An investigation of proteoforms in health and disease using peptide-level readouts

Sophie Moggridge

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Judit Villén, Chair

Devin Schweppe

Christine Queitsch

Program Authorized to Offer Degree:

Genome Sciences

University of Washington

Abstract

An investigation of proteoforms in health and disease using peptide-level readouts

Sophie Moggridge

Chair of the Supervisory Committee:
Judit Villén
Department of Genome Sciences

Proteins are the biomolecules that drive the functions of life. To fully appreciate and understand the role of protein diversity in health and disease, we must seek a deeper understanding of proteoforms - the molecular variants of canonical proteins. Although humans have around 20,000 protein-coding genes, millions of proteoforms arise through mutations, splicing, and post-translational modifications. My thesis work focuses on two proteoform types: missense mutations and phosphorylation. In Chapter 2, I demonstrate the utility of pooled mass spectrometry (MS)-based assays to measure solubility and thermal stability of missense mutations. Using ten disease-causing mutants of the human phosphoglucomutase 1 (PGM1) protein, I achieved improved resolution using our pooled MS assays compared to previously published studies that relied on individually purified mutants. Scaling of this approach to larger mutant libraries and diverse biochemical assays will significantly enhance our understanding of how missense mutations affect protein function and contribute to variant classification in disease contexts. Chapter 3 discusses proteoforms generated by phosphorylation, a post-translational

modification that enables proteins to rapidly and reversibly alter their properties and functions. I investigate the phosphorylation signatures in response to osmotic, heat, and oxidative stresses. I identify shared and stress-type specific phosphorylation signatures that align with previously reported data. I also identify phosphorylation sites on proteins known to localize to stress granules, providing a candidate list of stress granule phosphorylation sites for mechanistic investigation. Together, these chapters highlight the power of MS-based methods for characterizing proteoforms and their roles in health and disease.

TABLE OF CONTENTS

LIST OF FIGURES & TABLES

DEDICATION

To my lab partner in this joyous and messy experiment that is life, Pat Salembier.

ACKNOWLEDGEMENTS

I am immensely grateful to those who have supported me throughout my scientific journey. I am privileged to be supported by a community of scientists, friends, and loved ones who have given their time, energy, and expertise. This would not have been possible without you.

Firstly, thank you to the fantastic science educators who taught me in classrooms: **Sumi Aota** of Nepean High School, **Doug Briant**, **Perry Howard**, and **Greg Beaulieu** of the University of Victoria. Thank you to the research mentors who hired me as a bright-eyed undergraduate co-op education student and showed me the joy of scientific research. To **Sue Twine**, working in your lab at the National Research Council of Canada profoundly changed my life trajectory. Thank you for believing in me, for being a role model, and for your unwavering support for over ten years. Thank you to **Chris Hughes** and **Gregg Morin,** who supervised me at the British Columbia Cancer Research Centre who taught me to think creatively, dream bigger, and explore graduate school options outside of Canada.

For the past five years, I had the privilege of working in the lab of **Judit Villén**. Judit, thank you for teaching me how to think like a scientist, think independently, critically, and outside the box. Throughout my time in the lab, you have encouraged me to develop my research ideas and directions. Thank you for trusting me and being patient. Thank you for your scientific guidance and supervision and for providing the resources and funding for my work in the lab. I am forever grateful for the training I received from you.

From 2019-2025, I had the privilege of working with many brilliant scientists in the Villén Lab. **Ricard Rodriguez,** you have a joy and excitement for brainstorming that I will carry wherever I go. Thank you for cheering on my weird ideas and tangents. And thank you for your endless patience with instrumentation. To **Alexis Chang**, the ultimate hype-girl and brainiac,

thank you for being by my side at the bench for the past few years. Your willingness to help and inquisitive spirit will always be treasured. To **Matt Berg**, thank you for the scientific feedback, guidance and friendship. To the graduate students and postdocs who welcomed me like family into the Villén lab - **Ian Smith**, **Bianca Ruiz**, **Anthony Villente**, **Kyle Hess, Alex Hogrebe** and **Mario Leutert** - thank you. Your encouragement, friendship, and advice are deeply appreciated.

Thank you to **Devin Schweppe**. You have profoundly shaped my scientific identity and values during graduate school. You have taught me to think critically and stay positive. I will always treasure the kindness and generous spirit you showed me, and I aim to emulate that wherever I go.

Thank you to **Christine Queitsch** for providing infectious energy and excitement about my projects, especially when I was down. Thank you for taking me seriously and for advocating for me and the graduate students in the department as the co-director of the graduate program.

Thank you to my thesis committee; Devin Schweppe, Christine Queitsch, **Brian Beliveau**, and **Jesse Zalatan** - your feedback, critiques and advice have been constructive. To all the Genome Sciences Department members, including the 2019 GS cohort, support staff, IT and admin - thank you for welcoming me into this special community with open arms.

Thank you to my best friend, **Katie Fullmer,** who I have the pleasure of laughing with, holding mitts for, and punching every Saturday at Emerald City Boxing Gym. Katie, you are are a thoughtful, loyal and fun friend. Thanks to the wonderful community of boxers at **Emerald City Boxing Gym,** who have helped me stay sane and made Seattle feel like home. Thank you to my longtime best friend and undergraduate lab partner from UVic, **Darcy Sutherland**. Your steadfast support and encouragement have been outstanding.

To my supportive parents, **Angela Walker** and **Tim Moggridge**, thank you for always nurturing and encouraging my curiosity. I am privileged to be your daughter and grateful for all you have sacrificed for me. Thank you for giving me cases of Hornby Island chocolate espresso power bars to fuel long days (and overnights) in the lab. To my late grandfather and lifetime learner, **John Alexander Walker**. Your Einstein bobblehead sits on my lab desk as a daily reminder of your kindness and excitement for science every day. Thank you for always sharing how proud you were of me. I take great comfort knowing that if you were still here. you would have read this thesis cover to cover and come to me with highlighted notes and follow-up questions. To my in-laws, **Chantale Courcy** and **Gerry Salembier**, thank you for the relentless cheerleading and encouragement.

Above all else, thank you to my partner, **Pat Salembier**. I love you, Bub. Half of this PhD truly belongs to you for your support and love over the past six years. Moving countries and jobs takes a lot of courage, even without a global pandemic. These six years have been challenging, but we have coevolved to be stronger and more resilient. Thank you for the grocery and boba deliveries to the lab on the weekend, for the late-night pickups, for drying my tears, and for hyping me up on the darkest days. I am so excited to see what adventures life has for us and our little cat, Smudge.

CHAPTER 1. INTRODUCTION

1.1 Moving from the proteome to proteoforms

There are 20,000 protein-coding genes in the human genome, but when variation at the levels of DNA (mutations), RNA (splicing, mistranslation, alternative start sites), and protein (post-translational modifications), are taken into account, the estimated number of proteoforms expands to the millions(Aebersold et al., 2018). Different proteoforms of the same protein can adopt different conformations and have different functions. Functional changes afforded by proteoforms are critically important for our understanding of cellular processes and disease, thus we must develop methods to survey proteoforms. However, the detection and characterization of proteoforms remains challenging. This necessitates the development of methods to identify and characterize proteoforms on a large scale using unbiased methods such as mass spectrometry (MS).

1.2 Analysis of proteoforms arising from missense mutations

A common type of proteoform is created by a missense mutation, whereby the canonical/ consensus amino acid is replaced by a different amino acid at a given position within the protein sequence. The consequence of a missense mutation depends on its position within the protein sequence, and the biochemical differences between the original and substituted amino acids. A missense mutation can have no effect (benign), or it can affect a variety of different protein attributes, including structure, function, and/or localization. Depending on the severity of the missense mutation, they can also potentially contribute to protein dysfunction and disease.

Unfortunately, the functional significance of most missense mutations in the human genome is unknown, and this adversely affects health outcomes(Starita et al., 2017).

Previously, patient cohorts and family histories were used to determine variant effects. However, patients with novel or rare variants cannot receive the rapid diagnosis and tailored interventions afforded to those with large cohorts or detailed family histories. Variant effect annotations (pathogenic, benign, unknown significance), can be used by clinicians to diagnose, treat, and predict outcomes. Furthermore, the mechanism of variant dysfunction can provide a rich dimension of actionable information. Mechanisms of dysfunction can be represented by changes in biophysical properties (such as solubility, stability, modifications, localization, etc.) compared to the wild-type protein. For example, the treatment of cystic fibrosis is dependent on the type of variant dysfunction(Southern et al., 2023). Mutations to protein encoded by the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) gene can cause disruptions to salt homeostasis via several mechanisms of dysfunction. Depending on the position of the mutation, dysfunction can arise via issues with stability at the membrane, channel gating, conductance defects - for which there are different classes of drugs available(Southern et al., 2023). Variant-specific therapies are employed to treat a diversity of diseases from cancer to sickel cell anemia and beyond. Mechanisms of dysfunction can be critically important for tailoring treatments and developing new therapeutics. For these reasons, new technologies are needed that enable large-scale functional analysis of protein variants.

Until the introduction of deep mutational scanning (DMS) by (Fowler et al., 2010) over 15 years ago, variants were assessed using biochemical assays that often required purification of individual variants (classical biochemical assays, alanine scanning ect). These methods were time-intensive and laborious. However, in the past 15 years DMS has contributed to tremendous

strides in the high throughput variant phenotyping using sequencing-based methods of variants in dozens of disease-relevant human proteins such as BRCA1, PTEN, and TP53(Funk et al., 2025; Matreyek et al., 2018; Starita et al., 2015). While these assays provide information on pathogenicity, they cannot report protein-level properties such as mechanisms of dysfunction or protein inactivation (impairment in folding or interactions etc). In contrast, classical biochemical assays, while informative, are low throughput and often require protein purification.

1.3 Recent advances in Mass Spectrometry to analyze proteoforms

Conveniently, there are many MS-compatible biochemical assays that report on protein properties, such as solubility(Sridharan et al., 2022), thermal stability(Savitski et al., 2014), small molecule binding(Martinez Molina et al., 2013), protein-protein interactions(Roux et al., 2012), etc. These assays can potentially be scaled to library sizes comparable with DMS. In addition, there are many ways to create protein variants. These include cloning variant libraries for all possible mutations within a target proteins or even inducing errors in protein translations, through a process known as mistranslation, to create protein variants proteome-wide in a single sample(Berg et al., 2019; Cozma et al., 2023; Rodriguez-Mias et al., 2022; Zimmerman et al., 2018). MS is a powerful technique that can provide protein-level readouts. It also offers insight into biochemical protein properties and modifications.

1.4 Importance of proteoforms that arise from phosphorylation

Proteoforms can also arise via post-translational modifications (PTMs) such as phosphorylation, glycosylation, ubiquitinylation, acylation, and methylation, among others(Aebersold et al., 2018). The addition of chemical groups to a protein via PTMs serves to dynamically regulate cellular processes and proteome organization. For example, the addition of ubiquitin group(s) on lysine residues can alter protein properties and alter the cell cycle control, signalling, autophagy, or proteasomal degradation(Damgaard, 2021). Glycosylation results from the addition of polysaccharide groups to proteins. Glycosylation can protein properties such as protein folding, stability, and interactions, and is critical for cellular adhesion, immune response, and signalling(He et al., 2024).

Phosphorylation occurs when an enzyme called a kinase catalyzes the transfer of a phosphate group from adenosine triphosphate to the amino acid side chain of a target protein(Ramazi & Zahiri, 2021). Phosphorylation primarily occurs at serine, threonine, or tyrosine (Ser, Thr, Tyr) residues in eukaryotes. This process involves the replacement of a neutral hydroxyl group with a negatively phosphate group in the amino acid side chain. The phosphate is also more hydrophilic and polar than the hydroxyl group. The biochemical changes afforded by phosphorylation serve as a rapid, reversible, and dynamic mechanism to alter protein properties without altering protein sequence. Phosphorylation can alter the stability, folding, activity, localization, or interactions of a protein(Bilbrough et al., 2022). For example, in an intermediate signalling protein called LYN kinase, phosphorylation of Tyr397 in the activation loop causes a conformational change in the ATP binding pocket(Xu et al., 1999) and also reduces the protein's thermal stability(Potel et al., 2021). Certain protein domains such as WW, forkhead, and

Src-homology 2 recognize phosphorylated amino acids to enable protein binding interactions and formation of higher order protein complexes.

Phosphorylation can also alter non-specific interactions between proteins (and other biomolecules). Hyperphosphorylation of residues located in intrinsically disordered protein regions can promote (and disrupt) protein-nucleic acid or protein-protein interactions that drive liquid-liquid phase separation (LLPS). For example, a nucleolar protein called Ki-67 has a disordered region (the central repeat domain) that has weak alternating charge blocks during interphase. However, during mitosis these disordered regions are hyperphosphorylated which results in repeated and alternating charge blocks that promote phase separation. The hyperphosphorylated Ki-67 phase separates and localizes to the chromosome periphery to prevent chromosome coalescence during mitosis(Yamazaki et al., 2022).

Additionally, phosphorylation serves as a method to rapidly transmit cellular signals through a process called phosphorylation signalling cascades. In this process, a stimulus causes the activation of the first kinase protein, the first kinase then phosphorylates a second kinase protein, the second kinase becomes activated and phosphorylates a third kinase and so forth. This serial propagation of phosphorylation signal continues until it reaches an effector protein that carries out a cellular response. A well-characterized example of this occurs when epidermal growth factor (EGF) binds a transmembrane receptor protein called epidermal growth factor receptor (EGFR, HER1). Binding of EGF to the EGFR (monomer or preformed dimer) causes a conformational change and rotation of the intracellular kinase domains within the EGFR homodimer(Purba et al., 2017). In this new position, the kinase domains phosphorylate specific tyrosine residues within the c-terminal protein tail(Yarden & Schlessinger, 1987). The combination and positions of the phosphotyrosine residues are recognized by different adaptor or

enzyme proteins which in turn activate one of several signalling pathways such as the ERK-RAS-MAPK, PI3K-AKT, PLCγ1-PKC, SRC, or JAK-STAT pathways(Wee & Wang, 2017). These pathways terminate in effector proteins and result in regulation of numerous biological processes such as growth, proliferation, cell cycle, calcium signalling, survival and more(Wee & Wang, 2017).

Phosphorylation signalling cascades are critical infrastructure for transmitting information and signals throughout the cell. Cascades are equally as important for adapting to changes in the cellular environment and for normal cellular functioning. Dysregulation of signalling cascades occurs in many diseases. For example, mutations causing the constitutive activation of a HER2 (ERB2) is a driving oncogenic mutation in several types of cancer(Cocco et al., 2019) and is the target of numerous chemotherapeutics. Insight into the phosphorylation status of proteins can serve as a dynamic snapshot of how that protein contributes to processes within the cell. Since phosphorylation can be specific to cell type, environment, stress conditions, age etc., there is ongoing motivation to understand differences in phosphorylation states of all proteins.

1.5 Historical methods in phosphoproteomics

The field of protein phosphorylation was largely pioneered here at the University of Washington by Krebs and Fisher, who discovered that reversible protein phosphorylation is a regulatory mechanism(Fischer & Krebs, 1955; Krebs et al., 1958). The pair later received the Nobel prize in 1992 for that discovery. Much of that and subsequent work of subsequent decades utilized radiolabeled phosphorus P^{32} for visualization and study of phosphoproteins. The advent

of phosphorylation-specific antibodies in the 1980s enabled detection of phosphorylated proteins using western blots and, eventually, protein arrays. Since the advent of MS-based phosphoproteomics in *S. cerevisiae* in 2002 by (Ficarro et al., 2002), several advancements in sample preparation, instrumentation, annotation, and computation have enabled improved detection and quantification of phospho-sites.

The main challenge in phosphoproteomics is the relatively low abundance of phosphorylated peptides compared to unphosphorylated peptides. Reproducible detection from small sample amounts has been achieved using bead-based phosphoproteomic enrichment strategies such as immobilized metal affinity chromatography(Zhou et al., 2013), metal-oxide affinity chromatography(Gates et al., 2010) and optimization thereof(Arribas Diez et al., 2021; Tsai et al., 2014). Automation of bead-based phosphopeptide enrichment methods(Bortel et al., 2024; Chang et al., 2023; Leutert et al., 2019; Tape et al., 2014) offer vast improvements and eased sample handling for large experiments. Advancements in stable isotope reagents (such as stable amino acids in cell culture(Ong & Mann, 2007), and Tandem Mass Tags(Dayon et al., 2008) have enabled multiplexed quantitative comparison of up to 35 samples in a single injection(Zuniga et al., 2024). Advancements in instrument acquisition types such as data-independent acquisition (DIA), multiple reaction monitoring (MRM), and parallel reaction monitoring (PRM), along with improved computational tools for scoring positional isomers such as Thesaurus(Searle et al., 2019), AScore(Barente & Villén, 2022; Beausoleil et al., 2006), PTMProphet(Shteynberg et al., 2019) have increased phosphopeptide identifications, sensitivity and localization. Together, these improvements have created several tour-de-force phosphorylation datasets in diverse sample types and conditions(Leutert et al., 2022; Ochoa et al., 2020; Potel et al., 2018).

However, as the identification and quantification of phospho-sites has drastically increased in the past two decades, we are in the dark as to the function, regulation and annotation of most of the phosphoproteome. Less than 5% of reported phosphosites have identified kinases, and less than 0.1% of phospho-sites have annotated functions(Needham et al., 2019). One way scientists have tried to illuminate the dark phosphoproteome is through performing large perturbation experiments in diverse sample types and conditions This endeavour to uncover functions of the phosphoproteome will be critical for understanding cellular processes in health and disease states.

1.6 Towards elucidating the role of phosphorylation in stress granules

One fascinating example of a cellular compartment dynamically organized by PTMs is the stress granule, or SG(Hofweber & Dormann, 2019). SGs are a type of supramolecular condensate that forms as a result of stress, whereby mRNAs, translation initiation complexes, and other proteins within the cytosol partition to form condensates(Collier & Schlesinger, 1986; Kedersha et al., 2005). SGs benefit cells by sequestering and protecting proteins and mRNAs during stress conditions or by sequestering signalling proteins. Sequestration inhibits signalling pathways, altering signalling and metabolism in the cell under stress(Kedersha et al., 2013). SG dysregulation is implicated in a variety of diseases, which include viral infections, cancer metastasis, ALS, and Alzheimer's disease(Riggs et al., 2020). Despite the importance in stress response, there is limited understanding of SG dynamics and function.

The protein composition of SGs is stress-type specific, but the regulatory mechanisms by which proteins are included/excluded from granules in different conditions remain unclear.

MS-based proximity labelling techniques such as APEX and BioID(Markmiller et al., 2018; Youn et al., 2018) have been used to identify proteins localized to stress granules. There is a database of RNA-condensate proteins that amalgamates dozens of studies with various types of genomic and proteomic evidence(Millar et al., 2023). While recent studies have shown that some PTMs such as phosphorylation, acetylation, methylation(Hofweber & Dormann, 2019) drive SGs formation and dissolution, no one has systematically studied phosphorylation in SGs in multiple stress types. Previous examination of SG phosphorylation has been site-specific, and thus required prior knowledge of sites to be measured. Fluorescence microscopy has been the primary method of interrogation, but it is limited by the availability of PTM-specific antibodies and multiplexing capacity. Other PTMs such as ubiquitination have been examined in multiple stresses and have stress-type specific signatures(Gwon et al., 2021; Maxwell et al., 2021).

Given that ubiquitination and phosphorylation are commonly involved in PTM crosstalk, it is likely that phosphorylation is also involved in SGs. One phosphoproteomic study of stress granules was performed by(Wippich et al., 2013) and identified DYRK3 as a regulator of SG disassembly. However, no one has compared SG stress types to identify differences in phosphorylated proteins. Identification of candidate phosphosites in proteins known to localize to stress granules may provide insight into stress-type specific regulation.

1.7 Organization of this thesis

Following this introduction chapter (Chapter 1), this thesis comprises two projects employing MS to investigate proteoforms. In Chapter 2, I develop a method for determining the effect of missense mutations on protein thermal stability and solubility. Using disease-causing

variants of a human metabolic protein, phosphoglucomutase 1 (PGM1), I demonstrate how pooled MS-based assays compare to classical biochemical assays on purified protein variants. In Chapter 3, I investigate the differences in phosphorylation responses between human osteosarcoma cells (U2OS) in different stress conditions known to induce stress granules, such as heat, osmotic and oxidative stress. In addition to global analysis, I use computational tools and parse stress granule literature to determine if differentially abundant phosphorylation sites occur on proteins that localize within stress granules. Finally, in Chapter 4, I summarize my thesis work, discuss its impact, and discuss my hopes for future studies of proteoforms using MS.

CHAPTER 2. ASSESSING PGM1 PROTEIN VARIANTS USING MASS SPECTROMETRY-BASED PROTEOMICS

This chapter is based on the following preprint expected to be uploaded to BioRxiv in March/ April 2025:

Establishing mass spectrometry-based assays to assess the impact of protein variation in pooled format

Sophie Moggridge, Kyle Hess, Matthew D. Berg, Ricard A. Rodriguez-Mias*, Judit Villén*

Department of Genome Sciences, University of Washington, Seattle, USA

Department of Genome Sciences, University of Washington
*corresponding authors

2.1 Abstract

Protein variants arising from single amino acid substitutions contribute to many human diseases. These substitutions can affect protein properties such as structure, stability, activity, subcellular localization, and interactions with other biomolecules. Previous work has characterized the effects of substitutions using traditional biochemical assays with individually expressed and purified protein variants. Here, we have streamlined this laborious process by coupling biochemical assays to mass spectrometry analysis, enabling the characterization of multiple protein variants in a single experiment. To benchmark this approach, we selected a set of thirteen previously characterized missense variants of human phosphoglucomutase 1 (PGM1) and conducted protein solubility and thermal stability assays in a pooled format. Our mass spectrometry-based assays successfully recapitulate known PGM1 mutational effects on solubility and stability. Moreover, we achieve improved resolution compared to previously published results with purified variants. Future scaling of this approach to larger variant libraries combined with diverse biochemical assays will greatly enhance our understanding of how single

amino acid substitutions affect protein function and contribute to better classification of variants in the context of disease.

