevice, the media was switched to 4% glucose. To account for

increased nuclear fluorescence, we defined the 25 brightest pixels in the cytosol as the nucleus and subsequently excluded them from the analysis.

**Chapter 6.4 - Methods from Chapter 4**

**Molecular Biology**

SpyCatcher-GFP was created by amplifying SpyCatcher from the bacterial expression vector, and inserting into the SpyoIPD-GFP pfa6His3 plasmid, removing SpyoIPD in the process and fusing to GFP. SnoopCatcher-GFP was created using the same strategy, also inserting into SpyoIPD-GFP. Linker-SnoopTag was inserted into pfa6Kan using the same strategy used to make Linker-SpyTag.

Monomeric AzamiGreen (V123T, Y188A, F190K) was amplified from pRSFDuet-[114], mCherry from pNAS1b, mEOS3.2 from pNAS1B_B[72]. Sequences encoding the fluorescent proteins were amplified and inserted into the pFA6a-SCGFP-HIS3MX6 vector so as to replace GFP with the desired FP.

**Chapter 6.5 - Methods from Chapter 5**

**Molecular Biology**

To make a construct to use as a template for attaching linker-MEEVF onto genomic DNA targets, CYC1T was amplified from p424GAL1, attaching a flexible linker and a sequence coding for MEEVF in the process (codon optimized for yeast expression). This resulting product was then inserted into the pFA6KANMX6 destination vector using CPEC. To generate TRAP-mCherry, mCherry was amplified from pNAS1b, and inserted into a modified pNAS1b vector containing TRAP 4-mEOS (pNAS1B_B[72]) using CPEC, so as to replace mEOS with mCherry. TRAP 4-mCherry

was amplified from this vector and inserted into p424GAL1 by CPEC, and the GAL1-TRAP-mCherry-$CYC1_T$ cassette was amplified from this vector and inserted into pfa6His3MX6 using BamHI and SmaI restriction sites.

# Appendix

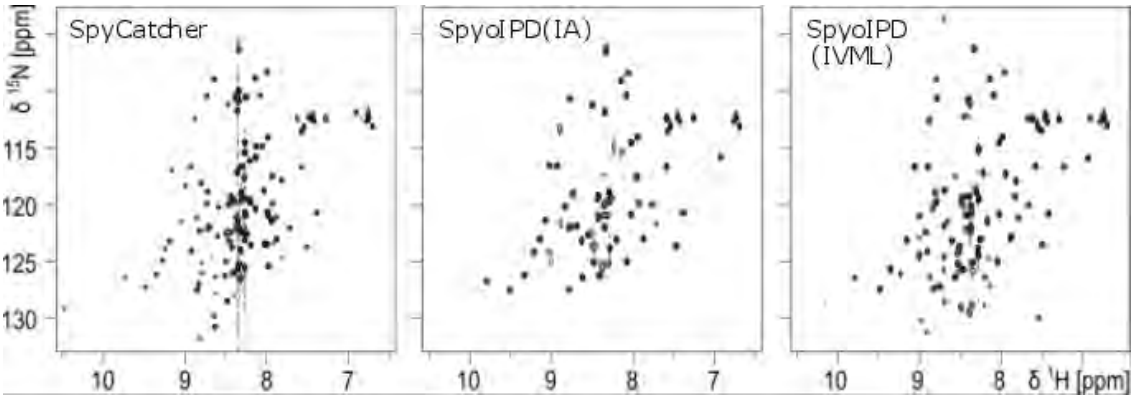



**Figure A1**. $^{1}H$-$^{15}N$ HSQC NMR spectra of SpyCatcher, SpyoIPD(IA) and SpyoIPD(IVML), collected at room temperature and in the absence of SpyTag.

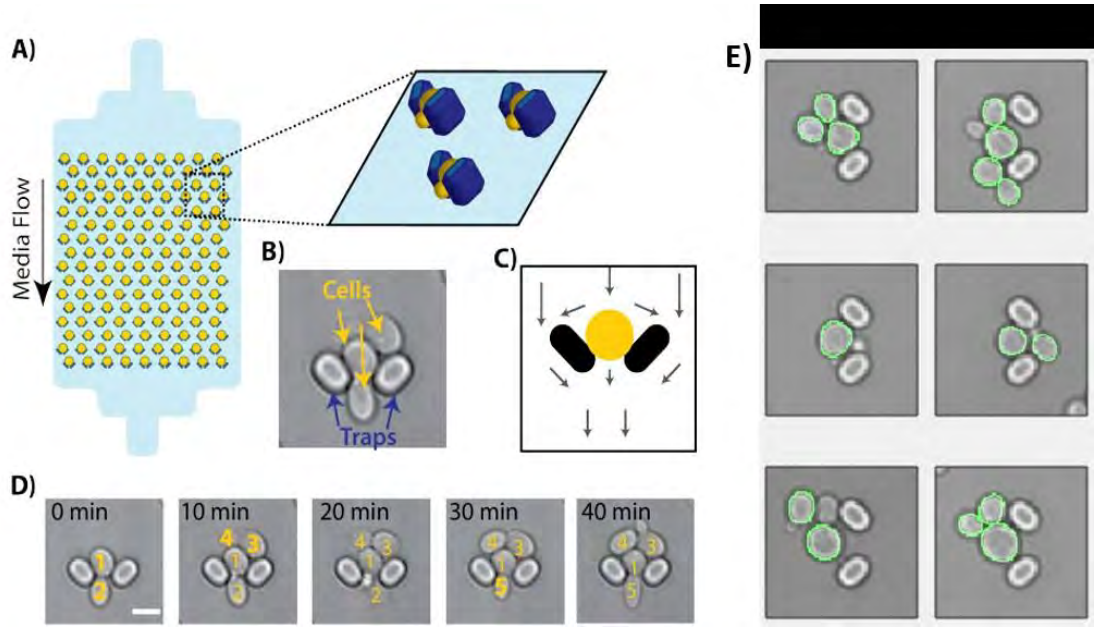



**Figure A2:** Layout of a single chamber of a microfluidic device and cell detection. A) 3D-graphic of microfluidic device with pillars (blue) and trapped yeast cells (yellow). B) An image of several yeast cells trapped in a microfluidic device. C) A schematic view of media flow (indicated by arrows) which creates specific physical constraints. Via media flow cells are kept in place and buds are being washed away. D) Time-lapse images of a single trap show the appearance of new cells washed in from above and the disappearance of daughters that are washed away after birth. Cells are individually labelled to show the continuity between time points and the appearance of new cells (bold). Scale bar is 5 µm. E) Several pillars with trapped yeast cells are shown as they appear during image segmentation. A custom script detects the cells and creates an outline (green) which is subsequently used to extract cell information.

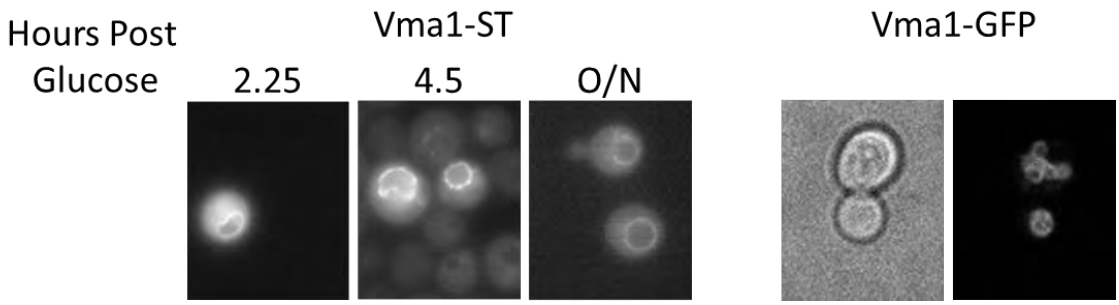



**Figure A3:** SpyoIPD-EGFP pulse-chase imaging of Vma1-SpyTag. Left panel of images: Yeast expressing Vma1-SpyTag and SpyoIPD-EGFP were grown overnight in noninducing media, diluted the next day into galactose containing media, and grown for 8 hours before adding glucose. Following the addition of glucose, samples were taken at several time points for imaging (times indicated above images, O/N ~ 16 hrs post glucose). Right panel: Yeast expressing Vma1-GFP.