2.2 Introduction

Protein variants caused by single amino acid substitutions in 25% of human proteins have been documented to contribute to Mendelian disease(Amberger et al., 2019), many of which are monogenic((Apgar & Sanders, 2022; Kroncke et al., 2015). However, we have a limited understanding of how each possible mutation within a protein affects protein function. Mechanisms of protein dysfunction are diverse, and substitutions can affect many protein properties including structure, stability, activity, subcellular localization, and interactions with other biomolecules. Indeed, understanding the impact of mutations on protein function is used to guide treatment in diseases as diverse as breast and ovarian cancers (BRCA1/2 gene)(Dorschner et al., 2013), cystic fibrosis (CFTR gene)(Dickinson & Collaco, 2021), and Li-Fraumeni Syndrome (TP53 gene)(Fortuno et al., 2021), for example. Unfortunately, more than three quarters of all variants in ClinVar are classified as variants of unknown significance (VUS)(Fayer et al., 2021). Mutations with ambiguous mechanisms of dysfunction pose a challenge for therapeutic intervention and treatment. To gain a deeper understanding of mechanisms through which missense mutations cause dysfunction, protein-centric assays that explore multiple protein properties conducted for thousands of protein variants simultaneously are needed.

High-throughput screening methods like deep mutational scanning(Fowler et al., 2010) have accelerated variant phenotyping. However, these methods do not measure protein variants

directly, thus limiting the protein properties that can be assayed and, therefore, deep mechanistic understanding. Conversely, traditional biochemical assays allow direct protein measurements but require protein purification and are low throughput. Biochemical assays that assess protein properties, including stability, post-translational modifications, and ligand binding, are compatible with mass spectrometry (MS) and can facilitate protein-centric analysis of mutational libraries. Previously, we applied these MS-based biochemical assays in combination with protein mistranslation to assay thousands of protein variants in the same lysate across the entire proteome(Rodriguez-Mias et al., 2022). In the present study, we aimed to apply these assays in a pool format to examine a well-defined set of genetically encoded protein variants within a single human protein, phosphoglucomutase 1 (PGM1).

Here we present the application of MS-based proteomics methods to determine molecular consequences of disease-causing variants of human PGM1 in a pooled format. PGM1 is a cytosolic, monomeric and highly soluble protein that catalyzes the bi-directional interconversion of glucose 1-phosphate (utilized in glycogen metabolism and important for protein glycosylation) and glucose 6-phosphate (a substrate for glycolysis and pentose phosphate pathway). Mutations in PGM1 can lead to PGM1-linked congenital disorder of glycosylation (PGM1-CDG)(Altassan et al., 2021). PGM1-CDG has multisystem phenotypes including cardiomyopathy, exercise intolerance, hypoglycemia, and physical malformations(Altassan et al., 2021). PGM1 has clinically relevant variants for which the molecular effects were determined with classical biochemical assays(Beamer, 2020; Lee et al., 2014; Wong et al., 2016) and broadly fall into two categories: those affecting folding and aggregation and those affecting catalytic activity. We focused on 13 diesase-associated missense variants that alter fundamental biochemical properties of PGM1.

The pooling of PGM1 variants in the same lysate presented several challenges that were anticipated using conventional bottom-up proteomics. As with many bottom-up experiments, it is not feasible to detect all peptides in a protein sequence using a single protease(Sun et al., 2022; Swaney et al., 2010) as peptides span a range of lengths and have variable ionization and fragmentation. Furthermore, a single amino acid substitution may introduce or eliminate protease cleavage sites or charges and significantly alter detection. Thus achieving full sequence coverage often requires the use of multiple proteases. The challenge is that while other assays may rely on the detection and quantitation of only one of several unique peptides in a protein, in our assay, each variant is represented by only one mutant peptide. Therefore the proximity to cleavage sites and ionization ability of one specific peptide dictates whether we can quantify the variant, as there are no other unique peptides for that variant. The second challenge in measuring a pool of protein variants is that the signal from the highly abundant wild-type peptides overwhelms the informative mutant peptides, impairing their detection. For example, if a protein produces sixty tryptic peptides, only one peptide contains the amino acid substitution while the remaining 59 peptides are identical to wild type. This issue exponentially grows with the size of the variant library, as the non-mutated peptides that are shared between variants are much more abundant than a single mutant peptide. Taken together, these challenges affect signal-to-noise ratios, potentially affecting peptide detectability, quantifiability, and ultimately the number and types of mutations that can be covered in a mutational library. To circumvent these detection challenges for PGM1 variants from a pooled library, we employed multiple digestion enzymes and implemented a TMT-compatible targeted MS acquisition method.

Here, we present an MS-based proteomics framework for profiling libraries of protein variants in a single experiment. As a proof of concept, we apply this framework to profile the

molecular consequences of multiple disease-causing variants of human phosphoglucomutase 1 (PGM1) in a pooled format. Using a pooled recombinant DNA library of PGM1 variants expressed in *Escherichia coli*, we applied MS-based solubility and thermal stability assays to characterize the variants. We benchmarked our data against the previously published biochemical measurements and found good agreement. Our work supports the idea that pooled MS-based assays can be applied to investigate mechanisms of protein variant dysfunction and serve as a stepping stone toward more complex libraries.

2.3 Results and Discussion

To characterize the effects of mutations on PGM1 function, we employed mass spectrometry to measure the solubility and thermal stability of 13 missense variants of PGM1. These thirteen variants are pathogenic and cause PGM1 Congenital Disorder of Glycosylation(Altassan et al., 2021).These variants were selected as they have been purified and biochemically characterized, making them fantastic benchmarks for our assays. The general workflow of the sample preparation and assays is depicted in Figure 2-1A. In this workflow, a library of *E. coli* strains, each expressing a PGM1 variant, was grown in a pooled culture. Biochemical assays were conducted at the protein level and the variant peptides containing the single amino acid substitution were measured by MS. Variant peptide abundance was representative of the biochemical property of the variant protein. A sequence map illustrating protein features of PGM1 and mutated residues in the variants we studied is provided in Figure 2-1B.

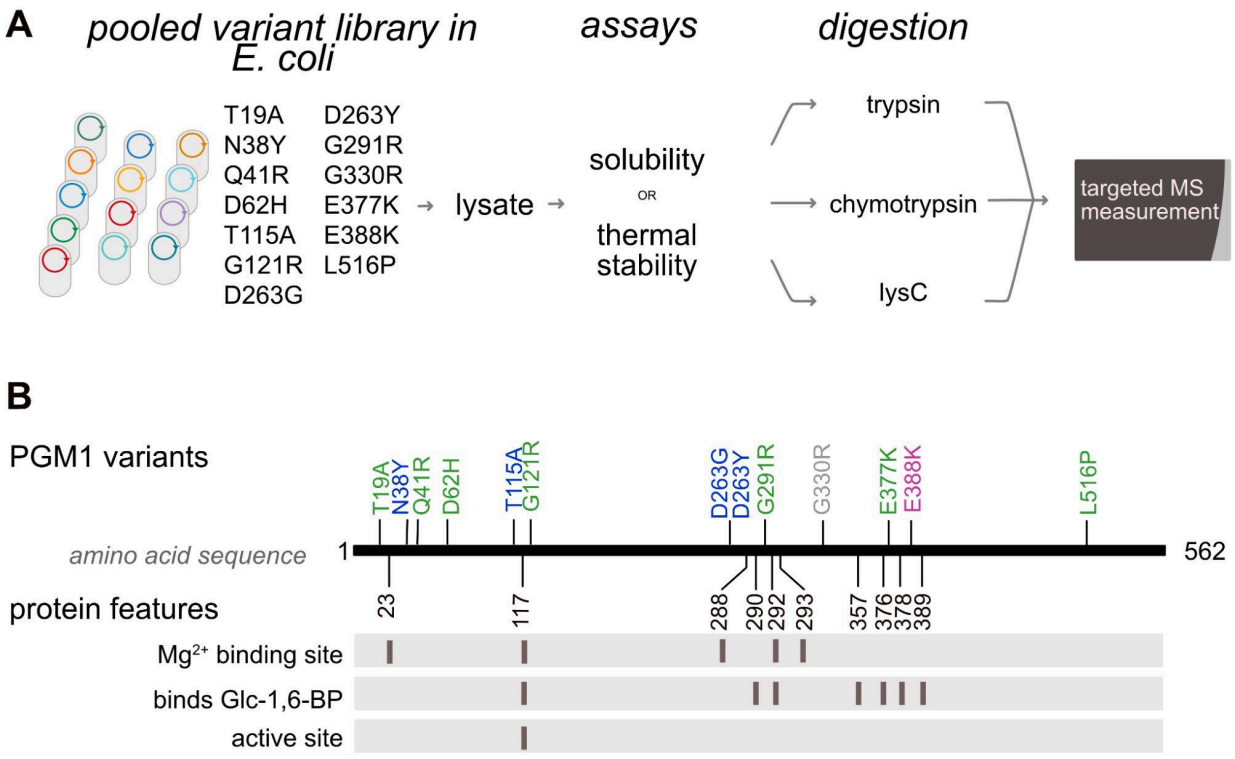


FIGURE 2-1. Exploring disease-associated PGM1 variants. ***(A)*** Experimental workflow for investigating the molecular mechanisms of variant dysfunction using pooled MS assays. A pooled library of 13 missense variants of PGM1 is expressed in *E. coli*, lysed, and assayed for protein solubility and thermal stability. After each biochemical assay, samples are split in three aliquots and each is digested with one protease (trypsin, chymotrypsin, or LysC). Resulting peptides are labeled with TMT11 isobaric mass tags and analyzed by MS. ***(B)*** Representation of PGM1 sequence highlighting protein features and mutation positions of the 13 variants analyzed. The colour of the variant represents the protease required for detection in our MS assays; trypsin (green), chymotrypsin (blue), LysC (magenta), and undetectable (grey).

2.3.1 Library characterization and variant detection

A typical TMT workflow with tryptic digestion, TMT-labelling, and data-dependent acquisition with standard SPS-MS3(McAlister et al., 2014) methods did not yield data that could be consistently matched to all PGM1 variants (data not shown). To increase coverage and quantification accuracy, we implemented digestion with multiple proteases and a targeted MS acquisition method. The library of protein variants was first assessed for coverage using trypsin digestion, and 7 out of 13 variants were reproducibly detected. Detection of additional variants

was achieved using chymotrypsin (two variants) and LysC (none). The utility of LysC was to improve the detection of E388K. However, this proved difficult. Therefore, using these three digestion enzymes we could detect 9 out of the 13 total variants. We used a combination of DDA and targeted methods to identify the most intense and reproducible peptide precursors and fragment ions for each variant. While G330R, D263G, and D263Y should theoretically generate detectable peptides by chymotrypsin digestion, those peptides were not confidently identified despite targeting multiple charge states. We moved forward with the nine detectable variants for all further analyses.

Given the successful identification of 9 out of 13 variant peptides in our library, we combined peptides after digestion with each of the enzymes (trypsin, LysC, or chymotrypsin). Quenched digests were mixed together and measured as a single MS sample, which enabled the detection of nine mutant peptides. While our approach allows us to measure many variants, we are unable to report for the wild-type protein as all the peptides are shared across all variants. In future studies, we propose to include measurements of the wild type protein by spiking in isotopically labelled wild-type protein.

2.3.2 Implementation of a targeted MS acquisition method for the analysis of variants in pooled biochemical assays.

We applied isobaric labelling with tandem mass tags (TMT) to multiplex the measurement of variants across multiple conditions as required by the biochemical assays (temperatures, replicates, buffers) in a single run. Mass spectrometry analysis with the commonly used TMT-MS3 acquisition method for TMT-tagged peptides(Erickson et al., 2017; McAlister et al., 2014) (Supplementary Figure A1) failed to detect most of the PGM1 variant

peptides (Supplemental Figure A2). This failure could be due to the extremely low abundance of variant peptides relative to their wild-type counterparts. Previous work has shown that the improved sensitivity of targeted MS methods can also be applied to TMT-labeled peptides(Erickson et al., 2017); (Yu et al., 2023). As we were successful with most variants using a targeted assay, we chose to develop a targeted MS acquisition approach for the TMT-labelled variants. This method, which we call Targeted-SPS-MS3 (Supplementary Figure A1), serves as a bridge between the sensitivity of PRM methods and the multiplexed quantification of SPS-MS3 acquisition methods. Upon comparison, our Targeted-SPS-MS3 method yields far more variant peptide scans per run, with significantly less MS2 interference in each isolation, compared to conventional SPS-MS3 methods (Supplementary Figure A2).

2.3.3 Solubility of PGM1 variants

Once we established the MS acquisition method (Targeted-SPS-MS3), we conducted several MS-based biochemical assays on the protein variants. We first assayed variant solubility, which serves as a proxy for protein folding and aggregation state, by quantifying the amount of soluble protein in various lysis buffers(Györkei et al., 2022; Listwan et al., 2009; Sridharan et al., 2019). We quantified all nine variants of interest and reported their solubility ratios, which represent the relative variant peptide abundance extracted from *E. coli* under gentle lysis versus denaturing conditions (Figure 2-2A). Solubility ratios spanned a range of 0-0.8. We reproducibly resolved highly soluble variants (T19A, T115A, G121R) from less soluble variants (N38Y, Q41R, E377K, L516P) (Figure 2-2B).

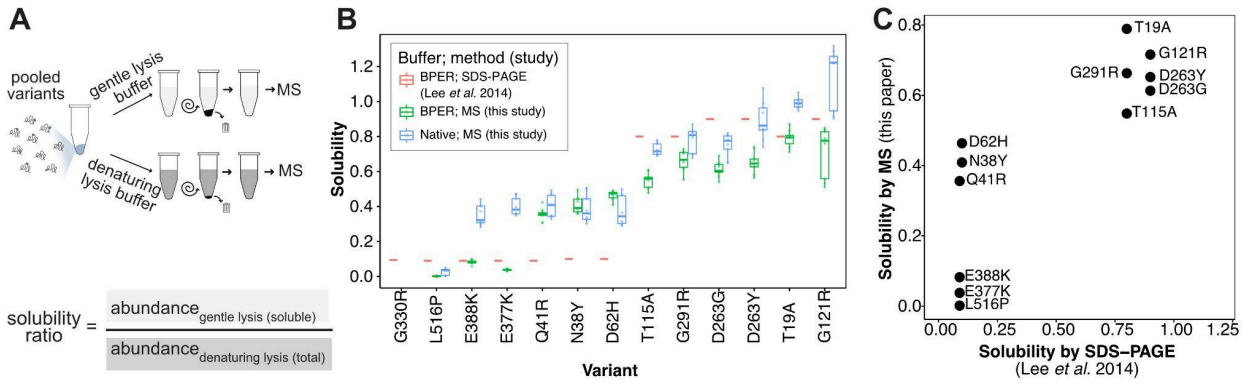


FIGURE 2-2. Protein solubility of PGM1 variants. ***(A)*** Protein solubility assay workflow. Cells were lysed in either a gentle (non-denaturing) lysis buffer to obtain the soluble protein fraction or a denaturing lysis buffer to obtain the total protein fraction. Lysates were centrifuged and the abundance of each variant in the supernatant was measured by MS. A solubility ratio was calculated based on the abundance of each variant in the gentle lysis buffer relative to the denaturing lysis buffer. ***(B)*** Benchmarking pooled MS assay against a purified protein assay (Lee et al., 2014) to assess the solubility of 13 PGM1 variants. In our study, a pool of *E. coli* each expressing a different PGM1 variant was lysed in either a proprietary native buffer (BPER; green) or a PBS-based native buffer (blue). We conducted three replicates and represented the top 3 most intense scans for each peptide variant as data points. ***(C)*** Scatter plot comparing average PGM1 variant solubility measurements in our study by MS (y-axis) and an SDS-PAGE-based assay on purified variants (x-axis) by (Lee et al., 2014). Both assays were performed using BPER buffer.

We compared our MS solubility results to previously published measurements from Lee et al. (2014). In that study each PGM1 variant was extracted in the same BPER buffer, but the relative abundance of the soluble and insoluble fractions were determined using SDS-PAGE and densitometry. Our results showed high correlation with these previously published data (Figure 2-2C, Spearman correlation R = 0.93). However, our MS-based method provided greater resolution of solubility values. This increase in resolution is expected given the higher quantitative capability of mass spectrometry compared to gel band densitometry.

Next, we wanted to explore how the buffer composition affects solubility. Given that the BPER non-denaturing buffer composition is proprietary, we performed the same solubility assay in a PBS-based lysis buffer with a defined composition referred to as 'native lysis buffer' (Figure 2-2B). The variant with a glutamic acid to lysine mutation had dramatically lower solubility in the BPER buffer than the native lysis buffer ($E337K_{BPER}$= 0.11, $E377K_{Native}$ = 0.40). It is possible

that the native buffer mitigates the effects of the charge switch in the active site, whereas BPER does not, resulting in higher variant solubilities. Without knowing the BPER composition, it is difficult to speculate how solubility is differentially affected in these buffers. Taken together, our findings illustrate that additional information can be gathered by assaying protein variants in multiple buffer conditions.

2.3.5 Thermal stability of PGM1 variants

Next, we measured the thermal stability of each variant using the thermal proteome profiling method(Savitski et al., 2014) as another dimension of protein stability. The method measured the extent to which heat treatment causes individual proteins to unfold and precipitate out of solution. Lysate from a pool of *E. coli* expressing PGM1 variants was heated to a range of commonly used temperatures (Figure 2-3A). We then separated the soluble protein fraction, quantified the abundance of each variant at each temperature condition, and constructed sigmoidal melting curves(Franken et al., 2015). We determined the melting temperature (Tm) for each variant, defined as the temperature at which 50% of the variant protein is soluble.

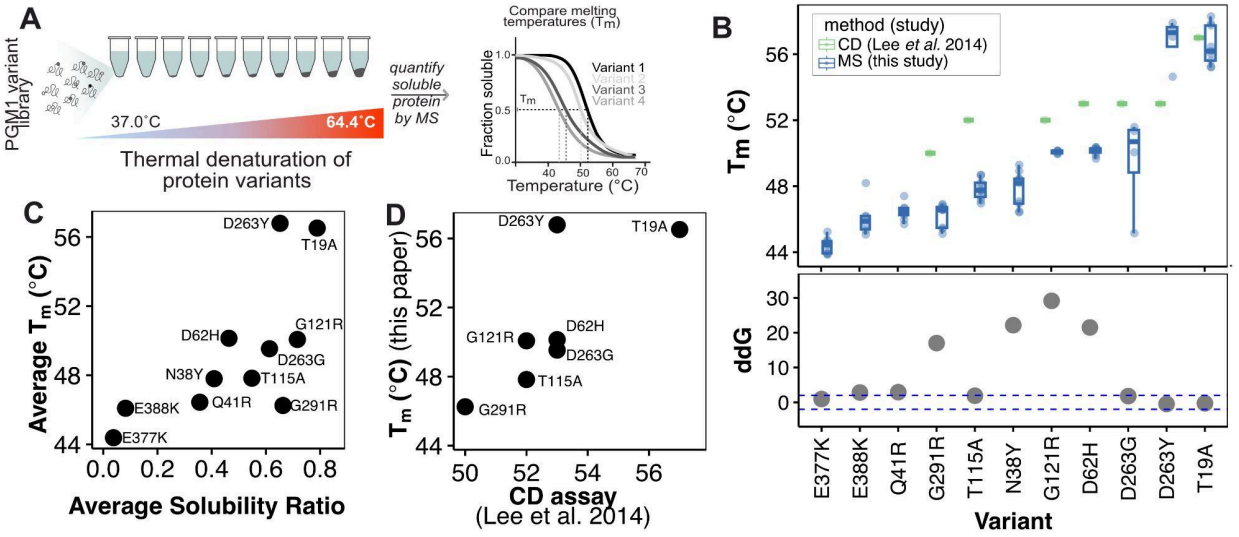


FIGURE *2*-3. Protein thermal stability of PGM1 variants. (*A*) Workflow for characterizing thermal solubility of protein variants. Lysate from the variant library was diluted to 2.5 mg/mL and aliquoted into ten tubes, each subjected to a different temperature for five minutes. Soluble protein was extracted from each tube and MS was used to quantify the variants across the temperature range. Melting temperature (Tm) was determined from melting curves as the temperature at which half the protein variant was soluble (as reported by the corresponding variant peptide). This assay was performed in triplicate. ***(B)*** Thermal stability and change in change in Gibb's Free Energy **ΔΔ**G for protein variants. Tm measurements of pooled PGM1 variants (blue) benchmarked against data from thermal denaturation circular dichroism performed on individually purified variants ((Lee et al., 2014); green). For each of the three replicates, the top three3 most intense variant scans were selected and are represented as points. Predicted **ΔΔ**G between the unfolded and folded states in the mutant versus wildtype protein ($\Delta\Delta G_{Mut \rightarrow WT}$) calculated using mutfunc(Wagih et al., 2018) ***(C)*** Scatter plot comparing average PGM1 variants thermal stability measurements from a pooled MS assay (this study) versus those made from individually purified variants measured by thermal denaturation circular dichroism (Lee et al., 2014). ***(D)*** Scatterplot comparing the average measurements obtained by MS for the thermal stability and solubility assays.