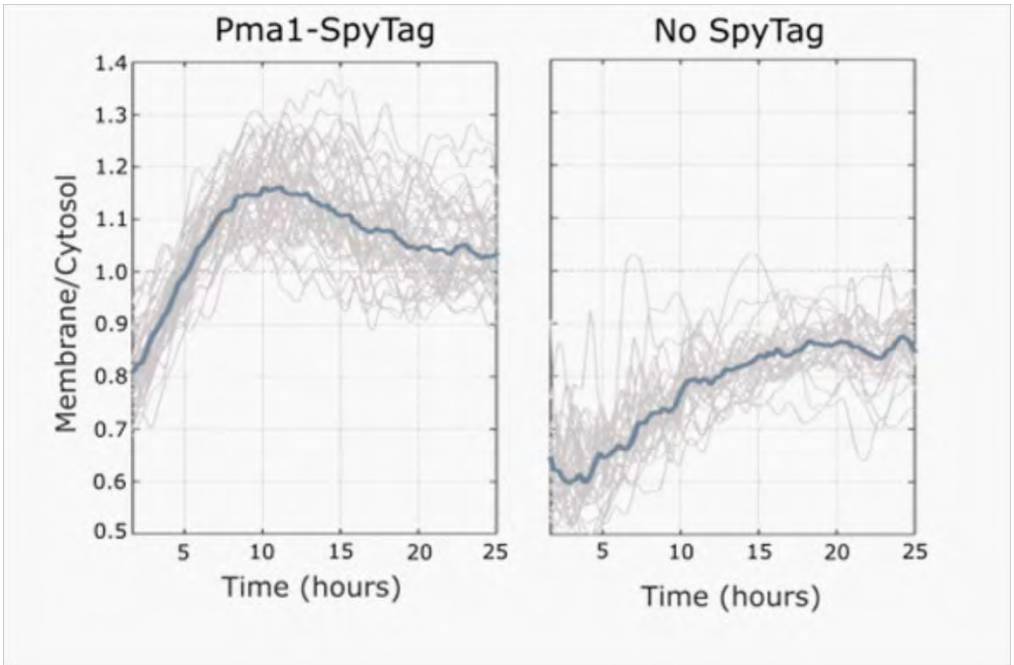



**Figure A3**: Comparing the ratio of membrane to cytosol signal over time for cells expressing SpyoIPD-GFP and either PMA1-SpyTag (PMA1-SpyTag, left plot) or untagged PMA1 (No SpyTag, right plot). Data was collected in a yeast microdevice with multiple chambers, allowing data to be collected on both strains simultaneously. The dark gray line corresponds to the average ratio of 44 cells (PMA1-SpyTag) and 31 cells (No SpyTag). Cell segmentation was determined using an active countour extraction method[115]. Plotting in this fashion allows easier visualization of the difference between tagged and untagged strains. To estimate the half-life of in membrane Pma1, the mean fluorescence decay from the membrane of cells expressing untagged Pma1 was substracted from the mean fluorescence decay of cells expressing tagged Pma1.

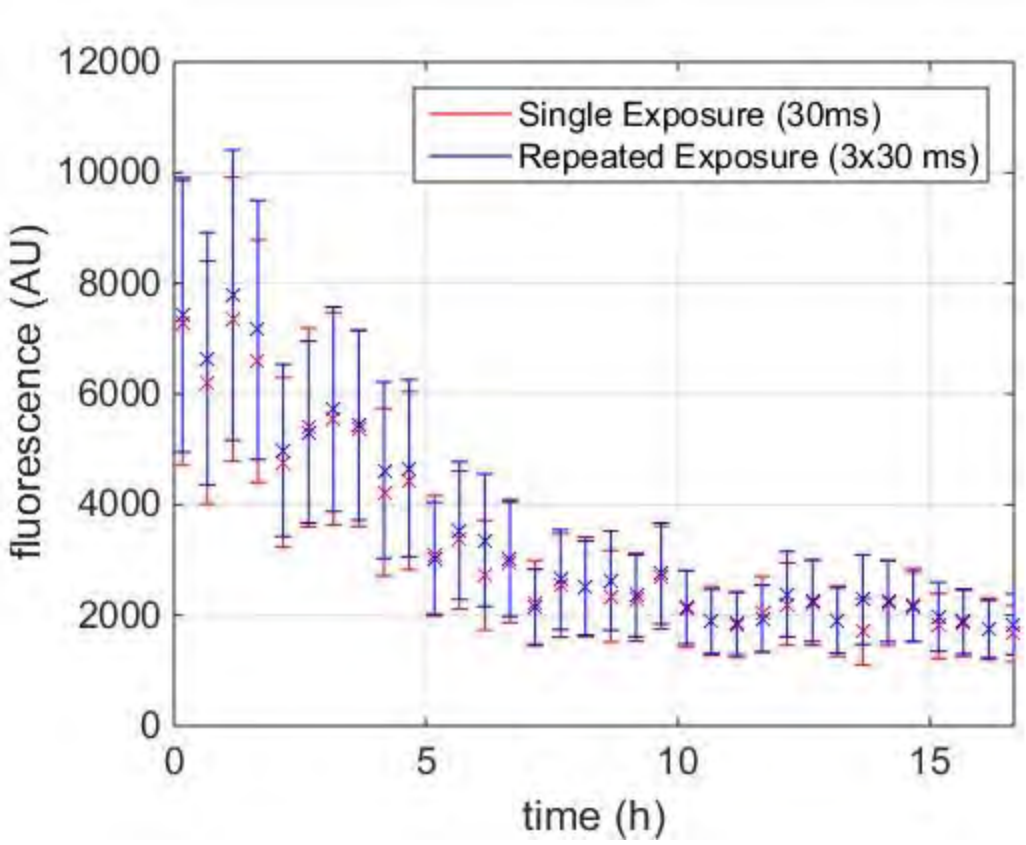



**Figure A4:** Investigating the effect of photobleaching in PMA1 turnover experiments. Cell expressing PMA1-SpyTag and SpyoIPD-GFP were grown overnight in galactose containing media, loaded onto yeast microdevices, and switched into media supplemented with glucose. Total cellular fluorescence was followed over time, exciting cells with either one 30 ms pulse (red bars) or three 30 ms pulses (blue bars) of excitation light at each time point. Vertical error bars show standard deviation for 17 (1 pulse) and 20 (3 pulses) individual cells. Mean values are indicated by crosses. Comparing the two data sets reveals that increasing exposure times does not accelerate loss of fluorescence (one-sided Fisher's t-test, $p = 0.05$, $n_1 = 17$, $n_2 = 20$), indicating that photobleaching does not contribute significantly to fluorescent decay.

| # | Name | Sequence | Description |
|---|---|---|---|
| 1 | Gal_SC_F | CGGATTCTAGAACTAGTGGATCCATGCGTCTGTCGTACTACCATCAC | Transferring SpyCatcher into p424GAL1 |
| 2 | SC_Gal_R | GAGTCATGTAATTAGTTATGTCACGCGTGACGCTCATATTAGATCTTAGTGA | Transferring SpyCatcher into p424GAL1 |
| 3 | 424_Sp_F | GGAGAAAAAACCCCGGATTC ACTTTAAGAAGGAGATATACATATGTCG | Transferring SpyoIPD (IA or IVML) into p424GAL1 |
| 4 | IPD_424_R | GCTTGATATCGAATTCCTGCAGAGCTCGAATTCCGGATCC | Transferring SpyoIPD (IA or IVML) into p424GAL1 |

**Table A1**: Oligos used to make transfer His6 tagged constructs from bacterial expression vectors into yeast shuttle vectors, for testing *in vivo* activity.

| # | Name | Sequence | Description |
|---|---|---|---|
| 1 | FA6LnkCYCF | GCTGAAGCTTCGTACGCT GGTGGATCAGGCTCTGGTT TGCAA TAA GTCATGTAATTAGTTATGTC ACGC | Inserting Linker-CYC1 into pFA6a-KANMX6 |
| 2 | CYC_pFA6_R | CTGGCGCGCCTTAATTAAC CGCAAATTAAAGCCTTCGA GC | Inserting Linker, ST, and mCherry-CYC1 into pFA6a-KANMX6 |
| 3 | Lnk_ST_CYCR | GTAAGCGTGACATAACTAA TTACATGATCATTATTTCGT CGGTTTATACGCATCCACC ATGACAATGTGAGCTTGCA AACCAGAGCCTG | Inserting SpyTag into pFA6a-Linker-HIS3 |
| 4 | FA6_F | GGTTATTGTCTCATGAGCG G | Inserting SpyTag into pFA6a-linker-HIS3 |
| 5 | 415_V5_GFP_F | GGATCCACTAGTTCTAGAT CCGATGGGTAAACCAATTC CAAATCCATTGTTGGGTTTG GATTCTACTGGTTCTAGTAA AGGAGAAGAACTTTTCACT G | Attaching V5 epitope to GFP and transferring to pCu415CUP1 |
| 6 | GFP_ST_415_R | CTATTAAAGCTTATCGATAC CGTCGACCTTAGTAGGTTT ATAAGCATCAACCATAACA ATATGAGCAGAACCTTTGT ATAGTTCATCCATGCCATG | Attaching ST to GFP and transferring to pCu415CUP1 |

**Table A2**: Oligos used to make template vectors for tagging proteins with either flexible linker or SpyTag, and oligos to make V5-EGFP-SpyTag and insert into the pCu415CUP1 plasmid.

| # | Name | Sequence | Description |
|---|---|---|---|
| 1 | PMA1_ST_F | ATGGCTGCTATGCAAAGAGTCTCTACTCAACACGAAAAGGAAACCGGTGGATCAGGCTCTGG | Attaching SpyTag to genomic PMA1 |
| 2 | PMA1_ST_R | AAAATGTGACAAAAATTATGATTAAATGCTACTTCAACAGGATTATTAGAAAAACTCATCGAGCATC | Attaching SpyTag to genomic PMA1 |
| 3 | HTB2_ST_F | GAAGGTACTAGGGCTGTTACCAAATACTCCTCCTCTACTCAAGCCGGTGGATCAGGCTCTGG | Attaching SpyTag to genomic HTB2 |
| 4 | HTB2_ST_R | GATGCTCGATGAGTTTTTCTAAGTCACTCACTAGGTATTGTGATTTAGTCATGTTTTCTTTTTATTA | Attaching SpyTag to genomic HTB2 |
| 5 | CDC12_ST_F | GAAGAGCAGGTCAAAAGCTTGCAAGTAAAAAAATCCCATTTAAAAGGTGGATCAGGCTCTGG | Attaching SpyTag to genomic CDC12 |
| 6 | CDC12_ST_R | GATGCTCGATGAGTTTTTCTAATGATTAATTAATGTCTTCCTCTTTGTCTCGTCAATTTCAACGCCT | Attaching SpyTag to genomic CDC12 |
| 7 | PMA1_CT_F | GAGGGTCACGAGAACACC | Checking C-terminus of genomic PMA1 |
| 8 | PMA1_CT_R | GAAAAATTAAACCAGAAAAATCAAGTTG | Checking C-terminus of genomic PMA1 |
| 9 | HTB2_CT_F | GCAAACTCACCCAGACAC | Checking C-terminus of genomic HTB2 |
| 10 | HTB2_CT_R | CCAAACTGCTCAAGATAAGATCG | Checking C-terminus of genomic HTB2 |
| 11 | CDC12_CT_F | GAGGGTCACGAGAACACC | Checking C-terminus of genomic CDC12 |
| 12 | CDC12_CT_R | CAGTTACTTCTGCTGGTTCC | Checking C-terminus of genomic CDC12 |
| 13 | GAL2_SpG_F | GGAGAAAAAACCCCGGATTCATGTCGTACTACCATCACCATC | Inserting SpyoIPD-GFP at GAL2 locus |
| 14 | GAL2_SpG_R | CCGCTGCCGCTGCCGCCAGCAGCAACCATGACAGC | Inserting SpyoIPD-GFP at GAL2 locus |

**Table A3**: Oligos used to tag target POIs with SpyTag or flexible linker, and screen resulting transformants for proper incorporation.

| # | Name | Sequence | Description |
|---|---|---|---|
| 1 | PFA6_mCher_F | GCTGAAGCTTCGTACGCTGCATCCGTGAGCAAGGG | Inserting mCherry into pFA6a-KANMX6 |
| 2 | mCher_CYC_R | CATAACTAATTACATGACTTATCACTTGTACAGCTCGTCC | Inserting mCherry into pFA6a-KANMX6 |
| 3 | TRAP_mChrF | CAGCGGCAGCGGCCTGCAGGCATCCGTGAGCAAGGG | Used to amp mCher from pNAS1b and stick into TRAP-mEOS bacterial vector (Pratt et al.) |
| 4 | Chr_pNAS_R | GCACGCGTACCATGAAGCTTTTATCACTTGTACAGCTCGTCCATGC | Used to amp mCher from pNAS1b and stick into TRAP-mEOS bacterial vector (Pratt et al.) |
| 5 | Pgal_TRAP_F | GGAGAAAAAACCCCGGATTCATGAAGCAGGCACTGAAAG | Used to amplify TRAP-mCherry from above vector and transfer into pFA6a-HIS3 |
| 6 | mCher_pGal_R | GCTTGATATCGAATTCCTGCAGTTATCACTTGTACAGCTCGTCC | Used to amplify TRAP-mCherry from above vector and transfer into pFA6a-HIS3 |
| 7 | Fa6LnkTgCYCF | GCTGAAGCTTCGTACGCTGGTGGATCAGGCTCTGGTTTGCAA TAA GTCATGTAATTAGTTATGTCACGC | Amp CYC, attach linker and linker-MEEVF, then use CPEC to stick into pfa6 vector |
| 8 | CYC_pfa6_R | CTGGCGCGCCTTAATTAACCGCAAATTAAAGCCTTCGAGC | Amp CYC, attach linker and linker-MEEVF, then use CPEC to stick into pfa6 vector |
| 9 | BamHI_Gal1_F | CTAGTT GGATCC GAGCTCTAGTACGGATTAGAAGC | Oligos to amp TRAP-EGFP and insert into pFA6a-HIS3 plasmid |
| 10 | CYC_SmaI_R | agttcACCCGGGCGCAAATTAAAGCCTTCGAGC | Oligos to amp TRAP-EGFP and insert into pFA6a-HIS3 plasmid |

**Table A4**: Oligos used to fuse TRAP 4 to various FPs, and generate template vectors for inserting into yeast.

| # | Name | Sequence | Description |
|---|---|---|---|
| 1 | GAL1_SC_F | GGAGAAAAAACCCCGGATTC ATGCGTCTGTCGTACTACC | Used to make SpyCatcher-GFP |
| 2 | SC_GS_R | CCGCTGCCGCTGCCGCCAGATC TAATATGAGCGTCACCTTTAG | Used to make SpyCatcher-GFP |
| 3 | Gal1_snoop_F | GGAGAAAAAACCCCGGATTC ATGGGCAGCAGCCATC | Used to make SnoopCatcher-GFP |
| 4 | Snoop_GS_R | CCGCTGCCGCTGCCGCCTTTCG GCGGTATCGGTTC | Used to make SnoopCatcher-GFP |
| 5 | Lnk_SnoT_CYC_R | GTAAGCGTGACATAACTAATTAC ATGATTATTTGTTCACCTTAATAAA TTCTATATCTCCCAGCTTGACAAT GTGAGCTTGCAAACCAGAGCCTG | Used to make SnoopTag in pfa6 vector |