We obtained Tms for nine out of ten detectable PGM1 variants, from 44°C to 57°C (Figure 2-3B). We classified PGM1 variants into three groups according to their thermal stability: stable (Tm > 53°C) T19A; moderately stable (Tm between 49°C and 53°C) D62H, G121R; and not stable (Tm < 49°C) T115A, N38Y, G291R, Q41R, E377K. Variant L516P showed low solubility at room temperature (as shown in the solubility assay, Figure 2-2B).

One metric commonly used to predict the effect of a point mutation on protein stability is the change in the difference of Gibb's free energy between the unfolded and folded state in the wild-type versus the mutant protein ($\Delta\Delta G_{Mut \rightarrow WT}$). We calculated the $\Delta\Delta G_{Mut \rightarrow WT}$ of all PGM1

variants and found only partial agreement with the previously determined Tms(Lee et al., 2014) (Figure 2-3B). This highlights the importance of conducting experimental determinations of protein stability to assess variant dysfunction.

We compared the thermal stability results from our pooled variant MS-based method to published data obtained by circular dichroism spectroscopy (CD) on purified variants(Lee et al., 2014). The relative ranking of variant thermal stabilities was similar between the two methods (Spearman correlation r = 0.975, Figure 2-3C). We found that despite the expanded range of temperatures used in the CD assay, the reported CD Tms spanned a narrower range (50 to 57°C) than obtained with our MS-based assay (44 to 57°C). The differences between the two assays may be due to differences in analytical technique, buffer composition, and/or protein concentration.

The MS-based assay was able to better detect difficult-to-purify variants, such as E377K, Q41R, and N38Y, which were not measured by CD due to their low expression and propensity to aggregate. Furthermore, our MS-based assay differentiated the Tms of several variants, which were previously reported to be identical by CD. We speculate the Tm difference between methods was due to the addition of 2 mM magnesium in our buffer versus no magnesium in the previously reported CD measurements.

We also observed Tm differences for variants with mutations near the catalytic phosphoserine (S117) compared to CD results. We observed a Tm for the G121R variant of 50.1°C and a Tm for T115A of 48.0°C, while CD measurements reported identical Tms for both variants. T115A was shown to have increased structural disorder compared to G121R(Lee et al., 2014), but the catalytic activity of T115A was only mildly impacted compared to G121R.

Without structural data for T115A, it is difficult to pinpoint the structural features responsible for the likely differences in thermal stabilities between these variants. We belirve that exploring thermal stabilities of variants in buffers with and without PGM1 cofactor (Mg^{2+}) and substrate (glucose-1-phosphate) may uncover evidence of dysfunction relating to binding and/or catalysis.

As expected, we found good correlation between Tm and solubility measurements for the PGM1 variants (Spearman correlation R = 0.667; Figure 2-3D). However, the observed differences for the D62H and G291R variants suggest that the assays can be complementary. The D62H variant has moderate thermal stability (Tm_{D62H} = 50.2°C), but relatively low solubility (Sol_{D62H_BPER} = 0.46). Conversely, G291R has moderate solubility (Sol_{G291R_BPER} = 0.67), and low thermal stability (Tm_{G291R} = 46.3°C).

2.4.6 Application of MS methods to mutational scanning of proteins

Proteins harbor diverse functions and there are several ways in which a single amino acid mutation can alter a protein's structural and biochemical properties, causing dysfunction. MS-based assays afford many improvements to characterize protein variants compared to classical biochemical or genomics methods. Compared to biochemical methods, MS allows for the simultaneous measurement of an entire variant library without purification. Compared to mutational approaches that rely on DNA sequencing, MS-based assays provide a direct readout of protein properties and mechanism of dysfunction rather than requiring assays that maintain the genotype-phenotype link.

Even at the scale of nine protein variants, the pooled MS-based method we present here offers significant utility compared to classic biochemical assays on individual variants. Expression and measurement of variants in a pooled format saves considerable time and money

compared to purification. Furthermore, our MS-based assays have higher sensitivity than circular dichroism or SDS-PAGE densitometry, so much less material is required. This improved sensitivity affords the measurement of poorly soluble proteins that cannot be purified and measured by circular dichroism. Furthermore, by mass spectrometry one can collect many measurements of mutant peptides, whereas the previously described low-throughput methods take only a single measurement. Together the benefits provide a higher throughput method for determining mechanisms of protein variant dysfunction. One could imagine prioritizing variants using VAMP-seq(Matreyek et al., 2018) and then running several pools of a dozen variants through our MS-based assays.

Despite the many advantages of MS-based assays for protein mutational scanning, considerable challenges remain for wide-scale implementation. Peptides have varying masses, charges, cleavage sites, and ionization efficiencies; thus some peptides are not detectable by MS. Another challenge that exists in implementing MS-based methods is the low abundance of variant peptides compared to the high abundance non-informative backbone peptides. This issue escalates with increasing library size and with the use of multiple proteases. For the small variant library presented here, we have circumvented this issue by developing a targeted MS acquisition method that combines the multiplexed quantification of SPS-MS3 methods with the high selectivity and sensitivity of parallel reaction monitoring. Our method works for comparing a dozen variants, at a time, and could be used to assay more if the libraries were split into sub pools with a bridging channel. This approach would still offer considerable advantage over methods requiring protein purification. In order to scale these assays up to hundreds or thousands of variants, additional method tailoring would be required.

Genetically encoded peptide barcodes called Flycodes(Egloff et al., 2019; Matsuzaki et al., 2021) have been used to determine protein binding characteristics in hundreds nanobodies. Flycodes aere short peptide sequences with demonstrated detectability in the mass spectrometer that ideally do not influence the protein of interest. We envisage that Flycodes could be used to assay other properties besides binding, and could be applied to the thermal stability and solubility assays described in this paper. These peptide barcodes could also be measured using alternative protein sequencing methods(Chinnaraj et al., 2025; Motone et al., 2024). High throughput protein-based methods for variant scanning, like those presented in our study, will further our understanding of the molecular basis of disease

2.5 Conclusions

In this study, we demonstrate that pooled MS-based assays for determining protein variant solubility and stability provide comparable results to traditional biochemical methods without the effort required to purify variants. The exquisite sensitivity of MS enables the measurement of unstable and poorly expressed variants that cannot be measured by other methods. We demonstrate the utility of protein-centric approaches to study variant effects and curtail many bottlenecks in traditional biochemical variant analyses and deep mutational scanning. Although additional improvements will be required to implement these methods for large libraries of protein variants, our work provides a proof of concept for using MS for variant scanning at the protein level. By integrating data from assays that report on different protein properties, a more detailed view of the molecular mechanisms underlying the dysfunction of disease-associated variants can be obtained. In this regard, high throughput protein assays coupled with MS detection are well-positioned as the next frontier in mutational scanning.

2.6 Acknowledgements

We thank members of the Villén lab and Schweppe lab for enriching scientific discussions and feedback on this project. We thank the lab of Prof. Lesa Beamer at the University of Missouri for providing plasmids containing the PGM1 variants. Research reported in this publication was primarily supported by the National Human Research Institute of the National Institutes of Health under award number RM1HG010461, by the National Institute of General Medicine of the National Institutes of Health under award number R35GM152061, and by a Medical Research Program grant from the W.M. Keck Foundation. The work and personnel involved in the project were additionally supported by NIH grants R35GM119536 from the National Institute of General Medicine and R01AG056359 from the National Institute of Aging; and by Human Frontiers Science Program grant RGP0034/2018. MDB was supported by a Canadian Institutes of Health Research Postdoctoral Fellowship (193932). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or other funding agencies.

2.7 Author contributions

This project was conceived by RR and JV. Experiments were designed by RR and KH. Experiments, sample preparation, and mass spectrometry analysis were performed by SM with assistance from KH, MDB, and RR. The targeted-SPS-MS3 acquisition method was developed by RR and implemented by SM and RR. Data analysis was conducted by SM and MDB, with support from RR. Figures were made by SM with feedback from all authors. SM wrote the initial

draft of the manuscript and all authors edited it. JV provided overall project supervision, infrastructure, and funding.

2.8 Methods

2.8.1 Library plasmid design

The plasmid library was ordered from Twist Biosciences in a pET-28a(+) (Novagen) plasmid backbone with N-terminal His-tag and a thrombin cleavage site. Codon-optimized PGM1 variant sequences were cloned between NdeI and XhoI sites under the control of the T7-lacO promoter.

2.8.2 Expression of PGM1 mutant library in *E. coli*

Expression was carried out as previously described (Lee et al., 2014). The pooled plasmid library was transformed into T7 Express lysY competent *E. coli* (NEB C3010I) using the manufacturer's protocol. Following overnight growth at 37°C, cells were grown to OD_{600} of 0.6 - 0.8. The culture was then cooled to 19°C and PGM1 expression was induced overnight at 220 rpm agitation with 0.2 mM IPTG (isopropyl β- d-1-thiogalactopyranoside). Cells were split into 25 mL (5.6 mg dry) or 50 mL (11.2mg dry) aliquots and centrifuged for 10 mins at 7200 x g, rinsed with ice-cold PBS, and pellets were snap-frozen in liquid nitrogen and stored at -80°C.

2.8.3 Solubility assay

One *E. coli* pellet expressing PGM1 library (5.6 mg) was thawed on ice and resuspended in 100 µL of either native buffer, denaturing buffer, or BPER buffer (ThermoFisher Scientific). Buffer compositions were as follows: Native buffer (1x phosphate-buffered saline, 0.25x cOmplete™

protease inhibitor cocktail (Roche), 2 mM $MgCl_2$, 50 units/mL benzonase, 0.02 µg/µL lysozyme); denaturing buffer (100 mM HEPES, 100 mM NaCl, 8 M urea, 2 mM $MgCl_2$, 0.25x cOmplete™ protease inhibitor cocktail (Roche), 50 units/mL benzonase, 0.02 µg/µL lysozyme); BPER buffer (ThermoFisher Scientific), 0.25x cOmplete™ protease inhibitor cocktail (Roche), 2 mM $MgCl_2$, 50 units/mL benzonase, 0.02 µg/µL lysozyme). Resuspended pellets were split into three replicates and buffer was added to each for a total volume of 500 µL per replicate. For the lysis with BPER, the cell suspension was pipetted until homogenous and incubated for 10 min at room temperature, lysates were clarified by centrifugation at 15,000 x g for 5 min at 4°C, and 450 µL of supernatant was recovered. For the native and denaturing lysis, samples were thoroughly vortexed and nutated at room temperature for 20 min, before cells were lysed on wet ice using sonication at 30% power for eight rounds of 30 seconds, with 30 seconds of rest in between. Lysates were clarified by centrifugation at 15,000 x g for 5 min at 4°C and 450 µL of supernatant was recovered. In order to have uniform buffer composition for digestion, each lysate was diluted with equal parts of the other two lysis buffers (without benzonase and lysozyme) and quantified using the Pierce™ BCA protein assay (ThermoFisher Scientific). Equal volumes of protein were digested for each replicate and buffer condition.

2.8.4 Thermal stability assay

E. coli expressing PGM1 library pellet (11.2mg) was thawed on ice and resuspended in 1 mL of native lysis buffer (1x phosphate-buffered saline, 0.25x cOmplete™protease inhibitor cocktail (Roche), 2 mM $MgCl_2$, 0.02 µg/µL lysozyme). Lysis was carried out in bulk, samples were thoroughly vortexed and nutated at room temperature for 20 min. Cells were lysed on wet ice using sonication at 20% power for ten rounds of 30 seconds, with 30 seconds rest in between. Lysates were clarified by centrifugation at 15,000 x g for 10 min at 4°C. Samples were split into

three replicates and kept on ice while protein quantification was performed using Pierce™ BCA protein assay Kit (ThermoFisher Scientific). The lysate was diluted in native lysis buffer to 2.5 µg/µL. TPP sample preparation was similar to the previously described method (Savitski et al., 2014). For each of the three technical replicates, 100 µL of 2.5 µg/µL lysate was added to ten PCR tubes on ice. Samples were pre-equilibrated at 37°C for 5 min. Lysates were treated for 5 min at the following temperatures: 37.0°C, 37.7°C, 39.5°C, 42.4°C, 46.3°C, 50.1°C, 53.8°C, 57.6°C, 61.5°C, and 64.4°C. Following temperature treatment, samples were incubated for 5 min at 25°C. Soluble protein extraction was performed by centrifuging PCR tubes at 21,000 x g for 30 min at 4°C. The soluble portion was recovered from each PCR tube and diluted with an equal volume of urea buffer (100 mM HEPES, 100 mM NaCl, 8 M urea, 2 mM $MgCl_2$,). An extra control channel, which was subjected to the lowest temperature treatment (37°C) was used for protein quantification using the Pierce™ BCA protein assay kit.

2.8.5 Proteomic sample preparation

2.8.5.1 Reduction and Alkylation

Following solubility and thermal stability assays, all lysate samples were reduced with dithiothreitol (DTT; 5 mM final concentration) at 55°C for 30 min with agitation at 300 rpm and subsequently alkylated with iodoacetamide (final concentration 15 mM) at RT in the dark for 30 min without agitation. Alkylation was quenched with an additional 5 mM DTT and incubated for 30 min at room temperature with agitation at 300 rpm.

2.8.5.2 Desalting and Digestion

Reduced and alkylated lysates were split into three before digestion and cleanup. Solubility samples were desalted by the manual SP3 method(Hughes et al., 2019). Thermal stability samples were cleaned up and digested with an automated implementation of the SP3 protocol(Leutert et al., 2019). Samples were split in thirds for three separate digestions by different enzymes. Digestion occurred in 50 mM HEPES at a ratio of 1:25 enzyme to protein (by mass) using either trypsin (Promega), LysC (Wako Chemicals) or chymotrypsin (Promega) for 4 hours at 37°C. Digestions were quenched by acidification, and peptides were dried and stored at -20°C.

2.8.5.3 TMT labelling and MCX cleanup

For each condition and replicate, 10 µg of dried peptides were resuspended in 30% ACN and labelled with 80 µg of TMT11plex (ThermoFisher), as per manufacturer's instructions. The TMT labelling scheme for both assays can be found in Supplementary Table Ai. Labelling was quenched with hydroxylamine to a final concentration 1.25%, dried and frozen. Samples were resuspended in 0.1% TFA, and cleaned up using mixed-mode cation exchange material on stage tips(Ishihama et al., 2006). Briefly, 20 µL pipette tips were packed with eight cookies of Empore™ polystyrene-divinylbenzene, reversed-phase sulfonate solid phase extraction discs (3M). Tips were conditioned with 30 µL of each: 100% ethanol; 100% acetonitrile (ACN); 75% ACN, 5% NH4OH; 75% ACN, 1% AcOH, 0.1% trifluoroacetic acid (TFA). After adding the solution, tips were centrifuged at 420 x g until only a small amount of liquid remained above the cookies. Dried TMT-labelled peptides (10 µg) were resuspended in 100 µL 4% formic acid, 3% ACN and added to the tips. Tips were washed twice with 120 µL of 0.1% TFA, followed by two

washes with 60 µL of 75% ACN-AcOH. Peptides were eluted with 60 µL 75% ACN, 5% NH_4OH and dried. Labelling efficiency was determined to be > 98%.

2.8.6 LC-MS/MS analysis of peptides

Dried peptides were resuspended in 4% formic acid (FA), and 3% ACN to a concentration of 0.5 µg/µL. Samples were measured on an Orbitrap Eclipse (ThermoFisher Scientific) tribrid mass spectrometer with an in-line EASY-nLC™ 1200 System (ThermoFisher Scientific) fitted with a 100 µm ID × 3 cm precolumn packed with Reprosil C18 3 µm beads (Dr. Maisch GmbH) and a 100 µm ID × 30 cm analytical column packed with Reprosil C18 1.9 µm beads (Dr. Maisch GmbH) housed into a column heater set at 50°C. Peptides were separated by reverse-phase chromatography at 400 nL/min on a 60 min gradient with the following solvents: A (0.1% FA), B (0.1% FA, 80% ACN): 0-2 min 3-13% B; 2-34 min 13-30% B; 34-42 min 30-50% B; 42-44 min 50-90% B; 44-49 min 90% B; 49-50 min 90-3% B; 50-60 min 3%B.

2.8.7 SPS-MS3 acquisition

Standard SPS-MS3 acquisition was adapted from(McAlister et al., 2014) and employed a high-resolution MS1 scan in the orbitrap, followed by MS2 in the ion trap and high-resolution MS3 scans of selected precursors in the orbitrap. Acquisition starts with a full MS1 scan on the Orbitrap (scan range 400–1600 m/z; resolution 120,000; 100% Standardized AGC target 4×10^5; and maximum injection time 50 ms). We used dynamic exclusion of 30 sec and cycle time of 3 sec. The MS2 consisted of collision-induced dissociation (CID); fragmentation-based quadrupole ion trap analysis; scan speed Rapid; 100% Normalized AGC target 1×10^4; isolation window 0.7 m/z; NCE 36% with fixed mode; and maximum injection time of 50 ms. Synchronous precursor isolation of the most intense MS2 ions was carried out and an MS3 scan with high collision

dissociation (HCD) fragmentation was analyzed in the Orbitrap (resolution 30,000; 2 m/z isolation window, mass range 100–500 m/z; 100% Normalized AGC target 5×10^4; NCE 55%; and auto injection time setting; and maximum number of SPS ions 10).

2.8.8 Targeted-SPS-MS3 acquisition

Our Targeted-SPS-MS^3 acquisition method comprises MS1, MS2 and MS3 scans all carried out in the Orbitrap mass analyzer and leverages targeted mass inclusion lists for the MS2 and MS3 scans, as well as a mass trigger list for the MS3 scans to select PGM1 variant peptide precursors. The method starts with a full MS1 scan in the Orbitrap (400-1800 m/z; resolution 120,000; normalized AGC target 100%, and auto setting for maximum injection time). Using a target mass filter in unscheduled mode and 25 ppm mass tolerance, precursors matching our inclusion list (see Supplemental Table Aii) were selected and up to 4 precursors per cycle were subjected to subsequent MS2 and MS3 data-dependent scans. The MS2 analysis of precursors consisted of collision-induced dissociation (CID) (isolation window 0.8 m/z, fixed collision NCE 35%) acquired in the Orbitrap (resolution 30,000, 100% normalized ACG target, and automatic injection time setting). A mass trigger filter was applied (Supplementary Table Aii) such that only precursors that produced diagnostic MS2 fragment ions and a 25 ppm tolerance were selected for a subsequent MS3 scan. For these precursors, MS3 scan proceeded by synchronous precursor selection of a maximum of 5 fragment ions listed in MS2 targeted mass filter (Supplementary Table Aii) (0.8 m/z MS isolation window), HCD fragmentation (NCE 55%), and analysis of reporter ions in the Orbitrap (50,000 resolution, scan range 100-500 m/z, normalized ACG target 200%, maximum injection time 500 ms).

2.8.9 Mass spectrometry data analysis

Each raw file was searched three times for each of the three proteases: trypsin, LysC and chymotrypsin. All MS data were searched using Comet version 2019.01.2(Eng et al., 2013, 2015) against the *E. coli* K12 protein sequence database with PGM1 added (wild type sequence and thirteen additional entries containing each single amino acid variant). Comet parameters for all searches included 50 ppm precursor mass tolerance, 2 allowed missed cleavages, fragment bin tolerance 0.02 Da, digest mass range 600-5000 m/z, maximum precursor charge 6, maximum fragment charge 5, variable methionine oxidation modification, constant modifications of cysteine carbamidomethyl, TMT11 n-terminal, and TMT11-lysine. Reporter ion intensities were processed using IsobaricQuant(Hogrebe et al., 2022) with the following parameters: most_intense ion for search method, search.ppm 20ppm, search.ppm.1 20ppm, search.ppm.2 20ppm, precSignalWindow 1, ms2 precursor window 2.5, ms1 depth search 10, topX 10.

2.8.10 Bioinformatics

2.8.10.1 General

Bioinformatic analyses were performed using R (version 4.1.1 2021-08-10). Hinges on boxplots correspond to the first and third quartiles, whiskers correspond to 1.5 IQR.

2.8.10.2 Solubility Analysis

For solubility samples, data were filtered (precursor mass error $1 < PPM < 11$ based on instrument performance, Comet (XCorr >= 1, msnTotalSignal(MS3)>= 10,000, totalSignalSPSWind >= 100,000, expect < 0.01). Normalization was performed to correct for any variability in starting sample amounts for replicates in the same solubility buffer. Normalization

factors were calculated from runs with the SPS-MS3 method, by dividing the total TMT reporter intensity measured at the MS3 for all peptide spectral matches assigned to *E. coli* proteins for each replicate by the mean of the total reporter intensity for all replicates of the same buffer. For each variant, solubility ratios were calculated using the top 3 most intense MS3 scans. Solubility ratios for the given variant peptide were calculated as the reporter ion intensity in a given replicate divided by the mean reporter ion intensity in the three denaturing samples.

2.8.10.3 Thermal Stability Analysis

This data consisted of three separate MS runs, each run was a replicate of a TPP melting experiment. For TPP samples, data were filtered as follows: precursor mass error based on instrument performance (1 < PPM < 11); IsobaricQuant(totalSignalSPSWind >= 5,000, TMT126 channel (37.0°C) MS3 intensity > 0). Peptide intensity was not normalized between replicates. For each variant, the top 3 most intense MS3 scans were used for melting temperature determination in each replicate. Unlike conventional TPP data processing, the PSM intensities for a given replicate were not summed prior to curve fitting - each curve represents a single MS3 scan. Melting curves and Tms were calculated using the Bioconductor TPP package (Childs et al., 2020). The analyzeTPPTP function was used to generate Tms (method = meltcurvefit, normalize=FALSE, startPars (Pl = 0, a=550, b=10), splineDF(3:7)). Melting curves were filtered to have r_sq > 0.75.