**Table A5**: Oligos used to label target proteins with additional protein-peptide interactions (SpyCatcher, SnoopCatcher, and SnoopTag).

| # | Oligo Name | Sequence | Description |
|---|---|---|---|
| 1 | Sac_HXT1s_F | GCCATC GAGCTC CTAAATTCAAGGCGGATGTAAGG | Amplify HXT1s promoter and insert into p424 vector |
| 2 | HXT1_BAM_R | TAATGCGGATCCGATTTTACGTA TATCAACTAGTTGACG | Amplify HXT1s promoter and insert into p424 vector |
| 3 | Saci_HXT1L_F | GCCATC GAGCTC AAGCTTCCGATCCTCAAATAC | Amplify HXT1L promoter and insert into p424. Used HXT1_BAM_R as reverse primer. |
| 4 | Saci_HXT3_F | GCCATC GAGCTC GGATCCATCTAATCTGCAAAGTC | Amplify HXT3 promoter and insert into p424 vector |
| 5 | HXT3_smai_R | TAATGCCCCGGGGAATTCATGAT TGTTTAACTCAGATG | Amplify HXT3 promoter and insert into p424 vector |
| 6 | HXT1_seq_F | CCACCTTAAAATCTATAAAGATATCATA ATCG | Fwd seq primer for p424 HXT1 (s and L) |
| 7 | HXT3_seq_F | GAATCACAAACAAAATTTACATCTGAG | Fwd seq primer for p424 HXT3 |
| 8 | Sma_GFP_F | ACATC CCCGGG ATGAGTAAAGGAGAAGAACTTTTCACT G | Amplify and insert EGFP into p415HXT plasmids |
| 9 | GFP_Sal_R | ATGGCAGTCGACCTATTATTTGTATAGTT CATCCATGCCATG | Amplify and insert EGFP into p424HXT plasmids |
| 10 | Spe_Cher_F | TCAATG ACTAGT ATG GCATCCGTGAGCAAGGG | Amplify and insert mCherry into p424HXT plasmids |
| 11 | Cher_Bam_R | GATCGAGGATCCTTATCACTTGTACAGC TCGTCCATG | Amplify and insert mCherry into p424HXT plasmids |
| 12 | SpeNLSChr_F | TCAATG ACTAGT ATG cccaagaaaaagcgcaaggta GGT TCT GCATCCGTGAGCAAGGG | Amplify and insert NLS-mCherry into p424HXT plasmids |
| 13 | ChrST_Bam_R | GATCGAGGATCCTTATCATTTCGTCGGTT TATACGCATCCACCATGACAATGTGAGC CTGCAGGCCGCTGCCGCTGCCGCCCTTG TACAGCTCGTCCATG | Amplify and insert mCherry-ST into p424HXT plasmids |
| 14 | GAL1NLSSCF | GGAGAAAAAACCCCGGATTC Atg GGT TCT cccaagaaaaagcgcaaggta ATGCGTCTGTCGTACTACC | Forward primer used to make NLS-SC-GFP |

**Table A6**: Oligos used in chapter 3 to create various nuclear localizing proteins and glucose inducing promoters.

| Name | Parent | Description | Source |
|---|---|---|---|
| pFA6a-HIS3MX6 | | Yeast Insertion vector, His3 marker | Longtine et al. |
| pFA6a-SpyGFP-HIS3MX6 | pFA6His3MX6 | SpyoIPD-GFP insertion template. HIS3 selection marker | This Study |
| pFA6a-SCGFP-HIS3MX6 | pFA6His3MX6 | SpyCatcher-GFP insertion template. HIS3 selection marker | This Study |
| pFA6a-NLS-SCGFP-HIS3MX6 | pFA6His3MX6 | NLS-SpyCatcher-GFP insertion template. HIS3 selection marker | This Study |
| pFA6a-KANMX6 | | Yeast Insertion vector, KanR marker | Longtine et al. |
| pFA6a-Link-KANMX6 | pFA6a-KanMX6 | Template vector for tagging targets proteins with flexible linker. KANR marker | This Study |
| pFA6a-SpyTag-KANMX6 | pFA6a-KanMX6 | Template vector for tagging targets proteins with SpyTag. KANR marker | This Study |
| pFA6a-Cherry-KANMX6 | pFA6a-KanMX6 | Template vector for tagging targets proteins with mCherry. KanR marker | This Study |
| p424 GAL1 | | 2µ/Trp, GAL1, empty vector | Mumberg et al. |
| p424 GAL1 SpyCatcher | p424GAL1 | 2µ/Trp, GAL1, SpyCatcher | This Study |
| p424 GAL1 SpyoIPD | p424GAL1 | 2µ/Trp, GAL1, SpyoIPD | This Study |
| pCu415CUP1 | | CEN/Leu, CUP1, empty vector | Mumberg et al. |
| pCu415CUP1 GFPST | pcu415CUP1 | CEN/Leu, CUP1, GFP-SpyTag | This Study |
| p424 HXT1s | p424GAL1 | 2µ/Trp, GAL1, HXT1s promoter | This Study |
| p424 HXT1L | p424GAL1 | 2µ/Trp, GAL1, HXT1L promoter | This Study |
| p424 HXT3 | p424GAL1 | 2µ/Trp, GAL1, HXT3 promoter | This Study |

**Table A7**: Plasmids used in this work.

**Derivation of PMA1 half-life**

We used a Bayesian approach to estimate the half-life of PMA1 from the data of Fig. 6[116]. The median fluorescence of single cells falls over time. Each cell has a different initial and final value and 120 cells were followed over 58 time-points with fluorescence measurements at intervals of 30 minutes.

Considering first a single cell, the observed fluorescence will have three components: SpyoIPD-GFP covalently bound to Pma1 at the plasma membrane; SpyoIPD-GFP in the cytosol; and intrinsic cellular autofluorescence. At the membrane, GFP decays at a rate determined by the half-life of Pma1, because Pma1 is not transported into the daughter cell and GFP itself hardly decays; in the cytoplasm, GFP principally decays through entering and remaining in the daughter cell. We assume that autofluorescence does not decay. Mathematically, then, the observed fluorescence, $f$, obeys

$$f(t) = m_0 e^{-d_m t} + c_0 e^{-d_c t} + a_0 \tag{1}$$

where the initial GFP at the membrane, $m_0$, decays exponentially with a decay rate of $d_m$ and the initial GFP in the cytoplasm, $c_0$, decays exponentially with a decay rate of $d_c$. The autofluorecence, $a_0$, does not change with time $t$. The decay rates $d_m$ and $d_c$ are our focus, and $m_0$, $c_0$, and $a_0$ are `nuisance' parameters, which ideally we would integrate away.

For a cell indexed by $j$ with data $D^{(j)}$, the posterior probability of $d_m$ and $d_c$, which we write as $P^{(j)}\left(d_m, d_c \middle| D^{(j)}\right)$ to emphasize its dependence on cell $j$, is

$$P^{(j)}\left(d_m, d_c \middle| D^{(j)}\right) = \int dm_0 dc_0 da_0 P^{(j)}\left(m_0, c_0, a_0, d_m, d_c \middle| D^{(j)}\right)$$

$$\sim \int dm_0 dc_0 da_0 P^{(j)}\left(D^{(j)}, m_0, c_0, a_0 \middle| d_m, d_c\right) \tag{2}$$

assuming uniform, bounded prior probabilities for $d_m$ and $d_c$. We further assume that the errors in measuring fluorescence are identically and independently distributed with a Gaussian distribution of zero mean, but with a variance that changes with time. Assuming the measurement error is dominated by shot noise, we empirically estimate this variance, $\sigma_i^2$, as the mean fluorescence taken across all cells at each time point.

The likelihood for cell $j$ is then

$$P^{(j)}\left(D^{(j)} \middle| d_m, d_c\right) \sim \int dm_0 dc_0 da_0 \prod_i \exp\left[\frac{-(d_i - m_0 e^{-d_m t_i} - c_0 e^{-d_c t_i} - a_0)^2}{2\sigma_i^2}\right]$$

$$= \int dm_0 dc_0 da_0 \ \exp\left[\sum_i \frac{-(d_i - m_0 e^{-d_m t_i} - c_0 e^{-d_c t_i} - a_0)^2}{2\sigma_i^2}\right] \tag{3}$$

where $i$ runs over all time points and $\sigma_i$ is known.

By making $c_0$ and $m_0$ be independent variables in Eq. 3, we assume that the amount of GFP bound to Pma1 at the membrane, $m_0$, and the amount of GFP in the cytoplasm, $c_0$, are independent. SpyoIPD-GFP is covalently, essentially irreversibly, bound to Pma1 at the membrane, and our assumption of independence is strengthened if all the Pma1 at the membrane is bound by GFP. The integral in Eq. 3 should peak at positive values of $m_0$, $c_0$, and $a_0$, and therefore we can extend the range of integration to be over all real numbers, enabling an exact integration if an approximate calculation[116].

Integrating Eq. 3 and defining $M_i = e^{-d_m t_i}$ and $C_i = e^{-d_c t_i}$ we find ten sufficient statistics:

$$\begin{array}{rcl} T_1 = \sum_i C_i^2/\sigma_i^2 & ; & T_2 = \sum_i d_i C_i^2/\sigma_i^2 \\ T_3 = \sum_i C_i M_i/\sigma_i^2 & ; & T_4 = \sum_i d_i^2/\sigma_i^2 \\ T_5 = \sum_i d_i M_i/\sigma_i^2 & ; & T_6 = \sum_i M_i^2/\sigma_i^2 \\ T_7 = \sum_i C_i/\sigma_i^2 & ; & T_8 = \sum_i \sigma_i^{-2} \\ T_9 = \sum_i d_i/\sigma_i^2 & ; & T_{10} = \sum_i M_i \sigma_i^2 \end{array} \tag{4}$$

and that

$$P^{(j)}\left(D^{(j)}\middle|d_m, d_c\right) \sim \exp\left[\frac{-T_4}{2} + \frac{T_1 T_5^2 + T_2^2 T_6 - 2T_2 T_3 T_5}{2u_T}\right] \times \exp\left[\frac{(T_2 T_3 T_{10} - T_1 T_5 T_{10} + T_3 T_5 T_7 - T_2 T_6 T_7 + T_9 u_T)^2}{2u_T v_T}\right] / \sqrt{v_T} \tag{5}$$

with

$$u_T = T_1 T_6 - T_3^2 \tag{6}$$

and

$$v_T = T_8 u_T - T_6 T_7^2 + 2T_3 T_7 T_{10} - T_1 T_{10}^2 \tag{7}$$

To extend our analysis to more than one cell, we assume that $m_0$, $c_0$, and $a_0$ for each cell are independent of their values in other cells so that the posterior probability satisfies

$$\begin{aligned} P(d_m, d_c|D) &\sim P(D|d_m, d_c) \\ &= \prod_j P^{(j)}\left(D^{(j)}\middle|d_m, d_c\right) \end{aligned} \tag{8}$$

for $j$ running over all cells and $P^{(j)}\left(D^{(j)}\middle|d_m, d_c\right)$ being given by Eq. 5 with the sufficient statistics of Eq. 4 evaluated using the data from cell $j$.

For our data comprising 120 cells and 58 time-points per cell, plotting the negative logarithm of the posterior probability (Figure 6 – figure supplement 4) shows a

well-defined minimum (corresponding to a maximum of the probability). We are interested in the half-life of Pma1, $\tau_m$, which is determined by $d_m$. Integrating the posterior probability over $d_c$ and using $\tau_m = \log 2/\mathrm{d}_m$ and that

$$P(\tau_m) = \frac{\tau_m^2}{\log 2} P(d_m = \frac{\log 2}{\tau_m}) \tag{9}$$

through changing variables, we find that the marginal posterior probability for $\tau_m$ is sharply peaked at ≅ 11.5 hours with a 90% credible interval of 11.3 < $\tau_m$ < 11.7 hours (Figure 2.6 inset).

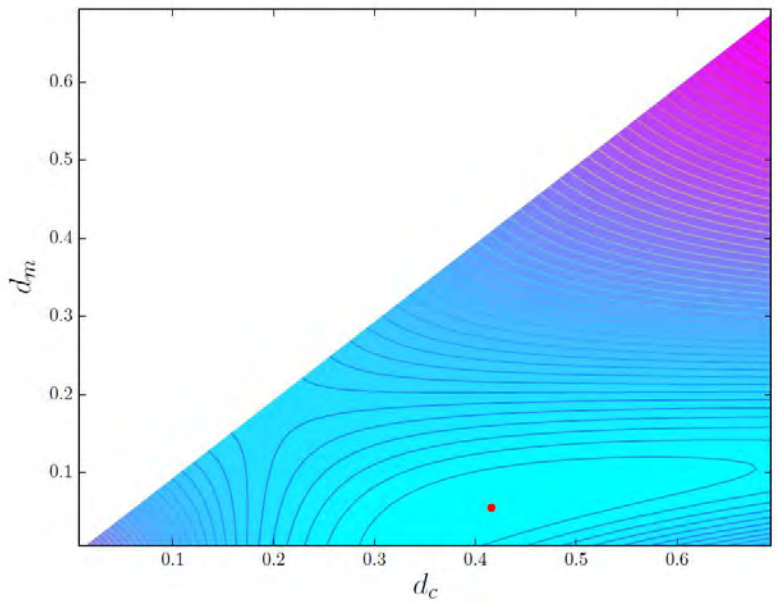



**Figure A5**: The posterior probability of $d_m$ and $d_c$ has a maximum at $d_c$ ≅ 0.4 hr$^{-1}$ and $d_m$ ≅ 0.06 hr$^{-1}$ (red dot). We plot the negative logarithm of the posterior probability, which has a minimum (light blue) at the most probable values of $d_m$ and $d_c$. Darker blue and violet shading correspond to higher values. The probability is symmetric in $d_m$ and $d_c$, and we plot only for $d_m < d_c$ because decay at the membrane is assumed to be slower than decay in the cytoplasm.

# References


1. Miyawaki, A., Llopis, J., Heim, R., McCaffery, J. M., Adams, J. A., Ikura, M., and Tsien, R. Y. (1997). Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. *Nature* 388, 882-887.

2. Kreis, T. E., Winterhalter, K. H., and Birchmeier, W. (1979). In vivo distribution and turnover of fluorescently labeled actin microinjected into human fibroblasts. *Proceedings of the National Academy of Sciences* 76, 3814-3818.

3. Feramisco, J. R. (1979). Microinjection of fluorescently labeled α-actinin into living fibroblasts. *Proceedings of the National Academy of Sciences* 76, 3967-3971.

4. Kreis, T. E. (1986). Preparation, assay, and microinjection of fluorescently labeled cytoskeletal proteins: Actin, α-actinin, and vinculin. *Methods in enzymology* 134, 507-519.

5. Modesti, M. (2011). Fluorescent labeling of proteins. *Single Molecule Analysis: Methods and Protocols*, 101-120.

6. Aroush, D. R.-B., Ofer, N., Abu-Shah, E., Allard, J., Krichevsky, O., Mogilner, A., and Keren, K. (2017). Actin Turnover in Lamellipodial Fragments. *Current Biology* 27, 2963-2973. e2914.

7. König, I., Zarrine-Afsar, A., Aznauryan, M., Soranno, A., Wunderlich, B., Dingfelder, F., Stüber, J. C., Plückthun, A., Nettels, D., and Schuler, B. (2015). Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nature methods* 12, 773-779.