2.8.10.4 Prediction of protein stability

Values for $\mathbf{\Delta\Delta G}_{Mut \rightarrow WT}$ also called DDG were calculated using the mutfunct tool(Wagih et al., 2018) on PGM1-1 structure (UniProt ID: P36871).

CHAPTER 3. PHOSPHORYLATION IN STRESS RESPONSE

3.1 Abstract

To understand stress-type specific phosphosignalling, we measured the proteome and phosphoproteome of SGs under diverse stress conditions in human osteosarcoma (U-2 OS) engineered cell line with a GFP-fused stress granule marker, Ras GTPase-activating protein-binding protein 1 (G3BP1). Using mass spectrometry, we quantified stress-specific changes to the total proteome and phosphoproteome in response to oxidative stress (sodium arsenite), osmotic stress (sorbitol), and heat stress in six biological replicates per condition. Comparing regulated phosphosites reported in the literature with our data confirmed several expected phosphorylation events. Moreover, we identified 158 candidate phosphosites on proteins that potentially localize to SGs and may be involved in SG regulation. These data serve as a foundation for further investigation and understanding of SG biology.

3.2 Introduction

Phosphorylation is a post-translational modification critical for normal cellular processes, the propagation of signals through cells, and responses to the environment. Phosphorylation is a reversible process and serves as a dynamic way for proteins to rapidly respond to stimuli. While phosphorylation is critical for normal cellular processing, aberrant phosphorylation can contribute to disease. A deep understanding of signalling and phosphorylation in health and disease creates avenues for developing new therapeutics. Unfortunately, despite the importance of phosphorylation, less than 0.1% of all reported phosphosites have annotated functions(Needham et al., 2019).

Responding to stress is a normal part of cellular functioning. Prolonged stress exposure, however, can lead to cellular death and contribute to disease. We decided to investigate the proteome and phosphoproteome response to three types of stresses: osmotic stress, heat stress, and oxidative stress. These types of stressors are physiologically relevant cellular states.

Hyperosmotic stress (henceforth called osmotic stress) occurs when an osmolyte or solute concentration is higher extracellularly than intracellularly. If the solute is membrane impermeable, the cell shrinks as water fluxes out to increase intracellular solute concentration(Reinehr & Häussinger, 2006). Shrinkage causes mechanical stress on the cytoskeleton of the cell, and the cell upregulates genes involved in the synthesis and transport of osmolytes. Cells also respond by upregulating growth arrest and DNA damage (GADD) genes, p53, and pro-inflammatory cytokines(Brocker et al., 2012). Transcription, translation and the cell cycle are also arrested. Renal and gastrointestinal cells have adapted to dramatic fluctuations in osmotic pressures. However, many other types of cells are susceptible to osmotic stress during inflammation and in numerous disease conditions (encephalopathy, diabetes, hypernatremia etc.(Brusilow & Cooper, 2011; Neuhofer, 2010).

Oxidative stress occurs when reactive oxygen species (ROS) accumulate inside the cell. While cells have mechanisms to tolerate low levels of ROS generated as byproducts of enzymatic processes, high levels of ROS can damage nucleic acids, amino acids, and cellular structures(Pizzino et al., 2017) and cause disease. Unchecked oxidative stress is causative of cancers and numerous cardiovascular, neurological, respiratory, and renal diseases.

Under stress conditions, cytosolic proteins and RNA condense to form membrane-less organelles called stress granules. SG formation is a reversible process whereby proteins and mRNA dissolve upon removal of the stressor (Figure 1A, schematic of SG formation). The

formation of stress granules benefits cells as they protect RNA and proteins from degradation, alter RNA processing, and disrupt cellular signalling. Dysfunction of stress granule formation/ dissolution can contribute to the pathology of several neurodegenerative diseases and cancer and can even be exploited by viruses for infection.

The identity of protein and RNA components within SGs is stress–type specific(Advani & Ivanov, 2020; Hofmann et al., 2021; Kedersha et al., 2005), leading to different SG subtypes. While mechanisms of initiation and disassembly/ resolution for different subtypes and some core shared proteins are known, the identity of protein components between stress granule subtypes is largely unexplored(Advani & Ivanov, 2020). Even if SGs contain similar proteins, they may differ in proteoforms, that is, the combination of post-translational modifications on the protein. PTMs such as acetylation, methylation, ADP ribosylation, ubiquitination and phosphorylation have been identified in SGs(Hofweber & Dormann, 2019). Ubiquitination has been the most heavily studied of PTMs in SG. U2 OS cells were shown to have stress-specific ubiquitination signatures, and ubiquitination is critical for priming cells to recover from heat stress(Gwon et al., 2021; Maxwell et al., 2021; Pan et al., 2024). Phosphorylation, while being the most common cellular PTM, has been relatively understudied in SGs.

Given that several kinases and phosphorylation sites are implicated in SG biogenesis, that SGs are hypothesized to disrupt canonical signalling pathways(Kedersha et al., 2013), and the timescale of phosphorylation aligns with that of SG formation, we predict that phosphorylation is much more prevalent in SGs than previously reported. Identifying proteins (and phosphorylation) in SGs has historically been challenging for several reasons. First, SGs are highly dynamic and stress-type specific, so it is critical that no method introduces additional cellular stress and confounds results. Reagents commonly used for cross-linking (formaldehyde) and ascorbate

peroxidase proximity labelling (hydrogen peroxide) can induce a stress response. Secondly, SGs are hard to isolate because they are membraneless. The cellular context and concentration are critical to SG composition. Therefore, one cannot use lysate-based fractionation or affinity purification methods. Despite these potentially confounding stress effects, labs have used cross-linking and oxidation-dependent proximity labelling approaches (APEX). To circumvent these issues, other labs have employed proximity BioID, TurboID, miniTurboID, and uMAP(Dionne & Gingras, 2022; Pan et al., 2024; Schreiber et al., 2024) to identify SG components. Lastly, phosphorylated peptides have relatively low abundance compared to the rest of the proteome and are hard to detect by MS without enrichment. Previous examination of SG PTMs has been site-specific, and thus required prior knowledge of sites to be measured. Fluorescence microscopy has been the primary interrogation method, but it is limited by the availability of PTM-specific antibodies and multiplexing capacity.

In this study, we present data on the stress response in cells under oxidative (sodium arsenite), heat (42℃), and hyperosmotic (sorbitol) stress conditions. First, we present global differences in phosphorylation under different stress conditions to identify shared and stress-specific responses. Second, we use publicly available SG databases and datasets to identify phosphorylation sites on proteins that localize to stress granules. Third, using computational prediction scores, we predict whether these candidate SG phosphosites would alter phase separation behaviour or whether they may just be signalling proteins sequestered into SGs without a biophysical role in SG formation/ dissolution. These data provide a resource for phosphorylation sites that could potentially localize to SGs and can be used for follow-up studies using proximity-labelling techniques.

3.3 Results & Discussion

3.3.1 Experimental overview

We first examined the difference between stress response types in the global phosphoproteome (Figure 3-1a) in human osteosarcoma (U-2 OS) cells expressing a core stress granule marker protein, G3BP1, fused to GFP (Gwon et al., 2021). Cells were exposed to either oxidative stress (0.5 mM sodium arsenite), heat stress (42^0C), hyperosmotic stress (0.4 M sorbitol) or control conditions for one hour. Live cell imaging confirmed the formation of SGs in stress conditions. Samples were harvested, digested and split into samples bound for whole proteome or phosphopeptide sample preparation before acquisition by mass spectrometry.

From these data, we quantified over seven thousand phosphosites in each treatment condition (Figure 3-1b). The vast majority of identified phosphosites were localized to serine (>75%), a small portion on threonine (~17%), with very few sites on tyrosine (< 8%) (Figure 3-1c). This distribution is consistent with what has been previously reported for ratios of phosphoserine, phosphothreonine and phosphotyrosine using metal affinity phosphopeptide enrichment(Khoury et al., 2011).

We were interested in identifying phosphorylation events involved in response to stress. The regulation was calculated using the phosphosite abundance fold change normalized to the corresponding protein-level fold change. Here, we report more than 300 significantly regulated phosphosites per stress type (Figure 3-1d). Strikingly, arsenite stress mostly up-regulates phosphosites, whereas heat and sorbitol stress up and down-regulate phosphosites. Mechanisms to explain these differences in directionality response remain unclear.

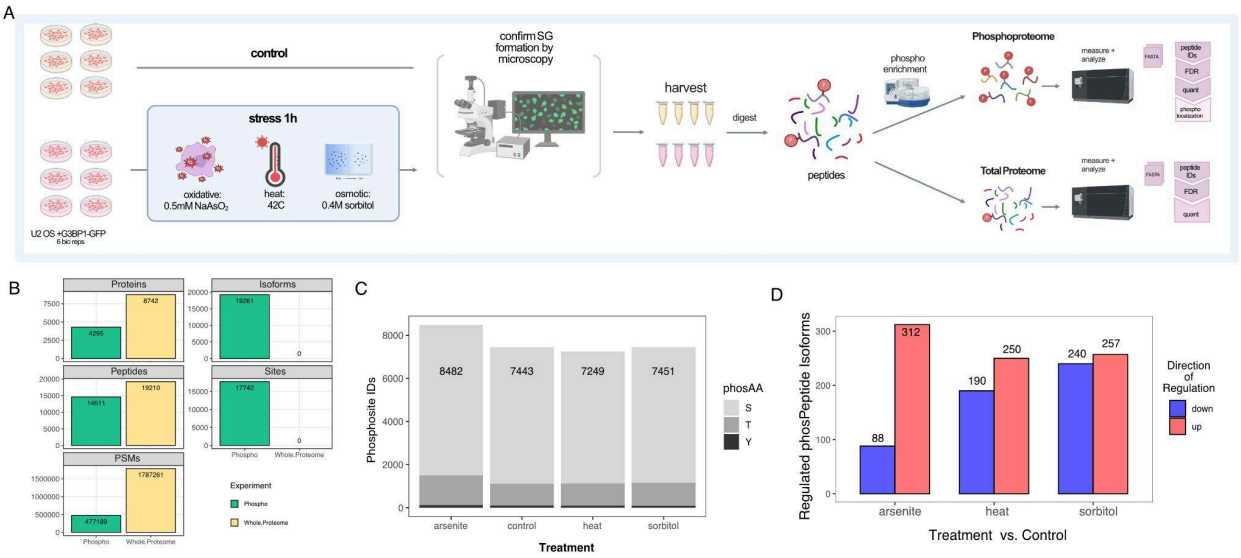


FIGURE 3-1. Proteomics and phosphoproteomics of U-2 OS cells under various stress conditions. ***(A)*** Experimental workflow for investigating the phosphorylation response in U-2 OS cells to osmotic (200mM sorbitol), oxidative (0.5mM sodium arsenite), and heat stress (42℃). ***(B)*** Identifications from global proteome and phosphoproteome experiments. ***(C)*** Distribution of phosphorylated residues identified in each condition. ***(D)*** Significantly regulated phosphoPeptide isoforms identified in each stress condition relative to control.

3.3.2 Shared and stress-type specific phosphorylation responses

To understand differences in stress-type phosphorylation responses, we performed hierarchical clustering of all observed phosphosites (Figure 3-2a). Three clusters (6, 13, and 17) are highlighted for further analysis to determine phospho-regulation distinct to specific stresses and regulation that are shared in all stress conditions. Distributions of z-scores across sample types were visualized in line plots and coloured by distance from the cluster centre (Figure 3-2b). This patterning shows that phosphosites in cluster 6 are part of a shared stress response, indicated by an increase in abundance in all stresses compared to the control.

To dig deeper into phosphorylation differences between shared and stress-type-specific responses, we performed Fisher exact tests on the enrichment of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and gene ontology (GO) term enrichment (cellular compartment, biological processes and molecular functions) (Figure 3-2c).

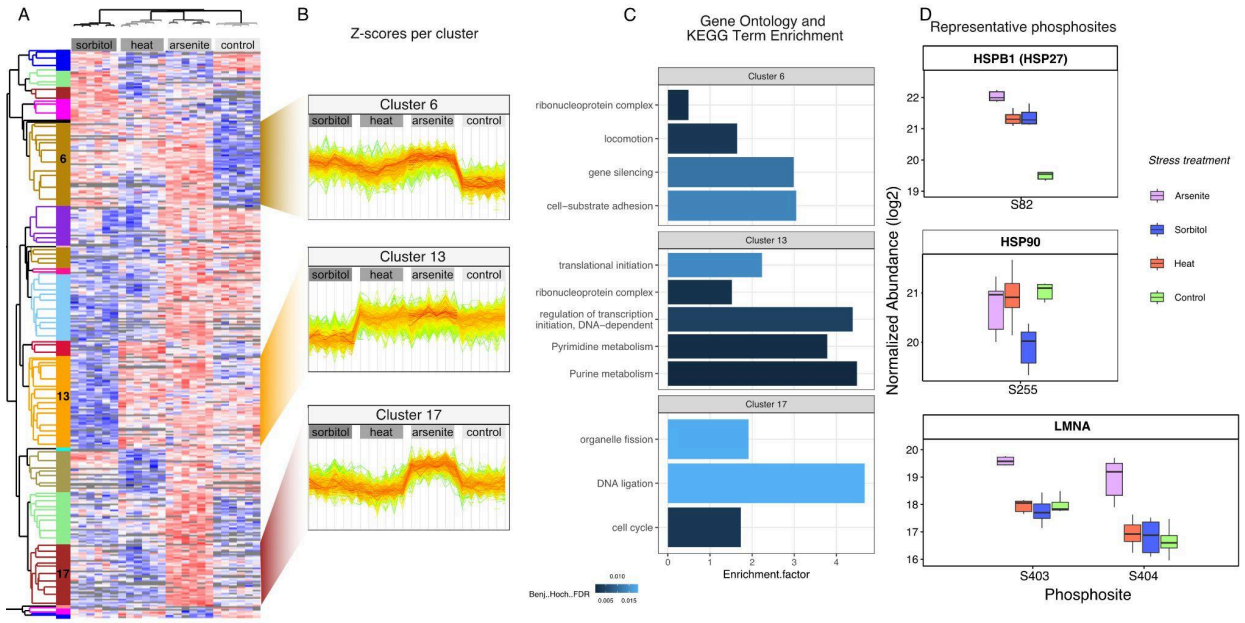


Figure 3-2. Overview of phosphorylation response to stress ***(A)*** Heatmap of phosphosites observed in all conditions. Phosphosites were row-normalized and z-score is displayed. Clustering was performed by *euclidean distance, with defined 21 clusters. ***(B)*** Line plot of relative z-scores for designated clusters, lines coloured by distance from centre value. ***(C)*** Enrichment of Gene Ontology annotations in designated clusters. Significance cutoff B-H FDR < 0.05 ***(D)*** Boxplots displaying log2 phosphosite intensities observed in different stress conditions. Protein names are designated in the grey bolded header.

Cluster 6, which is composed of phosphosites that increase in abundance to all stresses, is enriched for phosphosites on proteins involved in mRNA processing and metabolic processes, locomotion, cell junctions and adhesions. Stress-specific phosphorylation responses were observed in clusters 13 and 17. Phosphosites in cluster 13 decrease in abundance in response to sorbitol stress only. Enrichment in KEGG pathways involved in purine and pyrimidine metabolism are consistent with reported regulation of osmolyte synthesis in osmotic stress(Brocker et al., 2012). Enrichment of sites on proteins involved in transcription and translation is also expected given that osmotic stress halts both processes.

Phosphosites in cluster 17 increase in abundance in response to arsenite stress only and are enriched for phosphosites on proteins involved in organellar fission. This is consistent with reported reduction of mitochondrial biogenesis (fission and fusion) in response to oxidative

stress(Iqbal & Hood, 2014). This cluster is also enriched for phosphosites on proteins related to DNA ligation and cell cycle, both anticipated outcomes of DNA damage from oxidative stress(Mladenov et al., 2016; Young & Woodside, 2001).

In addition to recognizing gene ontologies shared between stress responses, we were interested in identifying regulation at the phosphosite-level, both novel and expected sites. First, we examined phosphosites that were previously annotated in stress response pathways. Examples of significantly regulated phosphosites representative of each cluster are shown in Figure 3-2d. An example of a phosphosite with a shared stress response from cluster 6 is phosphorylation at Ser-82 (pSer-82) in heat shock protein 27 (HSP27, also known as heat shock protein beta-1 HSPB1). HSP27 is a molecular chaperone involved in stress response(Almeida-Souza et al., 2010), cytoskeletal dynamics, and differentiation. In response to stress, phosphorylation of HSP27 at Ser82 by MAPKAPK2/3/5 occurs, resulting in reduced oligomerization size of HSP27(Gurgis et al., 2014), and decreased chaperone function. We observe a significant, regulated increase in pSer-82 on HSP27 in all stress conditions. This is consistent with previously reported phosphorylation analysis showing an increase in pSer82 in response to sorbitol (Boyd et al., 2023), heat (Landry et al., 1992; Reinhardt et al., 2007), and arsenite(Landry et al., 1992; Vertii et al., 2006). Together these data show that the general stress induces phosphorylation response in our experiment aligns with previous reports. In cluster 13 (Figure 3-2d), there is a decrease in phosphorylation in response to sorbitol stress. An example of this is at position pSer-255 in heat shock protein HSP 90-beta (HSP90 also known as HSBP90AB1). In response to sorbitol stress, we observed a significant negative regulation of pSer-255 in HSP90. Weidenauer and Quadroni (2021) showed that the pSer-255 site alters interaction partners, secretion and conformation(Weidenauer & Quadroni, 2021). Their data also

show that the pSer-255 site is not regulated in response to heat or oxidative stress conditions. In cluster 17, we observe an expected and arsenite-specific upregulation of phosphosites pSer-403 and pSer-404 in lamin A (LMNA). Lamin A proteins provide structural support within the nucleus by forming intermediate filaments(Shimi & Goldman, 2014). This is a regulatory site that Akt1 phosphorylates in mature lamin A or by PKC in prelamin. Upregulation seen in pSer-404 is known to be a regulatory site linked to oxidative stress and impacts protein degradation and inhibit transcription(Shimi & Goldman, 2014). We demonstrated that we see expected results for key regulatory sites involved in stress-specific responses.

3.3.2 Phosphorylation sites on proteins known to localize to stress granules

Many kinases have been shown to be involved in SG dynamics(Krisenko et al., 2015; Li et al., 2020; Wippich et al., 2013), further suggesting the importance of phosphorylation in SGs. Examples include is the phosphorylation of elongation initiation factor 4 alpha (eIF4**α**), which drives the formation of non-canonical pre-initiation complexes leading to the formation of SGs(Panas et al., 2016) and dual specificity tyrosine-phosphorylation-regulated kinase 3 (DYRK3) which is necessary for stress granule disassembly(Wippich et al., 2013). However, to our knowledge, no one has done a large-scale analysis of phosphorylation in SGs. Towards this end, we endeavoured to build a list of phosphosites on proteins that localize to SGs to serve as a foundation for further experiments using proximity labelling or localization techniques. .

SG formation was confirmed by fluorescence microscopy **(**Figure 3-3a), by the appearance of G3BP1-GFP condensation consistent with SG formation. While SG formation via p-eIF2alpha - dependent mechanisms occurs for both arsenite and heat shock stress, we observe visually distinct SG phenotypes (size and number of granules), which, at minimum, suggest differences in coarsening rates between stresses or potentially protein/proteoform-level

differences. We observed no granules in the control conditions but did observe granules in the oxidative, osmotic and heat stress conditions. Interestingly, we saw a difference in size and number of granules between conditions. Where osmotic stress had the greatest number of granules, but they were all quite small. Temporal dynamics of stress granule formation could be explored using live cell imaging to quantify SGs over time.

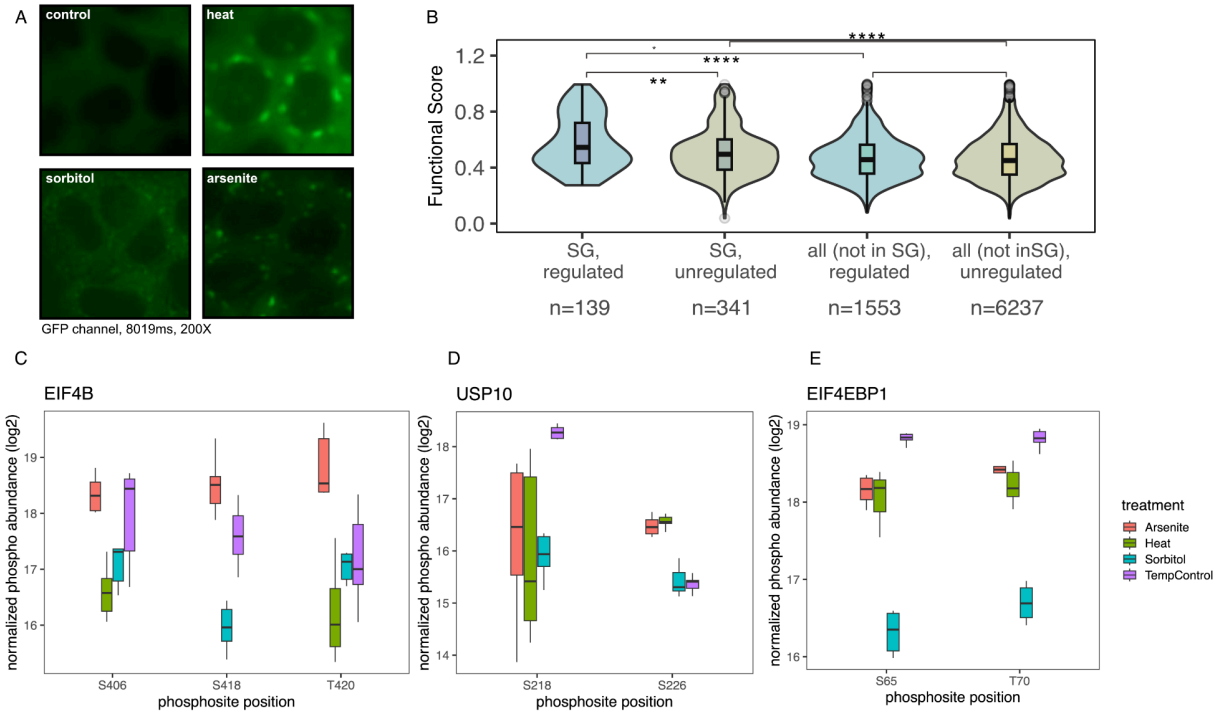


Figure 3-3. Phosphorylation in proteins known to localize to stress granules ***(A)*** schematic showing how phosphorylation may disrupt or promote interactions required for liquid-liquid phase separation. (***B)*** Imaging of U-2 OS cells expressing G3BP1-GFP. Cells were stressed for one hour, and GFP channel and brighfield images are overlaid. ***(C)*** Predicted functional score of phosphorylation sets subsetted by regulation and whether sites occur on proteins that localize to SGs. ***(D)*** Residue-level disorder score of phosphorylation sites subsetted by direction of regulation and protein annotations in the SG database. ***(D)*** Boxplots displaying log2 phosphosite intensities observed in different stress conditions.