8. Wang, S., and Hazelrigg, T. (1994). Implications for bcd mRNA localization from spatial distribution of exu protein in Drosophila oogenesis. *Nature* 369, 400-403.

9. Yu, D., Baird, M. A., Allen, J. R., Howe, E. S., Klassen, M. P., Reade, A., Makhijani, K., Song, Y., Liu, S., and Murthy, Z. (2015). A naturally monomeric infrared fluorescent protein for protein labeling in vivo. *Nature methods* 12, 763.

10. Kumagai, A., Ando, R., Miyatake, H., Greimel, P., Kobayashi, T., Hirabayashi, Y., Shimogori, T., and Miyawaki, A. (2013). A bilirubin-inducible fluorescent protein from eel muscle. *Cell* 153, 1602-1611.

11. Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173, 33-38.

12. Shaner, N. C., Lambert, G. G., Chammas, A., Ni, Y., Cranfill, P. J., Baird, M. A., Sell, B. R., Allen, J. R., Day, R. N., and Israelsson, M. (2013). A bright

monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nature methods* 10, 407-409.

13. Zacharias, D. A., Violin, J. D., Newton, A. C., and Tsien, R. Y. (2002). Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* 296, 913-916.

14. Megley, C. M., Dickson, L. A., Maddalo, S. L., Chandler, G. J., and Zimmer, M. (2008). Photophysics and dihedral freedom of the chromophore in yellow, blue, and green fluorescent protein. *The Journal of Physical Chemistry B* 113, 302-308.

15. Zimmer, M. (2012). What does it take to improve existing fluorescent proteins for in vivo imaging applications? *In Vivo Cellular Imaging Using Fluorescent Proteins: Methods and Protocols*, 235-241.

16. Ilagan, R. P., Rhoades, E., Gruber, D. F., Kao, H. T., Pieribone, V. A., and Regan, L. (2010). A new bright green-emitting fluorescent protein–engineered monomeric and dimeric forms. *The FEBS journal* 277, 1967-1978.

17. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nature biotechnology* 24, 79-88.

18. Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., and Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature biotechnology* 22, 1567-1572.

19. Shcherbo, D., Murphy, C. S., Ermakova, G. V., Solovieva, E. A., Chepurnykh, T. V., Shcheglov, A. S., Verkhusha, V. V., Pletnev, V. Z., Hazelwood, K. L., and Roche, P. M. (2009). Far-red fluorescent tags for protein imaging in living tissues. *Biochemical journal* 418, 567-574.

20. Heim, R., and Tsien, R. Y. (1996). Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current biology* 6, 178-182.

21. Shaner, N. C., Lin, M. Z., McKeown, M. R., Steinbach, P. A., Hazelwood, K. L., Davidson, M. W., and Tsien, R. Y. (2008). Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nature methods* 5, 545-551.

22. van de Linde, S., and Sauer, M. (2014). How to switch a fluorophore: from undesired blinking to controlled photoswitching. *Chemical Society Reviews* 43, 1076-1087.

23. Rodriguez, E. A., Campbell, R. E., Lin, J. Y., Lin, M. Z., Miyawaki, A., Palmer, A. E., Shu, X., Zhang, J., and Tsien, R. Y. (2017). The growing and glowing toolbox of fluorescent and photoactive proteins. *Trends in biochemical sciences* 42, 111-129.

24. Thestrup, T., Litzlbauer, J., Bartholomäus, I., Mues, M., Russo, L., Dana, H., Kovalchuk, Y., Liang, Y., Kalamakis, G., and Laukat, Y. (2014). Optimized ratiometric calcium sensors for functional in vivo imaging of neurons and T lymphocytes. *Nature methods* 11, 175-182.

25. Mahon, M. J. (2011). pHluorin2: an enhanced, ratiometric, pH-sensitive green florescent protein. *Advances in bioscience and biotechnology (Print)* 2, 132.

26. Sankaranarayanan, S., De Angelis, D., Rothman, J. E., and Ryan, T. A. (2000). The use of pHluorins for optical measurements of presynaptic activity. *Biophysical journal* 79, 2199-2208.

27. Shen, Y., Rosendale, M., Campbell, R. E., and Perrais, D. (2014). pHuji, a pH-sensitive red fluorescent protein for imaging of exo-and endocytosis. *J Cell Biol* 207, 419-432.

28. Benčina, M. (2013). Illumination of the spatial order of intracellular pH by genetically encoded pH-sensitive sensors. *Sensors* 13, 16736-16758.

29. Ghosh, I., Hamilton, A. D., and Regan, L. (2000). Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *Journal of the American Chemical Society* 122, 5658-5659.

30. Wilson, C. G., Magliery, T. J., and Regan, L. (2004). Detecting protein-protein interactions with GFP-fragment reassembly. *Nature methods* 1, 255-262.

31. Jackrel, M. E., Cortajarena, A. L., Liu, T. Y., and Regan, L. (2010). Screening libraries to identify proteins with desired binding activities using a split-GFP reassembly assay. *ACS chemical biology* 5, 553-562.

32. Sawyer, N., Gassaway, B. M., Haimovich, A. D., Isaacs, F. J., Rinehart, J., and Regan, L. (2014). Designed phosphoprotein recognition in Escherichia coli. *ACS Chemical Biology* 9, 2502-2507.

33. Speltz, E. B., Nathan, A., and Regan, L. (2015). Design of protein–peptide interaction modules for assembling supramolecular structures in vivo and in vitro. *ACS chemical biology* 10, 2108-2115.

34. Kodama, Y., and Hu, C.-D. (2012). Bimolecular fluorescence complementation (BiFC): a 5-year update and future perspectives. *Biotechniques* 53, 285-298.

35. Stadler, C., Rexhepaj, E., Singan, V. R., Murphy, R. F., Pepperkok, R., Uhlén, M., Simpson, J. C., and Lundberg, E. (2013). Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nature methods* 10, 315-323.

36. Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'shea, E. K., and Weissman, J. S. (2003). Global analysis of protein expression in yeast. *Nature* 425, 737-741.

37. Brohée, S., Barriot, R., Moreau, Y., and André, B. (2010). YTPdb: a wiki database of yeast membrane transporters. *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1798, 1908-1912.

38. Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature Biotechnology* 21, 86-89.

39. Gautier, A., Juillerat, A., Heinis, C., Corrêa, I. R., Kindermann, M., Beaufils, F., and Johnsson, K. (2008). An engineered protein tag for multiprotein labeling in living cells. *Chemistry & Biology* 15, 128-136.

40. Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., Wood, M. G., Learish, R., Ohana, R. F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G., Zhu, J., Darzins, A., Klaubert, D. H., Bulleit, R. F., and Wood, K. V. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chemical Biology* 3, 373-382.

41. Miller, L. W., Cai, Y., Sheetz, M. P., and Cornish, V. W. (2005). In vivo protein labeling with trimethoprim conjugates: a flexible chemical tag. *Nature methods* 2, 255-257.

42. Chen, Z., Jing, C., Gallagher, S. S., Sheetz, M. P., and Cornish, V. W. (2012). Second-generation covalent TMP-tag for live cell imaging. *Journal of the American Chemical Society* 134, 13692-13699.

43. Lacy, M. M., Baddeley, D., and Berro, J. (2017). Single-molecule imaging of the BAR-domain protein Pil1p reveals filament-end dynamics. *Molecular Biology of the Cell* 28, 2251-2259.

44. McMurray, M. A., and Thorner, J. (2008). Septin stability and recycling during dynamic structural transitions in cell division and development. *Current Biology* 18, 1203-1208.

45. Leng, S., Qiao, Q.-L., Gao, Y., Miao, L., Deng, W.-G., and Xu, Z.-C. (2017). SNAP-tag fluorogenic probes for wash free protein labeling. *Chinese Chemical Letters*.

46. Liu, Y., Miao, K., Dunham, N. P., Liu, H., Fares, M., Boal, A. K., Li, X., and Zhang, X. (2017). The cation− π interaction enables a halo-tag fluorogenic probe for fast no-wash live cell imaging and gel-free protein quantification. *Biochemistry* 56, 1585-1595.

47. Stagge, F., Mitronova, G. Y., Belov, V. N., Wurm, C. A., and Jakobs, S. (2013). SNAP-, CLIP-and Halo-tag labelling of budding yeast cells. *PloS one* 8, e78745.

48. Griffin, B. A., Adams, S. R., and Tsien, R. Y. (1998). Specific covalent labeling of recombinant protein molecules inside live cells. *Science* 281, 269-272.

49. Adams, S. R., Campbell, R. E., Gross, L. A., Martin, B. R., Walkup, G. K., Yao, Y., Llopis, J., and Tsien, R. Y. (2002). New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *Journal of the American Chemical Society* 124, 6063-6076.

50. Langhorst, M. F., Genisyuerek, S., and Stuermer, C. A. (2006). Accumulation of FlAsH/Lumio Green in active mitochondria can be reversed by β-mercaptoethanol for specific staining of tetracysteine-tagged proteins. *Histochemistry and cell biology* 125, 743.

51. Halo, T. L., Appelbaum, J., Hobert, E. M., Balkin, D. M., and Schepartz, A. (2008). Selective recognition of protein tetraserine motifs with a cell-permeable, pro-fluorescent bis-boronic acid. *Journal of the American Chemical Society* 131, 438-439.

52. Howarth, M., and Ting, A. Y. (2008). Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin. *Nature protocols* 3, 534-545.

53. George, N., Pick, H., Vogel, H., Johnsson, N., and Johnsson, K. (2004). Specific labeling of cell surface proteins with chemically diverse compounds. *Journal of the American Chemical Society* 126, 8896-8897.

54. Popp, M. W., Antos, J. M., Grotenbreg, G. M., Spooner, E., and Ploegh, H. L. (2007). Sortagging: a versatile method for protein labeling. *Nature chemical biology* 3, 707-708.

55. Antos, J. M., Chew, G.-L., Guimaraes, C. P., Yoder, N. C., Grotenbreg, G. M., Popp, M. W.-L., and Ploegh, H. L. (2009). Site-specific N-and C-terminal labeling of a single polypeptide using sortases of different specificity. *Journal of the American Chemical Society* 131, 10800-10801.

56. Uttamapinant, C., White, K. A., Baruah, H., Thompson, S., Fernández-Suárez, M., Puthenveetil, S., and Ting, A. Y. (2010). A fluorophore ligase for site-specific protein labeling inside living cells. *Proceedings of the National Academy of Sciences* 107, 10914-10919.

57. Sluder, G., and Wolf, D. E. (2013) *Digital microscopy*, Vol. 114, Academic Press.
58. Liu, C. C., and Schultz, P. G. (2010). Adding new chemistries to the genetic code. *Annual review of biochemistry* 79, 413-444.

59. Charbon, G., Brustad, E., Scott, K. A., Wang, J., Løbner-Olesen, A., Schultz, P. G., Jacobs-Wagner, C., and Chapman, E. (2011). Subcellular protein localization by using a genetically encoded fluorescent amino acid. *ChemBioChem* 12, 1818-1821.

60. Chatterjee, A., Guo, J., Lee, H. S., and Schultz, P. G. (2013). A genetically encoded fluorescent probe in mammalian cells. *Journal of the American Chemical Society* 135, 12540-12543.

61. Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D. J., Blackman, M. L., Fox, J. M., and Chin, J. W. (2012). Genetic encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels–Alder reactions. *Journal of the American Chemical Society* 134, 10317-10320.

62. Uttamapinant, C., Howe, J. D., Lang, K., Ber nek, V. c., Davis, L., Mahesh, M., Barry, N. P., and Chin, J. W. (2015). Genetic code expansion enables live-cell and super-resolution imaging of site-specifically labeled cellular proteins. *Journal of the American Chemical Society* 137, 4602-4605.

63. Park, H., Kang, H., Ko, W., Lee, W., Jo, K., and Lee, H. S. (2015). FRET-based analysis of protein-nucleic acid interactions by genetically incorporating a fluorescent amino acid. *Amino acids* 47, 729-734.

64. Summerer, D., Chen, S., Wu, N., Deiters, A., Chin, J. W., and Schultz, P. G. (2006). A genetically encoded fluorescent amino acid. *Proceedings of the National Academy of Sciences* 103, 9785-9789.

65. Amaro, M., Brezovsk , J., Kov čov , S., S kora, J., Bedn , D., N mec, V. c., Li kov , V., Kurumbang, N. P., Beerens, K., and Chaloupkov , R. (2015). Site-specific analysis of protein hydration based on unnatural amino acid fluorescence. *Journal of the American Chemical Society* 137, 4988-4992.

66. Wen, M., Guo, X., Sun, P., Xiao, L., Li, J., Xiong, Y., Bao, J., Xue, T., Zhang, L., and Tian, C. (2015). Site-specific fluorescence spectrum detection and characterization of hASIC1a channels upon toxin mambalgin-1 binding in live mammalian cells. *Chemical Communications* 51, 8153-8156.

67. Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H.-R., Haimovich, A. D., Kuznetsov, G., Mercer, J. A., Wang, H. H., Carr, P. A., and Mosberg, J. A. (2013). Genomically recoded organisms expand biological functions. *Science* 342, 357-360.

68. Sato, Y., Mukai, M., Ueda, J., Muraki, M., Stasevich, T. J., Horikoshi, N., Kujirai, T., Kita, H., Kimura, T., and Hira, S. (2013). Genetically encoded system to track histone modification in vivo. *Scientific reports* 3, 2436.

69. Rothbauer, U., Zolghadr, K., Tillib, S., Nowak, D., Schermelleh, L., Gahl, A., Backmann, N., Conrath, K., Muyldermans, S., and Cardoso, M. C. (2006). Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nature methods* 3, 887-889.

70. Koide, A., Abbatiello, S., Rothgery, L., and Koide, S. (2002). Probing protein conformational changes in living cells by using designer binding proteins: application to the estrogen receptor. *Proceedings of the National Academy of Sciences* 99, 1253-1258.

71. Kummer, L., Parizek, P., Rube, P., Millgramm, B., Prinz, A., Mittl, P. R., Kaufholz, M., Zimmermann, B., Herberg, F. W., and Plückthun, A. (2012). Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries. *Proceedings of the National Academy of Sciences* 109, E2248-E2257.

72. Pratt, S. E., Speltz, E. B., Mochrie, S. G., and Regan, L. (2016). Designed Proteins as Novel Imaging Reagents in Living Escherichia coli. *ChemBioChem* 17, 1652-1657.

73. Kaiser, P. D., Maier, J., Traenkle, B., Emele, F., and Rothbauer, U. (2014). Recent progress in generating intracellular functional antibody fragments to target and trace cellular components in living cells. *Biochimica Et Biophysica Acta (BBA)-Proteins and Proteomics* 1844, 1933-1942.

74. Cortajarena, A. L., Yi, F., and Regan, L. (2008). Designed TPR modules as novel anticancer agents. *ACS chemical biology* 3, 161-166.

75. Nizak, C., Martin-Lluesma, S., Moutel, S., Roux, A., Kreis, T. E., Goud, B., and Perez, F. (2003). Recombinant antibodies against subcellular fractions used to track endogenous Golgi protein dynamics in vivo. *Traffic* 4, 739-753.

76. Gross, G. G., Junge, J. A., Mora, R. J., Kwon, H.-B., Olson, C. A., Takahashi, T. T., Liman, E. R., Ellis-Davies, G. C., McGee, A. W., and Sabatini, B. L. (2013). Recombinant probes for visualizing endogenous synaptic proteins in living neurons. *Neuron* 78, 971-985.