Our data thus far has examined stress responses across the cell. To investigate phosphorylation sites that could be involved in stress granules, we filtered our data using a subset from the RNA granule database(Millar et al., 2023; Youn et al., 2019). We used a database of 284 proteins of tier 1 database hits that have been identified by colocalization studies in stress granules. In this study we identified 680 unique phosphosites on 120 proteins known to be in

stress granules. We utilized a phosphosite functional score prediction(Ochoa et al., 2020) that was developed using machine learning on over six thousand proteomics datasets. In running the score prediction against sites identified in our data, we found that phosphosites that are regulated and belong to proteins found in stress granules have higher predicted function than unregulated sites in SGs and within the whole cell (Fig 3-3b). These data suggest that many sites within stress granules warrant further investigation as they may contribute to SG biogenesis or disruptions of signalling cascades by sequestration of phosphoproteins in SGs.

We found several interesting examples of phosphorylation sites on proteins that localize to SGs (Supplemental Table B-i). An example in our list of potential SG phosphosites is eukaryotic translation initiation factor 4B (EIF4B) (Fig 3-3c). EIF4B is a component of the translation initiation complex and is necessary for mRNA to bind to ribosomes. Phosphorylation at Ser-406 and Thr-420 by RPS6KA1/2 increases affinity for EIF3 and contributes to the formation of the translation initiation complex(Ali et al., 2022). We see an increase in pSer-418 in arsenite conditions and a decrease in sorbitol. This sorbitol-specific signature is not surprising, given that others have shown that stress granule formation from osmotic stress proceeds through the EIF4(Hofmann et al., 2021).

Ubiquitin carboxyl-terminal hydrolase 10 (USP10) is an important deubiquitinase. In some stress conditions, it condenses with G3BP(Kedersha et al., 2016). We observe stress-specific phosphorylation on Ser-218 (Figure 3-3d) (a significant decrease in heat only) and on Ser-226 (a significant up-regulation in heat and arsenite conditions). These sites do not have an annotated function, but they could be responsible for stress type differences in SG biogenesis.

It must be noted that the database of SG proteins we used consists primarily of stress granules induced by oxidative stress. Given that the SG composition differs between stresses it is

possible that we are biased against proteins that uniquely localize to osmotic shock SGs or heat shock SGs. Another phosphoprotein of note is eukaryotic translation initiation factor 4E-binding protein 1 (EIF4EBP1). This protein is not listed in the SG database, but it is of particular interest due to its role in translation initiation. Phosphorylation of EIF4EBP1 reverses the repression of translation initiation(Patel et al., 2002). Interestingly, we see a downregulation of phosphorylation at Ser65 (by DYRK2, MAPK1, MAPK3 and MTOR) and Thr-70 (by MTOR), key sites that impact translation initiation, in all conditions. In our data (Fig 3-3e), sorbitol has the greatest downregulation at these sites, indicating that translation initiation may be more inhibited in sorbitol compared to other stresses. When not phosphorylated, it is ubiquitinated, leading to degradation. We note that we observe three significantly regulated phosphosites on proteins that localize to SGs which have not been reported previously (not in PhosphoSitePlus). They are as follows: (GENE_psite; ATXN2L_pS485, BAG3_pT307, NUP98_pS705).

3.4 Conclusion

In this study, we provided a list of candidate phosphosites that may localize to stress granules. There are caveats to these predicted sites, as they are biased towards proteins that localize to oxidative SGs. It is quite likely that we are missing many proteins that are only localised to SGs in heat or osmotic conditions. Furthermore, we identified sites on proteins that localize the SGs, but we cannot determine from this data whether the phosphorylated form actually partitions into granules. This is particularly important as phosphorylation is known to alter protein subcellular localization. The next step towards elucidating the SG phosphoproteome will be to perform proximity labelling experiment, with miniTurboID fused to G3BP1, a core SG marker protein. Prey protein pulldown can then be split for proteome and phosphoproteome enrichment. These data will result in a resource detailing the proteome and phosphoproteome in

SG composition in multiple stresses. From there we suggest performing phosphopeptide enrichment on proximity labelled SGs in various conditions to confidently localize sites. Kinase inhibitor assays and phosphomimetic could then be used to determine if the phosphosite is acting functionally in SG biogenesis.

3.5 Methods

3.5.1 Tissue culture, stress and harvesting

We used human osteosarcoma (U-2 OS) cells expressing a core stress granule marker protein, G3BP1, fused to GFP (Gwon et al., 2021). I processed 6 biological replicates per stress condition. Cells were cultured in 150mM plates at 37^0C, 5% CO2 in media containing DMEM (high glucose, with sodium pyruvate, Gibco) supplemented with 10% fetal bovine serum(Gibco) and 50 I.U/mL penicillin-streptomycin (Gibco). Two hours before stress, fresh media was added. Cells at ~80% confluence were exposed to either oxidative stress (0.5 mM sodium arsenite), heat stress (42^0C), or hyperosmotic stress (0.4M sorbitol), or control (37^0C) for one hour. Stressed cells were briefly washed with 10mL of Dulbecco's phosphate-buffered saline and then flash-frozen with liquid nitrogen to preserve signalling.

3.5.2 Imaging

SG formation was confirmed by live-cell imaging using fluorescence microscopy by the appearance of G3BP1-GFP condensation consistent with SG formation. We observed no granules in the control conditions but did observe granules in the oxidative, osmotic and heat stress conditions. Images were in the GFP channel at 200x magnification and exposed for 8019 ms. Images were normalized and processed using FIJI software(Schindelin et al., 2012).

3.5.3 Proteomic Sample preparation

3.5.3.1 Lysis

Frozen 150 mm petri plates containing cells were briefly thawed on ice for 2 min. 500 µL of lysis buffer (8M urea; 50 mM HEPES pH 8.0; 75 mM NaCl) was added to each plate prior to scraping and transferring to clean eppendorf tubes. plates and cells were collected by scraping on ice. Cells were lysed on wet ice using sonication at 30% power for at least five rounds of 30 seconds, with 30 seconds of rest in between. Lysates were clarified by centrifugation at 20,000 x g for 10 min at 15°C. Clarified lysates were diluted with sterile water up to 1mL final volume in order to reduce urea concentrations to be compatible with downstream methods. Lysates were quantified using the Pierce™ BCA protein assay (ThermoFisher Scientific). Equal volumes of protein were digested for each replicate and treatment condition.

3.5.3.2 Reduction, Alkylation and Digestion

Lysate samples were reduced with dithiothreitol (DTT; 5 mM final concentration) at 55°C for 30 min with agitation at 300 rpm and subsequently alkylated with iodoacetamide (final concentration 15 mM) at RT in the dark for 30 min without agitation. Alkylation was quenched with an additional 5 mM DTT and incubated for 30 min at room temperature with agitation at 300 rpm and then stored at -20°C. Lysates were thawed on ice and 200 µg of protein was diluted to achieve appropriate digestion buffer concentrations (final concentration 50 mM ammonium bicarbonate; 1.3 M urea; 8.38 µM HEPES, 12.56 µM NaCl, 5 µg/mL trypsin). Samples were digested with shaking at 300 rpm for 16 hours at 37°C. Digestions were quenched using trifluoroacetic acid (TFA) to reach an acidic pH less than 3. Peptides were centrifuged at 20,000

x g for 10 min at 4°C, pelleted material was discarded. Peptides were then dried and stored at -20°C.

3.5.3.3 Desalting

Digests were resuspended in 1% TFA and sonicated for three minutes. Samples were desalted using 60mg HLB Oasis plates (Waters) attached to a vacuum manifold. The plate was conditioned with 1 mL each of 100% methanol; 100% acetonitrile (ACN); 75% ACN, 0.5% acetic acid; 50% ACN, 0.5% acetic acid. The plate was rinsed three times with 1 mL of 0.1% TFA. Peptides were loaded onto the plate to be bound by the matrix. Sample wells were washed three times with 1mL of 0.1% TFA, and then eluted in 800uL of 75% ACN, 0.5% acetic acid. Peptides were then dried and stored at -20°C.

3.5.4 Whole proteome (WP) peptide processing

3.5.4.1 WP LC-MS

Dried peptides samples (from step 3.5.3.3) were resuspended in 4% formic acid (FA), and 3% ACN to a concentration of 0.1 µg/µL and sonicated for five minutes. Samples were measured on an Orbitrap Eclipse (ThermoFisher Scientific) tribrid mass spectrometer with an in-line EASY-nLC™ 1200 System (ThermoFisher Scientific) fitted with a 100 µm ID × 3 cm precolumn packed with Reprosil C18 3 µm beads (Dr. Maisch GmbH) and a 100 µm ID × 30 cm analytical column packed with Reprosil C18 1.9 µm beads (Dr. Maisch GmbH) housed into a column heater set at 50°C. Peptides were separated by reverse-phase chromatography at 400 nL/min on a 120 min gradient with the following solvents: A (0.1% FA), B (0.1% FA, 80% ACN): 0-93 min

7-38% B; 93-103 min 38-55% B; 103-104 min 55-95% B; 104-109 min 95% B; 109-110 min 95-3% B; 110-120 min 3% B.

3.5.4.2 WP MS Acquisition

MS/ MS acquisition for the whole proteome employed a high-resolution MS1 scan in the orbitrap, followed by high-resolution MS2 scans of selected precursors in the orbitrap. Acquisition starts with a full MS1 scan on the Orbitrap (scan range 375-1500 m/z; resolution 120,000; RF lens 35%, absolute AGC target 4.000e5). We used monoisotopic peak determination, a filter intensity threshold of 5.0e4, included charge states 2-6, dynamic exclusion of 30 sec and cycle time of 3 sec. The MS2 consisted of higher-energy collisional dissociation (HCD) fragmentation analyzed in the Orbitrap (resolution 30,000; 1.6 m/z isolation window, NCE 30%; standard ACG target, absolute ACG value 5.00e4, auto maximum injection time mode; and one microscan).

3.5.4.3 WP MS Analysis

Each raw file was searched with trypsin protease. All MS data were searched using Comet version 2019.01.2(Eng et al., 2013, 2015) against the human proteome (uniprot organism id 9606, retrieved 2023-09-22). Comet parameters for all searches included 20 ppm precursor mass tolerance, 2 allowed missed cleavages, fragment bin tolerance 0.02 Da, digest mass range 600-5000 m/z, maximum precursor charge 6, maximum fragment charge 5, variable modifications (methionine oxidation, N-terminal acetylation, STY phosphorylation), constant modification of cysteine carbamidomethyl. Extraction and quantification of peptide intensities was obtained using an in-house implementation of modest Feature Finder(Argentini et al., 2016) with the following parameters: peak search scan window 20, peak search RT window 30, peak

search ppm tolerance 10 ppm, peak search noise scan window 20, peak search noise MZ window 25, peak search max missed scans 2, peak search min scans 1.

3.5.5 Phosphoproteome peptide processing

3.5.5.1 Phosphoproteome enrichment

Dried peptides (250 µg from step 3.5.3.3) were resuspended in 900 µL 80% ACN, 0.1% TFA and sonicated for five minutes. Automated R2P2 phosphorylation enrichment was performed(Leutert et al., 2019) using the following parameters: 850 µL of peptide solution was mixed with 1X (70 µL = 5%) of beads; bound beads were washed three times with 80% ACN, 0.1% TFA and eluted in 50% ACN, 2.5% ammonium hydroxide and were immediately quenched with 60 µL of neutralization buffer (75% ACN, 10% formic acid).

3.5.5.2 Bead removal with C8 filters

Pipette tips (200 µL) were fitted with two cookies (layers) of C8 solid phase extraction discs (3M). Filters were preconditioned with 60 µL each of methanol; 100% ACN; 70% ACN, 0.25% acetic acid. Phosphopeptide samples were passed through the filters and collected in clean tubes. Filters were then chased with 50 µL of 70% ACN, 0.25% acetic acid and eluate was collected. Peptides were dried and stored at -20°C.

3.5.5.3 Phosphopeptide LC-MS

Dried peptides samples (from step 3.5.5.2) were resuspended in 4% formic acid (FA), and 3% ACN to a concentration of 0.1 µg/µL and sonicated for five minutes. Samples were measured on an Orbitrap Eclipse (ThermoFisher Scientific) tribrid mass spectrometer with an in-line EASY-nLC™ 1200 System (ThermoFisher Scientific) fitted with a 100 µm ID × 3 cm precolumn

packed with Reprosil C18 3 μm beads (Dr. Maisch GmbH) and a 100 μm ID × 30 cm analytical column packed with Reprosil C18 1.9 μm beads (Dr. Maisch GmbH) housed into a column heater set at 50°C. Peptides were separated by reverse-phase chromatography at 400 nL/min on a 90 min gradient with the following solvents: A (0.1% FA), B (0.1% FA, 80% ACN): 0-72 min 3-34% B; 72-73 min 34-60% B; 73-74 min 60-95% B; 74-79 min 95% B; 79-80 min 95-3% B; 80-90 min 3% B.

3.5.5.4 Phosphopeptide MS Acquisition

MS/ MS acquisition for the phosphoproteome samples employed a high-resolution MS1 scan in the orbitrap, followed by high-resolution MS2 scans of selected precursors in the orbitrap. Acquisition starts with a full MS1 scan on the Orbitrap (scan range 375-1500 m/z; resolution 120,000; RF lens 35%, absolute AGC target 4.000e5). We used peptide monoisotopic peak determination, a filter intensity threshold of 2.5e4, included charge states 2-6, dynamic exclusion of 30 sec with a mass tolerance ± 15ppm and cycle time of 3 sec. The MS2 consisted of higher-energy collisional dissociation (HCD) fragmentation analyzed in the Orbitrap (resolution 50,000; 1.6 m/z isolation window, NCE 30%; normalized ACG target 200%, absolute ACG value 1.00e5, maximum injection time 86ms; and one microscan).

3.5.5.5 Phosphopeptide MS Analysis

Each raw file was searched with trypsin protease. All MS data were searched using Comet version 2019.01.2(Eng et al., 2013, 2015) against the human proteome (uniprot organism id 9606, retrieved 2024-03-21). Comet parameters for all searches included 20 ppm precursor mass tolerance, 4 allowed missed cleavages, fragment bin tolerance 0.02 Da, digest mass range 600-5000 m/z, maximum precursor charge 6, maximum fragment charge 5, variable

modifications (methionine oxidation, N-terminal acetylation, STY phosphorylation), constant modification of cysteine carbamidomethyl. Results were filtered using Percolator(Käll et al., 2007) using a 1% FDR (PSM-level). Phosphosite localization was performed using an in-house implementation of AScore(Beausoleil et al., 2006) to filter phosphopeptides with AScore > 13. Extraction and quantification of peptide intensities was obtained using an in-house implementation of modest Feature Finder(Argentini et al., 2016) with the following parameters: peak search scan window 20, peak search RT window 30, peak search ppm tolerance 10 ppm, peak search noise scan window 20, peak search noise MZ window 25, peak search max missed scans 2, peak search min scans 1.

3.5.6 Bioinformatics

3.5.6.1 General

Bioinformatic analyses were performed using R (version 4.1.1 2021-08-10). Hinges on boxplots correspond to the first and third quartiles, whiskers correspond to 1.5 IQR.

3.5.6.2 Analysis

Phosphosite data was filtered for a localization score AScore > 13. For phosphosite data, the biological replicates within each treatment were median normalized using correction factors obtained from the summed intensities from the whole proteome data. Statistical hypothesis testing was performed using the Protti package(Quast et al., 2022). Significance of regulation was determined using a moderated t-test. Missingness was assigned as either complete, missing not at random (MNAR site present in < 25% of replicates), missing at random (MAR site present in at least half of the replicates, adjusted downward). Significantly regulated sites were imported

into Perseus(Tyanova et al., 2016) for further processing. Filters were applied such that site needed to be identified in at least three of six replicates. ANOVA using permutation based FDR of 0.05 with 250 randomizations was performed between the treatment types and data was filtered for FDR < 0.05. Z-scores were calculated for each p-site followed by hierarchical clustering of sites by Euclidean distance using average linkage. Fisher exact tests were used to determine significance of enrichment of annotated categories within clusters.

CHAPTER 4: CONCLUSION

4.1 Summary of thesis work

This thesis work has focused on investigating proteoforms originating at the levels of DNA and post-translational modification. I have shown that the properties of single amino acid variants can be measured in pooled format using mass spectrometry, and our results recapitulate those from biochemical assays on purified variants. I also characterized the phosphorylation response in three different stress conditions and used publicly available data to nominate candidate phosphosites that could potentially localize to stress granules. The work and methods presented in this thesis are a small part of a larger movement in the proteomics community migrating towards proteoform measurements.

4.2 Impact, scalability, and the future of profiling single amino acid variants by MS.

In Chapter 2, I presented an MS-based method for measuring single amino acid protein variants' biochemical properties (thermal stability and solubility). This proof-of-concept study illustrated the utility of applying MS-based assays to disease-associated protein variants to determine dimensions of PGM1 dysfunction. This protein-level property readout could not be achieved using traditional sequencing-based methods. I addressed challenges with variant detection using a novel MS acquisition method that combined the quantitative multiplexing capabilities of tandem mass tag synchronous precursor selection acquisition with the sensitivity of parallel reaction monitoring.

The first step in scaling up this method would be to increase the size of the variant library to hundreds or even thousands to be comparable to the throughput and depth of library coverage afforded by sequencing-based deep mutational scanning methods. Even in my modest library of

thirteen variants, I was unable to detect all tryptic variant peptides as some were of inappropriate length or had poor ionization for MS. Adding two proteases with different specificities (chymotrypsin and lysC) increased but did not fully cover the variant library. Peptide variant detection will also become increasingly complex with library size. For example, PGM1 yields 61 tryptic peptides, but only 1 out of 61 peptides contains the informative single amino acid variant. Therefore, my variant readout must be at the peptide level, as the 60 remaining peptides belong to the wild-type PGM1 "backbone" and will be shared across the rest of the library. The signal from highly abundant backbone peptides dominates the signal from the informative variant peptides, necessitating a targeted acquisition method we described. The relative ratio of backbone to variant peptides grows exponentially with library size. One clever way to overcome this is to use cleavable peptide barcodes, as had been proposed by a previous graduate student in the lab, Dr. Ian Smith(Smith, 2022). These would be C-terminal barcodes that do not affect solubility or thermal stability. Following assay selection, barcodes would be cleaved and enriched before measurement by MS, thus removing non-informative backbone peptides.

Beyond scaling the library size of PGM1 variants, we could envisage assaying other disease-relevant proteins. Our modular method of coupling functional protein assays with Targeted-SPS-MS3 acquisition could also be extended to additional assays of variant behaviour that reflect underlying disease biology. These could be properties such as metabolite binding, protein-protein interactions, or phosphorylation status. Complementary data from multiple MS-based functional assays would ultimately contribute to large multidimensional maps of protein variant function that serve to aid in annotating rare variants and deepen our understanding of the human proteome.

4.3 Stress-type specific phosphorylation: Impact and next steps.

In Chapter 3, I presented work examining phosphorylation response to short-term (1 hour) oxidative, osmotic and heat stress in human osteosarcoma cells. Clustering of phosphorylation responses co-regulated in different stresses revealed shared and distinct signatures of the three stresses. Phosphorylation signatures on known regulatory sites in our stress conditions agreed with published data. I was particularly interested in the role of phosphorylation in SGs. SG phosphosite localization is challenging as classical subcellular fractionation techniques such as sucrose density gradients or cross-linking are not amenable to phase-separated organelles. Because of challenges associated with the isolation of SGs, I decided to use the cell-wide phosphorylation dataset to try and predict which phosphosites could be localized to SGs.

I referenced published datasets of the SG proteome that were generated using computational predictions, physical association, genetic, and cell biological evidence(Youn et al., 2019). I created a list of 158 candidate phosphosites that are significantly regulated in stress and occur on proteins known to localize in SGs.

This candidate list provides an excellent foundation for understanding differences in stress-specific SG responses. However, there are limitations to this approach. First, we cannot confidently say that the phosphosites are localized to stress granules, as PTMs alter subcellular localization. Secondly, there are drawbacks to the published data we are referencing. Most of the studies are limited to oxidative stress, so it is possible that some proteins that are specific to heat SGs or osmotic SGs are being excluded, and we are missing potential phosphosites. Furthermore, the largest proximity dataset in the RNA granule database reports the SG proteome in the absence of stress, claiming that the condensates pre-exist in the cytosol without stress.