77. Freund, G., Desplancq, D., Stoessel, A., Weinsanto, R., Sibler, A. P., Robin, G., Martineau, P., Didier, P., Wagner, J., and Weiss, E. (2014). Generation of an intrabody-based reagent suitable for imaging endogenous proliferating cell nuclear antigen in living cancer cells. *Journal of Molecular Recognition* 27, 549-558.

78. Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., and Vale, R. D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159, 635-646.

79. Kamiyama, D., Sekine, S., Barsi-Rhyne, B., Hu, J., Chen, B., Gilbert, L. A., Ishikawa, H., Leonetti, M. D., Marshall, W. F., and Weissman, J. S. (2016). Versatile protein tagging in cells with split fluorescent protein. *Nature communications* 7, 11046.

80. Feng, S., Sekine, S., Pessino, V., Li, H., Leonetti, M. D., and Huang, B. (2017). Improved Split Fluorescent Proteins For Endogenous Protein Labeling. *bioRxiv*, 137059.

81. Hinrichsen, M., Lenz, M., Edwards, J., Miller, O., Mochrie, S., Swain, P., Schwarz-Linek, U., and Regan, L. (2017). A new method for post-translationally labeling proteins in live cells for fluorescence imaging and tracking. *Protein Engineering, Design and Selection* 30, 771-780.

82. Hagan, R. M., Björnsson, R., McMahon, S. A., Schomburg, B., Braithwaite, V., Bühl, M., Naismith, J. H., and Schwarz-Linek, U. (2010). NMR spectroscopic and theoretical analysis of a spontaneously formed Lys−Asp isopeptide bond. *Angewandte Chemie* 122, 8599-8603.

83. Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., and Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proceedings of the National Academy of Sciences* 109, E690-E697.

84. Hawkins, K. M., and Smolke, C. D. (2006). The regulatory roles of the galactose permease and kinase in the induction response of the GAL network in Saccharomyces cerevisiae. *Journal of Biological Chemistry* 281, 13485-13492.

85. Madden, K., and Snyder, M. (1998). Cell polarity and morphogenesis in budding yeast. *Annual Reviews in Microbiology* 52, 687-744.

86. Haarer, B. K., and Pringle, J. R. (1987). Immunofluorescence localization of the Saccharomyces cerevisiae CDC12 gene product to the vicinity of the 10-nm filaments in the mother-bud neck. *Molecular and Cellular Biology* 7, 3678-3687.

87. Mason, A. B., Allen, K. E., and Slayman, C. W. (2014). C-terminal truncations of the Saccharomyces cerevisiae PMA1 H+-ATPase have major impacts on protein conformation, trafficking, quality control, and function. *Eukaryotic Cell* 13, 43-52.

88. Seibel, N. M., Eljouni, J., Nalaskowski, M. M., and Hampe, W. (2007). Nuclear localization of enhanced green fluorescent protein homomultimers. *Analytical biochemistry* 368, 95-99.

89. Serrano, R., Kielland-Brandt, M. C., and Fink, G. R. (1986). Yeast plasma membrane ATPase is essential for growth and has homology with (Na++ K+), K+-and Ca2+-ATPases. *Nature* 319, 689-693.

90. Henderson, K. A., Hughes, A. L., and Gottschling, D. E. (2014). Mother-daughter asymmetry of pH underlies aging and rejuvenation in yeast. *Elife* 3, DOI: 10.7554/eLife.03504.

91. Ferreira, T., Mason, A. B., and Slayman, C. W. (2001). The yeast Pma1 proton pump: a model for understanding the biogenesis of plasma membrane proteins. *Journal of Biological Chemistry* 276, 29613-29616.

92. Crane, M. M., Clark, I. B., Bakker, E., Smith, S., and Swain, P. S. (2014). A microfluidic system for studying ageing and dynamic single-cell responses in budding yeast. *PLoS One* 9, DOI: 10.1371/journal.pone.0100042.

93. Benito, B., Moreno, E., and Lagunas, R. (1991). Half-life of the plasma membrane ATPase and its activating system in resting yeast cells. *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1063, 265-268.

94. Ozcan, S., and Johnston, M. (1995). Three different regulatory mechanisms enable yeast hexose transporter (HXT) genes to be induced by different levels of glucose. *Molecular and Cellular Biology* 15, 1564-1572.

95. Tomás-Cobos, L., Casadomé, L., Mas, G., Sanz, P., and Posas, F. (2004). Expression of the HXT1 low affinity glucose transporter requires the coordinated activities of the HOG and glucose signalling pathways. *Journal of Biological Chemistry* 279, 22010-22019.

96. Ferrer-Martínez, A., Riera, A., Jiménez-Chillarón, J. C., Herrero, P., Moreno, F., and Gómez-Foix, A. M. (2004). A glucose response element from the S. cerevisiae hexose transporter HXT1 gene is sensitive to glucose in human fibroblasts. *Journal of molecular biology* 338, 657-667.

97. Renicke, C., Schuster, D., Usherenko, S., Essen, L.-O., and Taxis, C. (2013). A LOV2 domain-based optogenetic tool to control protein degradation and cellular function. *Chemistry & biology* 20, 619-626.

98. Feng, S., Sekine, S., Pessino, V., Li, H., Leonetti, M. D., and Huang, B. (2017). Improved split fluorescent proteins for endogenous protein labeling. *Nature communications* 8, 370.

99. Veggiani, G., Nakamura, T., Brenner, M. D., Gayet, R. V., Yan, J., Robinson, C. V., and Howarth, M. (2016). Programmable polyproteams built using twin peptide superglues. *Proceedings of the National Academy of Sciences* 113, 1202-1207.

100. Jackrel, M. E. (2010) *Design of tetratricopeptide repeat proteins with novel binding specifications*, Yale University.

101. Speltz, E. B., and Regan, L. (2013). White and green screening with circular polymerase extension cloning for easy and reliable cloning. *Protein Science* 22, 859-864.

102. Fink, G. (2002). Guide to yeast genetics and molecular and cell biology. *Methods Enzymol* 350, 21.

103. Gietz, R. D., and Schiestl, R. H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature protocols* 2, 31-34.

104. Hoffman, C. S., and Winston, F. (1987). A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformaion of Escherichia coli. *Gene* 57, 267-272.

105. Mumberg, D., Müller, R., and Funk, M. (1994). Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Research* 22, 5767.

106. Labbé, S., and Thiele, D. J. (1999). Copper ion inducible and repressible promoter systems in yeast. *Methods in Enzymology* 306, 145-153.

107. Lee, M. E., DeLoache, W. C., Cervantes, B., and Dueber, J. E. (2015). A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synthetic Biology* 4, 975-986.

108. Longtine, M. S., McKenzie III, A., Demarini, D. J., Shah, N. G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J. R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* 14, 953-961.

109. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277-293.

110. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, P., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005). The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* 59, 687-696.

111. Huh, W.-K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., and O'Shea, E. K. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686-691.

112. Kushnirov, V. V. (2000). Rapid and reliable protein extraction from yeast. *Yeast* 16, 857-860.

113. Edelstein, A., Amodaj, N., Hoover, K., Vale, R., and Stuurman, N. (2010). Computer control of microscopes using μManager. *Current Protocols in Molecular Biology*, 14.20.11-14.20.17.

114. Karasawa, S., Araki, T., Yamamoto-Hino, M., and Miyawaki, A. (2003). A green-emitting fluorescent protein from Galaxeidae coral and its monomeric version for use in fluorescent labeling. *Journal of Biological Chemistry* 278, 34167-34171.

115. Crane, M. M., Clark, I. B., Bakker, E., Smith, S., and Swain, P. S. (2014). A microfluidic system for studying ageing and dynamic single-cell responses in budding yeast. *PloS one* 9, e100042.

116. Sivia, D., and Skilling, J. (2006) *Data analysis: a Bayesian tutorial*, OUP Oxford.