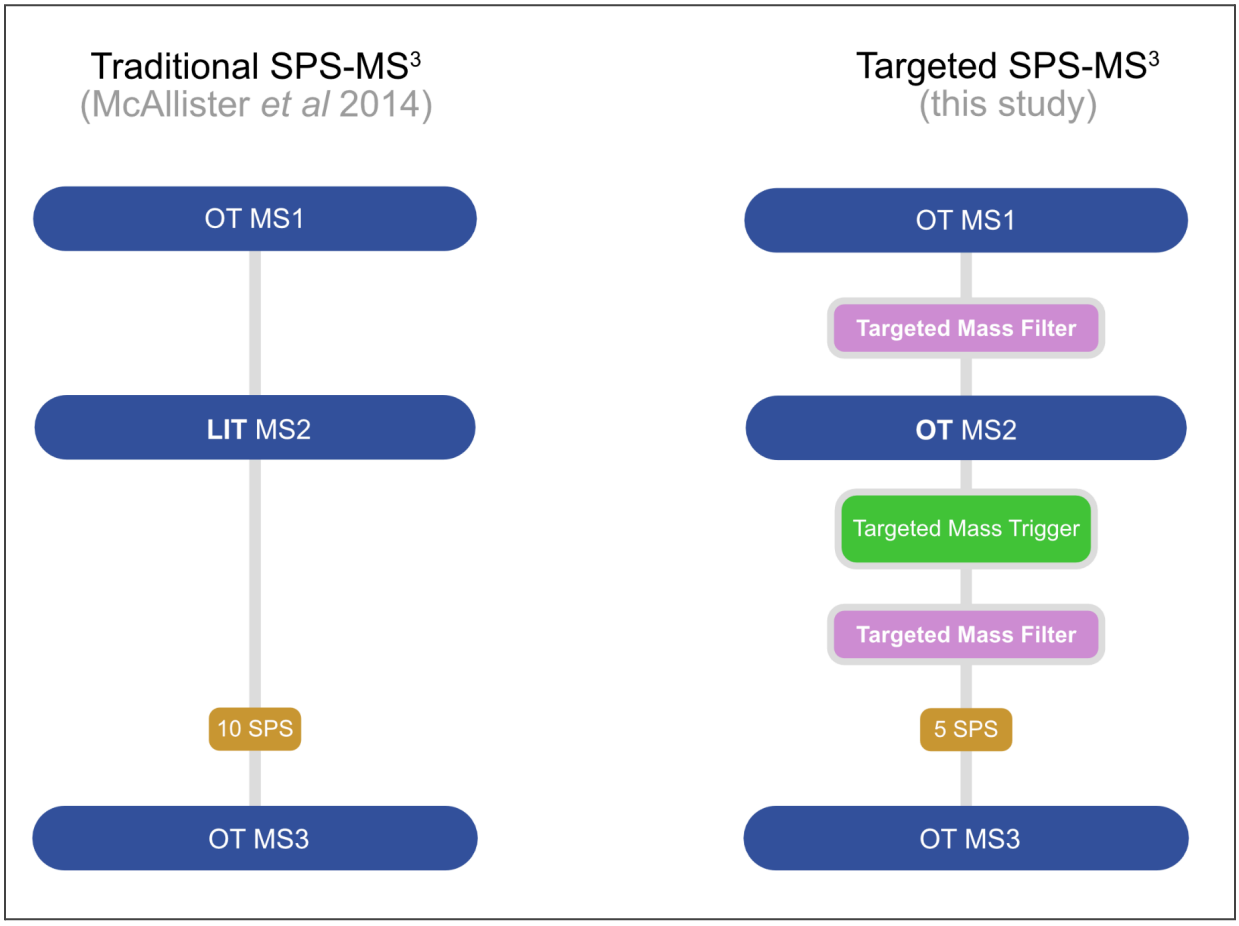
I have tried several computational analyses to identify shared features of candidate phosphosites, including structural predictions(Alphafold(Jumper et al., 2021), FuzDB(Hatos et al., 2022), metapredict (Emenecker et al., 2022)), kinase-substrate enrichment analysis(Wiredja et al., 2017), and motif enrichment. However, these analyses are preliminary and not yet included in this text. There are other tools I would like to explore, including metapredict residue disorder prediction using substituted phosphomimetic residues. I would also like to explore these data at the level of phosphoisoforms to determine the combinatorial impact of multiply phosphorylated peptides.

My goal for the remainder of my time in Genome Sciences is to combine miniTurboID proximity labelling with phosphopeptide enrichment in all three stress conditions. The ultimate goal of this project is to create a resource of SG-localized proteins and phosphopeptides in multiple stress types. SGs are implicated in various neurodegenerative diseases, cancer, viral infection, aging, and environmental stress. My project and data provide a foundation for others to investigate specific signalling mechanisms in SG regulation and dysfunction and ultimately annotate functional phosphosites.

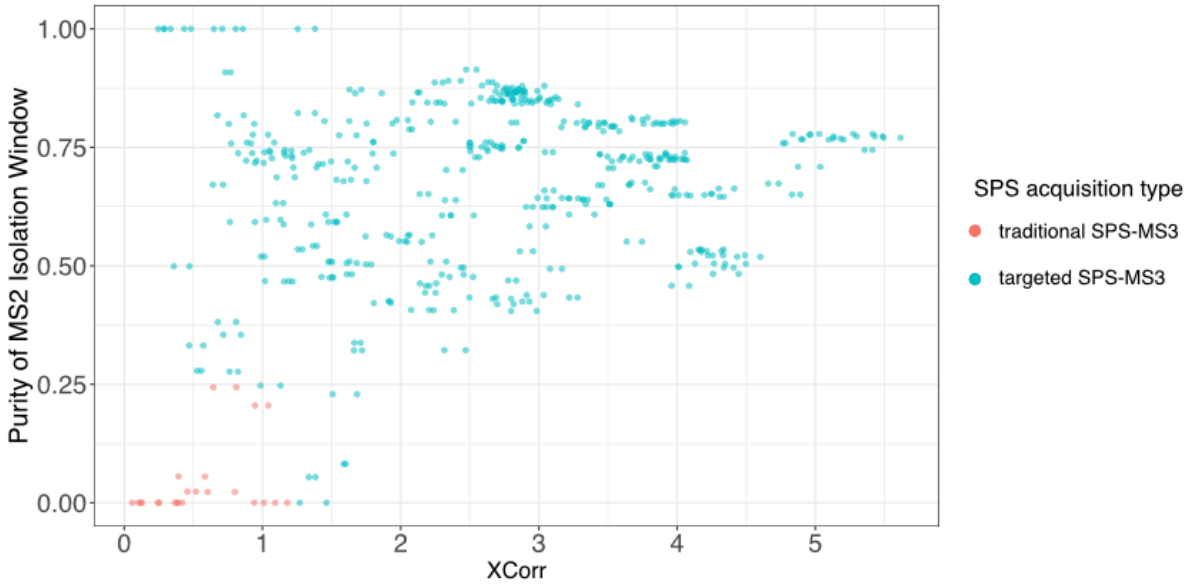
4.4 Concluding remarks

In summary, I have presented two distinct projects that examine proteoforms. One method examines the biophysical properties of disease-associated protein variants. The other examines the stress-specific differences in phosphorylation response in U2-OS cells and uses publicly available data to predict which phosphosites may belong to stress granules. While vastly different, these two projects contribute to the larger effort in proteomics to transition from protein-level readouts to proteoform-level readouts.

APPENDIX A: SUPPLEMENTAL MATERIAL FOR CHAPTER 2



Supplementary Figure A1. *Schematic of SPS-MS3 acquisition strategies.* The SPS-MS3 acquisition method (McAlister et al., 2014) entails a full MS1 (orbitrap OT), and several MS2 scans (linear ion trap LIT), where the number of MS2 scans is limited by duty cycle time. Ten ions are isolated by synchronous precursor selection (SPS) for MS3 in the OT. Our novel targeted SPS-MS3 method entails a full MS1 scan. If target masses are detected, a targeted MS2 scan is triggered in the OT. MS2 fragment masses are linked in the target list to MS1 precursors. If MS2 fragment masses belonging to the corresponding precursor target are detected, five of these MS2 ions undergo SPS and a full MS3 is performed. If, at any point, no targets are identified in the MS1 or MS2 scans, the duty cycle is reset.



Supplementary Fig A2. Comparison of metrics between regular and targeted SPS-MS3. All scans matching PGM1 variant peptides are shown and no filtering was applied. XCorr represents the goodness of fit of experimental versus theoretical spectra in the Comet database search. The Y axis represents the precRepIntRatio value from the IsobaricQuant package. It indicates the cleanliness of MS2 isolation by reporting the proportion of intensity within the MS2 isolation window that belongs to fragment peaks (Hogrebe et al., 2022).

Supplementary Table Ai. *TMT labelling scheme for samples.*

	Condition	
Label	**Solubility**	**Thermal Stability**
TMT126	BPER_Rep1	37 ℃
TMT127N	BPER_Rep2	37.7 ℃
TMT127C	BPER_Rep3	39.5 ℃
TMT128N	Native_Rep1	42.4 ℃
TMT128C	Native_Rep2	46.3 ℃
TMT129N	Native_Rep3	50.1 ℃

	Condition	
Label	**Solubility**	**Thermal Stability**
TMT129C	Denaturing_Rep1	53.8 ℃
TMT130N	Denaturing_Rep2	57.6 ℃
TMT130C	Denaturing_Rep3	61.5 ℃
TMT131N	NA	64.4 ℃
TMT131C	NA	all temperatures pooled

Supplementary Table Aii. *MS1 and MS2 mass targets for Targeted-SPS-MS3 acquisition*

variant	enzyme	sequence	MS1 charge	MS1 m/z	MS2 ion	MS2 charge	MS2 m/z
T19A	trypsin	TQAYQDQKPGASGLR	3	693.3847	y7	1	657.3678
T19A	trypsin	TQAYQDQKPGASGLR	3	693.3847	y11	2	693.3886
T19A	trypsin	TQAYQDQKPGASGLR	3	693.3847	y9	2	571.8458
T19A	trypsin	TQAYQDQKPGASGLR	3	693.3847	b4	1	693.3769
T19A	trypsin	TQAYQDQKPGASGLR	3	693.3847	b6	1	936.4624
N38Y	chymo	QSSANYAEY	2	631.2985	b2	1	445.2608
N38Y	chymo	QSSANYAEY	2	631.2985	b3	1	532.2928
N38Y	chymo	QSSANYAEY	2	631.2985	b4	1	603.33
N38Y	chymo	QSSANYAEY	2	631.2985	b5	1	717.3729
N38Y	chymo	QSSANYAEY	2	631.2985	b6	1	880.4362
Q41R	trypsin	SIISTVEPAQR	3	477.2786	y5	1	600.31
Q41R	trypsin	SIISTVEPAQR	3	477.2786	y4	1	471.2674

variant	enzyme	sequence	MS1 charge	MS1 m/z	MS2 ion	MS2 charge	MS2 m/z
Q41R	trypsin	SIISTVEPAQR	3	477.2786	b2	1	430.2863
Q41R	trypsin	SIISTVEPAQR	3	477.2786	b3	1	543.3704
Q41R	trypsin	SIISTVEPAQR	3	477.2786	b5	1	731.4501
Q41R	trypsin	QEATLVVGGDGR	3	477.5982	y7	1	659.3471
Q41R	trypsin	QEATLVVGGDGR	3	477.5982	y6	1	560.2787
Q41R	trypsin	QEATLVVGGDGR	3	477.5982	y5	1	461.2103
Q41R	trypsin	QEATLVVGGDGR	3	477.5982	b4	1	659.3562
Q41R	trypsin	QEATLVVGGDGR	3	477.5982	b5	1	772.4402
Q41R	trypsin	SIISTVEPAQR	2	715.4142	y8	1	887.4581
Q41R	trypsin	SIISTVEPAQR	2	715.4142	y4	1	471.2674
Q41R	trypsin	SIISTVEPAQR	2	715.4142	b2	1	430.2863
Q41R	trypsin	SIISTVEPAQR	2	715.4142	b5	1	731.4501
Q41R	trypsin	SIISTVEPAQR	2	715.4142	b7	1	959.5611
D62H	trypsin	QEATLVVGGHGR	3	484.9422	y6	1	582.3107
D62H	trypsin	QEATLVVGGHGR	3	484.9422	y5	1	483.2423
D62H	trypsin	QEATLVVGGHGR	3	484.9422	b3	1	558.3085
D62H	trypsin	QEATLVVGGHGR	3	484.9422	b4	1	659.3562
D62H	trypsin	QEATLVVGGHGR	3	484.9422	b5	1	772.4402
T115A	chymo	AASHNPGGPNGDF	2	735.3522	b2	1	372.2444

variant	enzyme	sequence	MS1 charge	MS1 m/z	MS2 ion	MS2 charge	MS2 m/z
T115A	chymo	AASHNPGGPNGDF	2	735.3522	b3	1	459.2765
T115A	chymo	AASHNPGGPNGDF	2	735.3522	b4	1	596.3354
T115A	chymo	AASHNPGGPNGDF	2	735.3522	b5	1	710.3783
T115A	chymo	AASHNPGGPNGDF	2	735.3522	b5	2	355.6928
G121R	trypsin	AIGGIILTASHNPR	3	550.3281	y8	1	895.4744
G121R	trypsin	AIGGIILTASHNPR	3	550.3281	y7	1	782.3904
G121R	trypsin	AIGGIILTASHNPR	3	550.3281	b4	1	528.3343
G121R	trypsin	AIGGIILTASHNPR	3	550.3281	b5	1	641.4184
G121R	trypsin	AIGGIILTASHNPR	3	550.3281	b6	1	754.5024
D263G	chymo	EDFGGHHPGPNLTY	3	590.6197	b6	1	872.41
D263G	chymo	EDFGGHHPGPNLTY	3	590.6197	b7	1	1009.4689
D263G	chymo	EDFGGHHPGPNLTY	3	590.6197	b7	2	505.2381
D263G	chymo	EDFGGHHPGPNLTY	3	590.6197	b9	2	582.2752
D263G	chymo	EDFGGHHPGPNLTY	3	590.6197	b11	2	687.8231
D263Y	chymo	EDFGGHHPYPNLTY	3	625.967	b6	1	872.41
D263Y	chymo	EDFGGHHPYPNLTY	3	625.967	b7	1	1009.4689
D263Y	chymo	EDFGGHHPYPNLTY	3	625.967	b7	2	505.2381
D263Y	chymo	EDFGGHHPYPNLTY	3	625.967	b9	2	635.2961
D263Y	chymo	EDFGGHHPYPNLTY	3	625.967	b11	2	740.844

variant	enzyme	sequence	MS1 charge	MS1 m/z	MS2 ion	MS2 charge	MS2 m/z
G291R	trypsin	SGEHDFGAAFDGDR	3	570.5954	y8	1	808.3584
G291R	trypsin	SGEHDFGAAFDGDR	3	570.5954	y3	1	347.1674
G291R	trypsin	SGEHDFGAAFDGDR	3	570.5954	b3	1	503.2663
G291R	trypsin	SGEHDFGAAFDGDR	3	570.5954	b4	2	320.6662
G291R	trypsin	SGEHDFGAAFDGDR	3	570.5954	b5	2	378.1797
G330R	chymo	QQTGVRRF	3	407.5736	y2	1	322.1874
G330R	chymo	QQTGVRRF	3	407.5736	y6	2	368.2167
G330R	chymo	QQTGVRRF	3	407.5736	b3	1	587.335
G330R	chymo	QQTGVRRF	3	407.5736	b2	3	162.7673
G330R	chymo	FQQTGVRRF	3	456.5964	y2	1	322.1874
G330R	chymo	FQQTGVRRF	3	456.5964	y7	2	432.2459
G330R	chymo	FQQTGVRRF	3	456.5964	b2	1	505.2972
G330R	chymo	FQQTGVRRF	3	456.5964	b7	1	1046.5944
G330R	chymo	FQQTGVRRF	3	456.5964	b7	2	523.8009
G330R	chymo	QQTGVRRF	2	610.8567	y4	1	577.3569
G330R	chymo	QQTGVRRF	2	610.8567	y2	1	322.1874
G330R	chymo	QQTGVRRF	2	610.8567	b2	1	486.2874
G330R	chymo	QQTGVRRF	2	610.8567	b3	1	587.335
G330R	chymo	QQTGVRRF	2	610.8567	b5	1	743.4249

variant	enzyme	sequence	MS1 charge	MS1 m/z	MS2 ion	MS2 charge	MS2 m/z
G330R	chymo	FQQTGVRRF	2	684.3909	b7	1	1046.5944
G330R	chymo	FQQTGVRRF	2	684.3909	b7	2	523.8009
E377K	trypsin	SFGTGSDHIR	3	435.8965	y6	1	684.3424
E377K	trypsin	SFGTGSDHIR	3	435.8965	y3	1	425.2619
E377K	trypsin	SFGTGSDHIR	3	435.8965	y2	1	288.203
E377K	trypsin	SFGTGSDHIR	3	435.8965	b2	1	464.2707
E377K	trypsin	SFGTGSDHIR	3	435.8965	b3	1	521.2921
E388K	LysC	LSLC[57]GEESFGTGSDHIRK	4	613.573	y11	2	717.3886
E388K	LysC	LSLC[57]GEESFGTGSDHIRK	4	613.573	y10	2	673.8725
E388K	LysC	LSLC[57]GEESFGTGSDHIRK	4	613.573	y9	2	600.3383
E388K	LysC	LSLC[57]GEESFGTGSDHIRK	4	613.573	y4	2	391.7635
E388K	LysC	LSLC[57]GEESFGTGSDHIRK	4	613.573	b4	1	703.401
L516P	trypsin	LSGTGSAGATIRPYIDSYEK	3	848.7952	y8	1	1243.6408
L516P	trypsin	LSGTGSAGATIRPYIDSYEK	3	848.7952	y5	1	870.4406
L516P	trypsin	LSGTGSAGATIRPYIDSYEK	3	848.7952	y4	1	755.4137
L516P	trypsin	LSGTGSAGATIRPYIDSYEK	3	848.7952	b12	2	651.3724
L516P	trypsin	LSGTGSAGATIRPYIDSYEK	3	848.7952	b16	2	895.4859

APPENDIX B: SUPPLEMENTAL MATERIAL FOR CHAPTER 3

Supplementary Table B-i. *Candidate phosphosites in stress granules.* The treatment columns provide information on the regulation of that site. Significance (significant or not significant NS) is determined based on the modified t-test. The direction of regulation is given by up, down or none. Data missingness is all complete, where the site is identified in at least three of the six biological replicates.

UniProt accession	Residue position	Protein	Gene	Arsenite	Sorbitol	Heat
C9JLW8	S21	FAM195B	MCRIP1	significant_up	significant_none	significant_up
C9JLW8	T16	FAM195B	MCRIP1	significant_up	significant_up	significant_up
O00571	S594	DDX3	DDX3X	significant_up	NS_none	NS_none
O43379	S1048	WDR62	WDR62	NS_none	significant_down	NS_none
O43379	T1053	WDR62	WDR62	NS_none	significant_down	NS_none
O60832	S494	dyskerin	DKC1	NS_none	significant_down	NS_none
O60832	S485	dyskerin	DKC1	NS_none	NS_none	significant_down
O75152	T762	ZC3H11A	ZC3H11A	significant_up	significant_up	significant_up
O75152	S761	ZC3H11A	ZC3H11A	significant_up	significant_up	significant_up
O75821	T38	eIF3-delta	EIF3G	NS_none	significant_down	NS_none
O95817	T285	BAG3	BAG3	significant_up	NS_none	significant_down
O95817	S173	BAG3	BAG3	NS_none	NS_none	significant_down
O95817	S381	BAG3	BAG3	significant_up	significant_up	significant_up
O95817	S377	BAG3	BAG3	significant_up	significant_up	significant_up
O95817	S279	BAG3	BAG3	NS_none	significant_down	significant_down
O95817	S291	BAG3	BAG3	NS_up	NS_none	significant_down
O95817	S275	BAG3	BAG3	NS_none	significant_down	significant_down
O95817	S289	BAG3	BAG3	significant_up	NS_none	significant_down
O95817	S386	BAG3	BAG3	significant_none	significant_up	significant_up
O95817	T307	NA	BAG3	significant_up	NS_none	NS_none
P04792	S82	HSP27	HSPB1	significant_up	significant_up	significant_up

UniProt accession	Residue position	Protein	Gene	Arsenite	Sorbitol	Heat
P07814	S883	EPRS	EPRS1	NS_none	NS_none	significant_down
P10636	S717	Tau	MAPT	NS_none	significant_down	significant_down
P10636	S713	Tau	MAPT	NS_none	significant_down	significant_down
P10636	S721	Tau	MAPT	NS_none	significant_down	significant_down
P18615	S115	RDBP	NELFE	significant_up	significant_none	significant_up
P18615	S181	RDBP	NELFE	significant_up	NS_none	NS_none
P23443	S447	p70S6K	RPS6KB1	significant_up	significant_up	significant_up
P23443	T444	p70S6K	RPS6KB1	significant_up	significant_up	significant_up
P23588	S497	EIF4B	EIF4B	significant_up	significant_up	NS_none
P23588	S498	EIF4B	EIF4B	significant_up	NS_none	NS_none
P23588	T420	EIF4B	EIF4B	significant_up	NS_none	NS_down
P23588	S425	EIF4B	EIF4B	significant_up	NS_none	NS_down
P23588	S422	EIF4B	EIF4B	significant_up	NS_none	NS_none
P23588	S418	EIF4B	EIF4B	significant_up	significant_down	NS_none
P23588	S409	EIF4B	EIF4B	significant_up	NS_none	significant_down
P23588	S406	EIF4B	EIF4B	NS_none	NS_none	significant_down
P23588	S430	EIF4B	EIF4B	significant_up	NS_none	significant_down
P23588	T450	EIF4B	EIF4B	NS_up	significant_up	NS_up
P35611	S431	ADD1	ADD1	significant_up	NS_none	NS_none
P35611	S436	ADD1	ADD1	significant_up	NS_down	NS_none
P38919	S12	DDX48	EIF4A3	significant_up	significant_down	NS_none
P49756	S677	RBM25	RBM25	NS_none	significant_down	NS_down
P49756	S683	RBM25	RBM25	significant_up	NS_down	NS_down
P49790	S333	NUP153	NUP153	significant_up	significant_up	significant_up
P49790	S338	NUP153	NUP153	significant_up	significant_up	significant_up
P49790	S334	NUP153	NUP153	significant_up	NS_up	significant_up
P50579	S60	METAP2	METAP2	significant_up	NS_none	NS_none
P51114	T483	FXR1	FXR1	NS_none	NS_none	significant_down

UniProt accession	Residue position	Protein	Gene	Arsenite	Sorbitol	Heat
P51114	S478	FXR1	FXR1	NS_none	NS_none	significant_down
P51114	S485	FXR1	FXR1	NS_none	NS_none	significant_down
P51114	S420	FXR1	FXR1	NS_none	NS_none	significant_down
P51114	T411	FXR1	FXR1	NS_none	NS_none	significant_down
P51116	S603	FXR2	FXR2	NS_none	significant_down	significant_down
P51116	S601	FXR2	FXR2	NS_none	significant_down	significant_down
P51116	S520	FXR2	FXR2	significant_up	NS_down	NS_none
P51116	S533	FXR2	FXR2	significant_up	NS_down	NS_none
P52292	S62	KPNA2	KPNA2	significant_up	significant_up	significant_up
P52948	S1060	NUP98	NUP98	significant_up	significant_up	significant_up
P52948	T1070	NUP98	NUP98	significant_up	significant_up	significant_up
P52948	S888	NUP98	NUP98	NS_none	significant_down	NS_none
P52948	S608	NUP98	NUP98	NS_none	NS_up	significant_up
P52948	S623	NUP98	NUP98	significant_up	NS_none	significant_up
P52948	S714	NUP98	NUP98	significant_up	NS_none	NS_up
P52948	S705	NA	NUP98	significant_up	NS_up	NS_up
P52948	S934	NUP98	NUP98	NS_up	NS_none	significant_up
P55884	S78	eIF3-eta	EIF3B	NS_none	NS_none	significant_down
P55884	S95	eIF3-eta	EIF3B	significant_down	NS_down	NS_down
P62753	T241	S6	RPS6	NS_none	significant_down	NS_none
P62753	S247	S6	RPS6	NS_none	NS_down	significant_down
P62753	S240	S6	RPS6	NS_none	significant_down	significant_up
P62753	S244	S6	RPS6	NS_none	significant_down	NS_none
Q04637	T205	eIF4G	EIF4G1	significant_none	NS_none	significant_down
Q04637	S1187	eIF4G	EIF4G1	significant_up	NS_none	significant_down
Q04637	T207	eIF4G	EIF4G1	NS_none	NS_none	significant_down
Q04637	S1194	eIF4G	EIF4G1	significant_up	significant_down	significant_down
Q04637	S1185	eIF4G	EIF4G1	significant_up	significant_down	significant_down

UniProt accession	Residue position	Protein	Gene	Arsenite	Sorbitol	Heat
Q04637	T211	eIF4G	EIF4G1	NS_none	NS_none	significant_down
Q07955	S201	SF2	SRSF1	NS_none	significant_down	NS_none
Q12906	S506	NFAT90	ILF3	significant_down	NS_none	NS_none
Q12906	S382	NFAT90	ILF3	NS_none	significant_down	NS_none
Q12906	T364	NFAT90	ILF3	significant_up	NS_none	NS_none
Q13242	S211	SFRS9	SRSF9	NS_none	significant_down	significant_down
Q13242	S216	SFRS9	SRSF9	NS_none	significant_down	significant_down
Q13769	S312	FMIP	THOC5	NS_none	NS_none	significant_down
Q13769	S314	FMIP	THOC5	NS_none	NS_none	significant_down
Q14694	S226	USP10	USP10	significant_up	NS_none	significant_up
Q14694	S218	USP10	USP10	NS_down	NS_down	significant_down
Q14978	S623	NOLC1	NOLC1	significant_up	NS_none	significant_down
Q14978	S622	NOLC1	NOLC1	significant_up	NS_none	NS_none
Q14978	T607	NOLC1	NOLC1	significant_up	NS_none	NS_none
Q14978	T610	NOLC1	NOLC1	significant_up	NS_none	NS_none
Q15637	S89	SF1	SF1	NS_none	NS_down	significant_up
Q15654	S142	TRIP6	TRIP6	significant_up	NS_none	NS_none
Q15654	T153	TRIP6	TRIP6	significant_up	NS_none	NS_none
Q16637	T25	SMN	SMN1	NS_none	significant_down	significant_down
Q16637	S28	SMN	SMN1	NS_none	significant_down	significant_down
Q16637	S31	SMN	SMN1	NS_none	significant_down	significant_down
Q6Y7W6	S26	TNRC15	GIGYF2	significant_up	NS_none	NS_none
Q6Y7W6	S30	TNRC15	GIGYF2	significant_up	NS_none	NS_none
Q6Y7W6	T382	TNRC15	GIGYF2	NS_none	significant_down	NS_none
Q6Y7W6	S236	TNRC15	GIGYF2	significant_up	significant_up	significant_up
Q6Y7W6	S340	TNRC15	GIGYF2	NS_none	NS_down	significant_down
Q6Y7W6	S335	TNRC15	GIGYF2	NS_none	NS_down	significant_down
Q6Y7W6	S337	TNRC15	GIGYF2	NS_none	NS_down	significant_down

UniProt accession	Residue position	Protein	Gene	Arsenite	Sorbitol	Heat
Q6Y7W6	S160	TNRC15	GIGYF2	significant_up	NS_none	NS_up
Q7Z2W4	S275	ZAP	ZC3HAV1	NS_none	NS_none	significant_down
Q7Z2W4	T273	ZAP	ZC3HAV1	NS_none	NS_none	significant_down
Q7Z417	S212	82-FIP	NUFIP2	NS_none	significant_down	significant_down
Q7Z417	S214	82-FIP	NUFIP2	NS_none	significant_down	significant_down
Q7Z417	T220	82-FIP	NUFIP2	significant_up	NS_down	NS_none
Q7Z434	S165	MAVS	MAVS	NS_none	NS_none	significant_up
Q8ND56	S178	alphaSNBP(A)	LSM14A	significant_up	NS_none	significant_none
Q8ND56	S216	alphaSNBP(A)	LSM14A	NS_none	significant_down	NS_none
Q8ND56	S192	alphaSNBP(A)	LSM14A	NS_none	significant_down	NS_none
Q8ND56	S183	alphaSNBP(A)	LSM14A	NS_none	significant_down	NS_none
Q8NE71	S140	ABCF1	ABCF1	NS_none	significant_down	NS_none
Q8NE71	S109	ABCF1	ABCF1	NS_none	significant_down	significant_down
Q8NE71	T108	ABCF1	ABCF1	NS_none	significant_down	significant_down
Q8NE71	S105	ABCF1	ABCF1	NS_none	significant_down	significant_down
Q8NI27	T1285	THOC2	THOC2	significant_down	NS_down	NS_none
Q8NI27	T1288	THOC2	THOC2	significant_down	NS_down	NS_none
Q8TB72	S136	Pumilio-2	PUM2	NS_none	significant_down	NS_none
Q8WWM7	S335	ataxin-2L	ATXN2L	significant_up	significant_up	NS_none
Q8WWM7	S423	ataxin-2L	ATXN2L	significant_up	NS_none	significant_none
Q8WWM7	S339	ataxin-2L	ATXN2L	significant_up	NS_none	significant_down
Q8WWM7	S680	ataxin-2L	ATXN2L	significant_up	significant_up	NS_none
Q8WWM7	S491	ataxin-2L	ATXN2L	significant_up	NS_up	NS_none
Q8WWM7	S485	NA	ATXN2L	significant_up	significant_up	NS_none
Q8WWM7	S496	ataxin-2L	ATXN2L	significant_up	NS_none	NS_none
Q8WWM7	S56	ataxin-2L	ATXN2L	significant_up	NS_up	NS_none
Q96HA1	S179	POM121	POM121	NS_none	significant_up	significant_up

UniProt accession	Residue position	Protein	Gene	Arsenite	Sorbitol	Heat
Q96HA1	S188	POM121	POM121	NS_none	significant_none	significant_up
Q96I25	S155	RBM17	RBM17	NS_none	significant_down	NS_none
Q96R06	S66	SPAG5	SPAG5	significant_up	NS_none	NS_none
Q96T37	S656	RBM15	RBM15	NS_none	significant_down	NS_none
Q96T37	S659	RBM15	RBM15	NS_none	significant_down	NS_none
Q96T37	S680	RBM15	RBM15	NS_up	significant_up	significant_up
Q96T37	S765	RBM15	RBM15	NS_none	significant_down	NS_none
Q99700	S758	ataxin-2	ATXN2	NS_none	significant_down	NS_none
Q99700	S772	ataxin-2	ATXN2	NS_none	significant_down	NS_none
Q9BYJ9	S291	YTHDF1	YTHDF1	significant_up	NS_none	NS_none
Q9H0D6	S487	XRN2	XRN2	significant_up	significant_up	NS_none
Q9H0D6	S448	XRN2	XRN2	significant_up	significant_none	NS_none
Q9H0D6	S499	XRN2	XRN2	NS_none	significant_down	significant_down
Q9H0D6	S501	XRN2	XRN2	NS_none	significant_down	significant_down
Q9H0D6	S475	XRN2	XRN2	NS_none	significant_up	NS_none
Q9UPU9	S272	SAMD4	SAMD4A	significant_up	significant_up	NS_none
Q9UPU9	S420	SAMD4	SAMD4A	significant_down	NS_none	significant_down
Q9Y520	T1244	PRRC2C	PRRC2C	significant_down	significant_down	NS_none
Q9Y520	S1242	PRRC2C	PRRC2C	significant_down	significant_down	NS_down
Q9Y520	S636	PRRC2C	PRRC2C	significant_up	significant_up	significant_none
Q9Y520	S633	PRRC2C	PRRC2C	significant_up	significant_up	significant_up
Q9Y520	S187	PRRC2C	PRRC2C	significant_up	significant_none	NS_none
Q9Y520	S926	PRRC2C	PRRC2C	significant_up	significant_down	NS_up
Q9Y520	T2682	PRRC2C	PRRC2C	significant_up	NS_none	NS_down
Q9Y520	S2685	PRRC2C	PRRC2C	significant_none	NS_none	significant_down
Q9Y520	S1246	PRRC2C	PRRC2C	NS_none	significant_down	NS_none

BIBLIOGRAPHY

Advani, V. M., & Ivanov, P. (2020). Stress granule subtypes: an emerging link to neurodegeneration. *Cellular and Molecular Life Sciences: CMLS*, *77*(23), 4827–4845. https://doi.org/10.1007/s00018-020-03565-0

Aebersold, R., Agar, J. N., Amster, I. J., Baker, M. S., Bertozzi, C. R., Boja, E. S., Costello, C. E., Cravatt, B. F., Fenselau, C., Garcia, B. A., Ge, Y., Gunawardena, J., Hendrickson, R. C., Hergenrother, P. J., Huber, C. G., Ivanov, A. R., Jensen, O. N., Jewett, M. C., Kelleher, N. L., … Zhang, B. (2018). How many human proteoforms are there? *Nature Chemical Biology*, *14*(3), 206–214. https://doi.org/10.1038/nchembio.2576

Ali, E. S., Lipońska, A., O'Hara, B. P., Amici, D. R., Torno, M. D., Gao, P., Asara, J. M., Yap, M.-N. F., Mendillo, M. L., & Ben-Sahra, I. (2022). The mTORC1-SLC4A7 axis stimulates bicarbonate import to enhance de novo nucleotide synthesis. *Molecular Cell*, *82*(17), 3284–3298.e7. https://doi.org/10.1016/j.molcel.2022.06.008

Almeida-Souza, L., Goethals, S., de Winter, V., Dierick, I., Gallardo, R., Van Durme, J., Irobi, J., Gettemans, J., Rousseau, F., Schymkowitz, J., Timmerman, V., & Janssens, S. (2010). Increased monomerization of mutant HSPB1 leads to protein hyperactivity in Charcot-Marie-Tooth neuropathy. *The Journal of Biological Chemistry*, *285*(17), 12778–12786. https://doi.org/10.1074/jbc.M109.082644

Altassan, R., Radenkovic, S., Edmondson, A. C., Barone, R., Brasil, S., Cechova, A., Coman, D., Donoghue, S., Falkenstein, K., Ferreira, V., Ferreira, C., Fiumara, A., Francisco, R., Freeze, H., Grunewald, S., Honzik, T., Jaeken, J., Krasnewich, D., Lam, C., … Morava, E. (2021). International consensus guidelines for phosphoglucomutase 1 deficiency (PGM1-CDG): Diagnosis, follow-up, and management. *Journal of Inherited Metabolic Disease*, *44*(1), 148–163. https://doi.org/10.1002/jimd.12286

Amberger, J. S., Bocchini, C. A., Scott, A. F., & Hamosh, A. (2019). OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Research*, *47*(D1), D1038–D1043. https://doi.org/10.1093/nar/gky1151

Apgar, T. L., & Sanders, C. R. (2022). Compendium of causative genes and their encoded proteins for common monogenic disorders. *Protein Science: A Publication of the Protein Society*, *31*(1), 75–91. https://doi.org/10.1002/pro.4183

Argentini, A., Goeminne, L. J. E., Verheggen, K., Hulstaert, N., Staes, A., Clement, L., & Martens, L. (2016). moFF: a robust and automated approach to extract peptide ion intensities. *Nature Methods*, *13*(12), 964–966. https://doi.org/10.1038/nmeth.4075

Arribas Diez, I., Govender, I., Naicker, P., Stoychev, S., Jordaan, J., & Jensen, O. N. (2021). Zirconium(IV)-IMAC revisited: Improved performance and phosphoproteome coverage by magnetic microparticles for phosphopeptide affinity enrichment. *Journal of Proteome Research*, *20*(1), 453–462. https://doi.org/10.1021/acs.jproteome.0c00508

Barente, A. S., & Villén, J. (2022). A Python Package for the Localization of Protein Modifications in Mass Spectrometry Data. *Journal of Proteome Research*. https://doi.org/10.1021/acs.jproteome.2c00194

Beamer, L. J. (2020). Enzyme dysfunction at atomic resolution: Disease-associated variants of human phosphoglucomutase-1. *Biochimie*. https://doi.org/10.1016/j.biochi.2020.08.017

Beausoleil, S. A., Villén, J., Gerber, S. A., Rush, J., & Gygi, S. P. (2006). A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nature Biotechnology*, *24*(10), 1285–1292. https://doi.org/10.1038/nbt1240

Berg, M. D., Zhu, Y., Genereaux, J., Ruiz, B. Y., Rodriguez-Mias, R. A., Allan, T., Bahcheli, A., Villén, J., & Brandl, C. J. (2019). Modulating mistranslation potential of tRNASer in Saccharomyces cerevisiae. *Genetics*, *213*(3), 849–863. https://doi.org/10.1534/genetics.119.302525

Bilbrough, T., Piemontese, E., & Seitz, O. (2022). Dissecting the role of protein phosphorylation: a chemical biology toolbox. *Chemical Society Reviews*, *51*(13), 5691–5730.

https://doi.org/10.1039/d1cs00991e
Bortel, P., Piga, I., Koenig, C., Gerner, C., Martinez-Val, A., & Olsen, J. V. (2024). Systematic optimization of automated phosphopeptide enrichment for high-sensitivity phosphoproteomics. *Molecular & Cellular Proteomics: MCP*, *23*(5), 100754. https://doi.org/10.1016/j.mcpro.2024.100754
Boyd, R. A., Majumder, S., Stiban, J., Mavodza, G., Straus, A. J., Kempelingaiah, S. K., Reddy, V., Hannun, Y. A., Obeid, L. M., & Senkal, C. E. (2023). The heat shock protein Hsp27 controls mitochondrial function by modulating ceramide generation. *Cell Reports*, *42*(9), 113081. https://doi.org/10.1016/j.celrep.2023.113081
Brocker, C., Thompson, D. C., & Vasiliou, V. (2012). The role of hyperosmotic stress in inflammation and disease. *Biomolecular Concepts*, *3*(4), 345–364. https://doi.org/10.1515/bmc-2012-0001
Brusilow, S. W., & Cooper, A. J. L. (2011). Encephalopathy in acute liver failure resulting from acetaminophen intoxication: new observations with potential therapy. *Critical Care Medicine*, *39*(11), 2550–2553. https://doi.org/10.1097/CCM.0b013e31822572fd
Chang, A., Leutert, M., Rodriguez-Mias, R. A., & Villén, J. (2023). Automated enrichment of phosphotyrosine peptides for high-throughput proteomics. *Journal of Proteome Research*, *22*(6), 1868–1880. https://doi.org/10.1021/acs.jproteome.2c00850
Childs, D., Kurzawa, N., Franken, H., Doce, C., Savitski, M., & Huber, W. (2020). *TPP: Analyze thermal proteome profiling (TPP) experiments* (Version 3.18.0). https://rdrr.io/bioc/TPP/
Chinnaraj, M., Huang, H., Hutchinson, S., Meyer, M., Pike, D., Ribezzi, M., Sultana, S., Ocampo, D., Ding, F., Carpenter, M. L., Chorny, I., & Vieceli, J. (2025). Protein barcoding and Next-generation protein sequencing for multiplexed protein selection, analysis, and tracking. In *bioRxiv* (p. 2024.12.31.630920). https://doi.org/10.1101/2024.12.31.630920
Cocco, E., Lopez, S., Santin, A. D., & Scaltriti, M. (2019). Prevalence and role of HER2 mutations in cancer. *Pharmacology & Therapeutics*, *199*, 188–196. https://doi.org/10.1016/j.pharmthera.2019.03.010
Collier, N. C., & Schlesinger, M. J. (1986). The dynamic state of heat shock proteins in chicken embryo fibroblasts. *The Journal of Cell Biology*, *103*(4), 1495–1507. https://doi.org/10.1083/jcb.103.4.1495
Cozma, E., Rao, M., Dusick, M., Genereaux, J., Rodriguez-Mias, R. A., Villén, J., Brandl, C. J., & Berg, M. D. (2023). Anticodon sequence determines the impact of mistranslating tRNAAla variants. *RNA Biology*, *20*(1), 791–804. https://doi.org/10.1080/15476286.2023.2257471
Damgaard, R. B. (2021). The ubiquitin system: from cell signalling to disease biology and new therapeutic opportunities. *Cell Death and Differentiation*, *28*(2), 423–426. https://doi.org/10.1038/s41418-020-00703-w
Dayon, L., Hainard, A., Licker, V., Turck, N., Kuhn, K., Hochstrasser, D. F., Burkhard, P. R., & Sanchez, J.-C. (2008). Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Analytical Chemistry*, *80*(8), 2921–2931. https://doi.org/10.1021/ac702422x
Dickinson, K. M., & Collaco, J. M. (2021). Cystic Fibrosis. *Pediatrics in Review / American Academy of Pediatrics*, *42*(2), 55–67. https://doi.org/10.1542/pir.2019-0212
Dionne, U., & Gingras, A.-C. (2022). Proximity-Dependent Biotinylation Approaches to Explore the Dynamic Compartmentalized Proteome. *Frontiers in Molecular Biosciences*, *9*, 852911. https://doi.org/10.3389/fmolb.2022.852911
Dorschner, M. O., Amendola, L. M., Turner, E. H., Robertson, P. D., Shirts, B. H., Gallego, C. J., Bennett, R. L., Jones, K. L., Tokita, M. J., Bennett, J. T., Kim, J. H., Rosenthal, E. A., Kim, D. S., National Heart, Lung, and Blood Institute Grand Opportunity Exome Sequencing Project, Tabor, H. K., Bamshad, M. J., Motulsky, A. G., Scott, C. R., Pritchard, C. C., … Jarvik, G. P. (2013). Actionable, pathogenic incidental findings in 1,000 participants' exomes. *The American Journal of Human Genetics*, *93*(4), 631–640. https://doi.org/10.1016/j.ajhg.2013.08.006
Egloff, P., Zimmermann, I., Arnold, F. M., Hutter, C. A. J., Morger, D., Opitz, L., Poveda, L., Keserue, H.-A., Panse, C., Roschitzki, B., & Seeger, M. A. (2019). Engineered peptide barcodes for in-depth analyses of binding protein libraries. *Nature Methods*, *16*(5), 421–428.

https://doi.org/10.1038/s41592-019-0389-8
Emenecker, R. J., Griffith, D., & Holehouse, A. S. (2022). Metapredict V2: An update to metapredict, a fast, accurate, and easy-to-use predictor of consensus disorder and structure. In *bioRxiv* (p. 2022.06.06.494887). https://doi.org/10.1101/2022.06.06.494887
Eng, J. K., Hoopmann, M. R., Jahan, T. A., Egertson, J. D., Noble, W. S., & MacCoss, M. J. (2015). A deeper look into Comet--implementation and features. *Journal of the American Society for Mass Spectrometry*, *26*(11), 1865–1874. https://doi.org/10.1007/s13361-015-1179-x
Eng, J. K., Jahan, T. A., & Hoopmann, M. R. (2013). Comet: an open-source MS/MS sequence database search tool. *Proteomics*, *13*(1), 22–24. https://doi.org/10.1002/pmic.201200439
Erickson, B. K., Rose, C. M., Braun, C. R., Erickson, A. R., Knott, J., McAlister, G. C., Wühr, M., Paulo, J. A., Everley, R. A., & Gygi, S. P. (2017). A strategy to combine sample multiplexing with targeted proteomics assays for high-throughput protein signature characterization. *Molecular Cell*, *65*(2), 361–370. https://doi.org/10.1016/j.molcel.2016.12.005
Fayer, S., Horton, C., Dines, J. N., Rubin, A. F., Richardson, M. E., McGoldrick, K., Hernandez, F., Pesaran, T., Karam, R., Shirts, B. H., Fowler, D. M., & Starita, L. M. (2021). Closing the gap: Systematic integration of multiplexed functional data resolves variants of uncertain significance in BRCA1, TP53, and PTEN. *The American Journal of Human Genetics*, *108*(12), 2248–2258. https://doi.org/10.1016/j.ajhg.2021.11.001
Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., & White, F. M. (2002). Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nature Biotechnology*, *20*(3), 301–305. https://doi.org/10.1038/nbt0302-301
Fischer, E. H., & Krebs, E. G. (1955). Conversion of phosphorylase b to phosphorylase a in muscle extracts. *The Journal of Biological Chemistry*, *216*(1), 121–132. https://doi.org/10.1016/s0021-9258(19)52289-x
Fortuno, C., Lee, K., Olivier, M., Pesaran, T., Mai, P. L., de Andrade, K. C., Attardi, L. D., Crowley, S., Evans, D. G., Feng, B.-J., Foreman, A. K. M., Frone, M. N., Huether, R., James, P. A., McGoldrick, K., Mester, J., Seifert, B. A., Slavin, T. P., Witkowski, L., … ClinGen TP53 Variant Curation Expert Panel. (2021). Specifications of the ACMG/AMP variant interpretation guidelines for germline TP53 variants. *Human Mutation*, *42*(3), 223–236. https://doi.org/10.1002/humu.24152
Fowler, D. M., Araya, C. L., Fleishman, S. J., Kellogg, E. H., Stephany, J. J., Baker, D., & Fields, S. (2010). High-resolution mapping of protein sequence-function relationships. *Nature Methods*, *7*(9), 741–746. https://doi.org/10.1038/nmeth.1492
Franken, H., Mathieson, T., Childs, D., Sweetman, G. M. A., Werner, T., Tögel, I., Doce, C., Gade, S., Bantscheff, M., Drewes, G., Reinhard, F. B. M., Huber, W., & Savitski, M. M. (2015). Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nature Protocols*, *10*(10), 1567–1593. https://doi.org/10.1038/nprot.2015.101
Funk, J. S., Klimovich, M., Drangenstein, D., Pielhoop, O., Hunold, P., Borowek, A., Noeparast, M., Pavlakis, E., Neumann, M., Balourdas, D.-I., Kochhan, K., Merle, N., Bullwinkel, I., Wanzel, M., Elmshäuser, S., Teply-Szymanski, J., Nist, A., Procida, T., Bartkuhn, M., … Stiewe, T. (2025). Deep CRISPR mutagenesis characterizes the functional diversity of TP53 mutations. *Nature Genetics*, *57*(1), 140–153. https://doi.org/10.1038/s41588-024-02039-4
Gates, M. B., Tomer, K. B., & Deterding, L. J. (2010). Comparison of metal and metal oxide media for phosphopeptide enrichment prior to mass spectrometric analyses. *Journal of the American Society for Mass Spectrometry*, *21*(10), 1649–1659. https://doi.org/10.1016/j.jasms.2010.06.005
Gurgis, F. M. S., Ziaziaris, W., & Munoz, L. (2014). Mitogen-activated protein kinase-activated protein kinase 2 in neuroinflammation, heat shock protein 27 phosphorylation, and cell cycle: role and targeting. *Molecular Pharmacology*, *85*(2), 345–356. https://doi.org/10.1124/mol.113.090365
Gwon, Y., Maxwell, B. A., Kolaitis, R.-M., Zhang, P., Kim, H. J., & Taylor, J. P. (2021). Ubiquitination of G3BP1 mediates stress granule disassembly in a context-specific manner. *Science*, *372*(6549),

eabf6548. https://doi.org/10.1126/science.abf6548
Györkei, Á., Daruka, L., Balogh, D., Őszi, E., Magyar, Z., Szappanos, B., Fekete, G., Fuxreiter, M., Horváth, P., Pál, C., Kintses, B., & Papp, B. (2022). Proteome-wide landscape of solubility limits in a bacterial cell. *Scientific Reports*, *12*(1), 6547. https://doi.org/10.1038/s41598-022-10427-1
Hatos, A., Monzon, A. M., Tosatto, S. C. E., Piovesan, D., & Fuxreiter, M. (2022). FuzDB: a new phase in understanding fuzzy interactions. *Nucleic Acids Research*, *50*(D1), D509–D517. https://doi.org/10.1093/nar/gkab1060
He, M., Zhou, X., & Wang, X. (2024). Glycosylation: mechanisms, biological functions and clinical implications. *Signal Transduction and Targeted Therapy*, *9*(1), 194. https://doi.org/10.1038/s41392-024-01886-1
Hofmann, S., Kedersha, N., Anderson, P., & Ivanov, P. (2021). Molecular mechanisms of stress granule assembly and disassembly. *Biochimica et Biophysica Acta, Molecular Cell Research*, *1868*(1), 118876. https://doi.org/10.1016/j.bbamcr.2020.118876
Hofweber, M., & Dormann, D. (2019). Friend or foe-Post-translational modifications as regulators of phase separation and RNP granule dynamics. *The Journal of Biological Chemistry*, *294*(18), 7137–7150. https://doi.org/10.1074/jbc.TM118.001189
Hogrebe, A., Hess, K. N., Llovet, A., Ramos, Y. J., Barente, A. S., Hernandez-Portugues, D., Smith, I. R., Rodríguez-Mias, R. A., & Villén, J. (2022). IsobaricQuant enables cross-platform quantification, visualization, and filtering of isobarically-labeled peptides. *Proteomics*, *22*(19-20), e2100253. https://doi.org/10.1002/pmic.202100253
Hughes, C. S., Moggridge, S., Müller, T., Sorensen, P. H., Morin, G. B., & Krijgsveld, J. (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nature Protocols*, *14*(1), 68–85. https://doi.org/10.1038/s41596-018-0082-x
Iqbal, S., & Hood, D. A. (2014). Oxidative stress-induced mitochondrial fragmentation and movement in skeletal muscle myoblasts. *American Journal of Physiology. Cell Physiology*, *306*(12), C1176–C1183. https://doi.org/10.1152/ajpcell.00017.2014
Ishihama, Y., Rappsilber, J., & Mann, M. (2006). Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics. *Journal of Proteome Research*, *5*(4), 988–994. https://doi.org/10.1021/pr050385q
Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2
Käll, L., Canterbury, J. D., Weston, J., Noble, W. S., & MacCoss, M. J. (2007). Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nature Methods*, *4*(11), 923–925. https://doi.org/10.1038/nmeth1113
Kedersha, N., Ivanov, P., & Anderson, P. (2013). Stress granules and cell signaling: more than just a passing phase? *Trends in Biochemical Sciences*, *38*(10), 494–506. https://doi.org/10.1016/j.tibs.2013.07.004
Kedersha, N., Panas, M. D., Achorn, C. A., Lyons, S., Tisdale, S., Hickman, T., Thomas, M., Lieberman, J., McInerney, G. M., Ivanov, P., & Anderson, P. (2016). G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. *The Journal of Cell Biology*, *212*(7), 845–860. https://doi.org/10.1083/jcb.201508028
Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M. J., Scheuner, D., Kaufman, R. J., Golan, D. E., & Anderson, P. (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *The Journal of Cell Biology*, *169*(6), 871–884. https://doi.org/10.1083/jcb.200502088
Khoury, G. A., Baliban, R. C., & Floudas, C. A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Scientific Reports*, *1*(1). https://doi.org/10.1038/srep00090

Krebs, E. G., Kent, A. B., & Fischer, E. H. (1958). The muscle phosphorylase b kinase reaction. *The Journal of Biological Chemistry*, *231*(1), 73–83. https://doi.org/10.1016/s0021-9258(19)77286-x

Krisenko, M. O., Higgins, R. L., Ghosh, S., Zhou, Q., Trybula, J. S., Wang, W.-H., & Geahlen, R. L. (2015). Syk Is Recruited to Stress Granules and Promotes Their Clearance through Autophagy. *The Journal of Biological Chemistry*, *290*(46), 27803–27815. https://doi.org/10.1074/jbc.M115.642900

Kroncke, B. M., Vanoye, C. G., Meiler, J., George, A. L., Jr, & Sanders, C. R. (2015). Personalized biochemistry and biophysics. *Biochemistry*, *54*(16), 2551–2559. https://doi.org/10.1021/acs.biochem.5b00189

Landry, J., Lambert, H., Zhou, M., Lavoie, J. N., Hickey, E., Weber, L. A., & Anderson, C. W. (1992). Human HSP27 is phosphorylated at serines 78 and 82 by heat shock and mitogen-activated kinases that recognize the same amino acid motif as S6 kinase II. *The Journal of Biological Chemistry*, *267*(2), 794–803. https://doi.org/10.1016/s0021-9258(18)48354-8

Lee, Y., Stiers, K. M., Kain, B. N., & Beamer, L. J. (2014). Compromised catalysis and potential folding defects in in vitro studies of missense mutants associated with hereditary phosphoglucomutase 1 deficiency. *The Journal of Biological Chemistry*, *289*(46), 32010–32019. https://doi.org/10.1074/jbc.M114.597914

Leutert, M., Barente, A. S., Fukuda, N. K., Rodriguez-Mias, R. A., & Villén, J. (2022). The regulatory landscape of the yeast phosphoproteome. In *bioRxiv* (p. 2022.10.23.513432). https://doi.org/10.1101/2022.10.23.513432

Leutert, M., Rodríguez-Mias, R. A., Fukuda, N. K., & Villén, J. (2019). R2-P2 rapid-robotic phosphoproteomics enables multidimensional cell signaling studies. *Molecular Systems Biology*, *15*(12), e9021. https://doi.org/10.15252/msb.20199021

Listwan, P., Terwilliger, T. C., & Waldo, G. S. (2009). Automated, high-throughput platform for protein solubility screening using a split-GFP system. *Journal of Structural and Functional Genomics*, *10*(1), 47–55. https://doi.org/10.1007/s10969-008-9049-4

Li, Y., Liang, R., Sun, M., Li, Z., Sheng, H., Wang, J., Xu, P., Liu, S., Yang, W., Lu, B., Zhang, S., & Shan, C. (2020). AMPK-dependent phosphorylation of HDAC8 triggers PGM1 expression to promote lung cancer cell survival under glucose starvation. *Cancer Letters*, *478*, 82–92. https://doi.org/10.1016/j.canlet.2020.03.007

Markmiller, S., Soltanieh, S., Server, K. L., Mak, R., Jin, W., Fang, M. Y., Luo, E.-C., Krach, F., Yang, D., Sen, A., Fulzele, A., Wozniak, J. M., Gonzalez, D. J., Kankel, M. W., Gao, F.-B., Bennett, E. J., Lécuyer, E., & Yeo, G. W. (2018). Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. *Cell*, *172*(3), 590–604.e13. https://doi.org/10.1016/j.cell.2017.12.032

Martinez Molina, D., Jafari, R., Ignatushchenko, M., Seki, T., Larsson, E. A., Dan, C., Sreekumar, L., Cao, Y., & Nordlund, P. (2013). Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science (New York, N.Y.)*, *341*(6141), 84–87. https://doi.org/10.1126/science.1233606

Matreyek, K. A., Starita, L. M., Stephany, J. J., Martin, B., Chiasson, M. A., Gray, V. E., Kircher, M., Khechaduri, A., Dines, J. N., Hause, R. J., Bhatia, S., Evans, W. E., Relling, M. V., Yang, W., Shendure, J., & Fowler, D. M. (2018). Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nature Genetics*, *50*(6), 874–882. https://doi.org/10.1038/s41588-018-0122-z

Matsuzaki, Y., Aoki, W., Miyazaki, T., Aburaya, S., Ohtani, Y., Kajiwara, K., Koike, N., Minakuchi, H., Miura, N., Kadonosono, T., & Ueda, M. (2021). Peptide barcoding for one-pot evaluation of sequence-function relationships of nanobodies. *Scientific Reports*, *11*(1), 21516. https://doi.org/10.1038/s41598-021-01019-6

Maxwell, B. A., Gwon, Y., Mishra, A., Peng, J., Nakamura, H., Zhang, K., Kim, H. J., & Taylor, J. P. (2021). Ubiquitination is essential for recovery of cellular activities after heat shock. *Science*, *372*(6549), eabc3593. https://doi.org/10.1126/science.abc3593

McAlister, G. C., Nusinow, D. P., Jedrychowski, M. P., Wühr, M., Huttlin, E. L., Erickson, B. K., Rad, R.,

Haas, W., & Gygi, S. P. (2014). MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Analytical Chemistry*, *86*(14), 7150–7158. https://doi.org/10.1021/ac502040v

Millar, S. R., Huang, J. Q., Schreiber, K. J., Tsai, Y.-C., Won, J., Zhang, J., Moses, A. M., & Youn, J.-Y. (2023). A New Phase of Networking: The Molecular Composition and Regulatory Dynamics of Mammalian Stress Granules. *Chemical Reviews*, *123*(14), 9036–9064. https://doi.org/10.1021/acs.chemrev.2c00608

Mladenov, E., Magin, S., Soni, A., & Iliakis, G. (2016). DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: Cell cycle and proliferation-dependent regulation. *Seminars in Cancer Biology*, *37-38*, 51–64. https://doi.org/10.1016/j.semcancer.2016.03.003

Motone, K., Kontogiorgos-Heintz, D., Wee, J., Kurihara, K., Yang, S., Roote, G., Fox, O. E., Fang, Y., Queen, M., Tolhurst, M., Cardozo, N., Jain, M., & Nivala, J. (2024). Multi-pass, single-molecule nanopore reading of long protein strands. *Nature*, *633*(8030), 662–669. https://doi.org/10.1038/s41586-024-07935-7

Needham, E. J., Parker, B. L., Burykin, T., James, D. E., & Humphrey, S. J. (2019). Illuminating the dark phosphoproteome. *Science Signaling*, *12*(565). https://doi.org/10.1126/scisignal.aau8645

Neuhofer, W. (2010). Role of NFAT5 in inflammatory disorders associated with osmotic stress. *Current Genomics*, *11*(8), 584–590. https://doi.org/10.2174/138920210793360961

Ochoa, D., Jarnuczak, A. F., Viéitez, C., Gehre, M., Soucheray, M., Mateus, A., Kleefeldt, A. A., Hill, A., Garcia-Alonso, L., Stein, F., Krogan, N. J., Savitski, M. M., Swaney, D. L., Vizcaíno, J. A., Noh, K.-M., & Beltrao, P. (2020). The functional landscape of the human phosphoproteome. *Nature Biotechnology*, *38*(3), 365–373. https://doi.org/10.1038/s41587-019-0344-3

Ong, S.-E., & Mann, M. (2007). Stable isotope labeling by amino acids in cell culture for quantitative proteomics. *Methods in Molecular Biology (Clifton, N.J.)*, *359*, 37–52. https://doi.org/10.1007/978-1-59745-255-7_3

Panas, M. D., Ivanov, P., & Anderson, P. (2016). Mechanistic insights into mammalian stress granule dynamics. *The Journal of Cell Biology*, *215*(3), 313–323. https://doi.org/10.1083/jcb.201609081

Pan, C., Knutson, S. D., Huth, S. W., & MacMillan, D. W. C. (2024). µMap proximity labeling in living cells reveals stress granule disassembly mechanisms. *Nature Chemical Biology*, 1–11. https://doi.org/10.1038/s41589-024-01721-2

Patel, J., McLeod, L. E., Vries, R. G. J., Flynn, A., Wang, X., & Proud, C. G. (2002). Cellular stresses profoundly inhibit protein synthesis and modulate the states of phosphorylation of multiple translation factors: Modulation of translation factors by cellular stress. *European Journal of Biochemistry*, *269*(12), 3076–3085. https://doi.org/10.1046/j.1432-1033.2002.02992.x

Pizzino, G., Irrera, N., Cucinotta, M., Pallio, G., Mannino, F., Arcoraci, V., Squadrito, F., Altavilla, D., & Bitto, A. (2017). Oxidative stress: Harms and benefits for human health. *Oxidative Medicine and Cellular Longevity*, *2017*, 8416763. https://doi.org/10.1155/2017/8416763

Potel, C. M., Kurzawa, N., Becher, I., Typas, A., Mateus, A., & Savitski, M. M. (2021). Impact of phosphorylation on thermal stability of proteins. *Nature Methods*, *18*(7), 757–759. https://doi.org/10.1038/s41592-021-01177-5

Potel, C. M., Lin, M.-H., Heck, A. J. R., & Lemeer, S. (2018). Widespread bacterial protein histidine phosphorylation revealed by mass spectrometry-based proteomics. *Nature Methods*, *15*(3), 187–190. https://doi.org/10.1038/nmeth.4580

Purba, E. R., Saita, E.-I., & Maruyama, I. N. (2017). Activation of the EGF receptor by ligand binding and oncogenic mutations: The “rotation model.” *Cells (Basel, Switzerland)*, *6*(2), 13. https://doi.org/10.3390/cells6020013

Quast, J.-P., Schuster, D., & Picotti, P. (2022). protti: an R package for comprehensive data analysis of peptide- and protein-centric bottom-up proteomics data. *Bioinformatics Advances*, *2*(1), vbab041. https://doi.org/10.1093/bioadv/vbab041

Ramazi, S., & Zahiri, J. (2021). Posttranslational modifications in proteins: resources, tools and prediction

methods. *Database: The Journal of Biological Databases and Curation*, *2021*, baab012. https://doi.org/10.1093/database/baab012

Reinehr, R., & Häussinger, D. (2006). Hyperosmotic activation of the CD95 death receptor system. *Acta Physiologica (Oxford, England)*, *187*(1-2), 199–203. https://doi.org/10.1111/j.1748-1716.2006.01541.x

Reinhardt, H. C., Aslanian, A. S., Lees, J. A., & Yaffe, M. B. (2007). p53-deficient cells rely on ATM- and ATR-mediated checkpoint signaling through the p38MAPK/MK2 pathway for survival after DNA damage. *Cancer Cell*, *11*(2), 175–189. https://doi.org/10.1016/j.ccr.2006.11.024

Riggs, C. L., Kedersha, N., Ivanov, P., & Anderson, P. (2020). Mammalian stress granules and P bodies at a glance. *Journal of Cell Science*, *133*(16). https://doi.org/10.1242/jcs.242487

Rodriguez-Mias, R. A., Hess, K. N., Ruiz, B. Y., Smith, I. R., Barente, A. S., Zimmerman, S. M., Lu, Y. Y., Noble, W. S., Fields, S., & Villén, J. (2022). Proteome-wide identification of amino acid substitutions deleterious for protein function. In *bioRxiv* (p. 2022.04.06.487405). https://doi.org/10.1101/2022.04.06.487405

Roux, K. J., Kim, D. I., Raida, M., & Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *The Journal of Cell Biology*, *196*(6), 801–810. https://doi.org/10.1083/jcb.201112098

Savitski, M. M., Reinhard, F. B. M., Franken, H., Werner, T., Savitski, M. F., Eberhard, D., Martinez Molina, D., Jafari, R., Dovega, R. B., Klaeger, S., Kuster, B., Nordlund, P., Bantscheff, M., & Drewes, G. (2014). Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science*, *346*(6205), 1255784. https://doi.org/10.1126/science.1255784

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Schreiber, K. J., Kadijk, E., & Youn, J.-Y. (2024). Exploring Options for Proximity-Dependent Biotinylation Experiments: Comparative Analysis of Labeling Enzymes and Affinity Purification Resins. *Journal of Proteome Research*. https://doi.org/10.1021/acs.jproteome.3c00908

Searle, B. C., Lawrence, R. T., MacCoss, M. J., & Villén, J. (2019). Thesaurus: quantifying phosphopeptide positional isomers. *Nature Methods*, *16*(8), 703–706. https://doi.org/10.1038/s41592-019-0498-4

Shimi, T., & Goldman, R. D. (2014). Nuclear lamins and oxidative stress in cell proliferation and longevity. *Advances in Experimental Medicine and Biology*, *773*, 415–430. https://doi.org/10.1007/978-1-4899-8032-8_19

Shteynberg, D. D., Deutsch, E. W., Campbell, D. S., Hoopmann, M. R., Kusebauch, U., Lee, D., Mendoza, L., Midha, M. K., Sun, Z., Whetton, A. D., & Moritz, R. L. (2019). PTMProphet: Fast and accurate mass modification localization for the Trans-Proteomic Pipeline. *Journal of Proteome Research*, *18*(12), 4262–4272. https://doi.org/10.1021/acs.jproteome.9b00205

Smith, I. R. (2022). *Developing Proteomic Methods to Assay Function of Proteoforms* (J. Villén (ed.)) [University of Washington]. https://www.proquest.com/pqdtglobal/docview/2652706672/3F1915EB3ED64176PQ/1?accountid=14784

Southern, K. W., Castellani, C., Lammertyn, E., Smyth, A., VanDevanter, D., van Koningsbruggen-Rietschel, S., Barben, J., Bevan, A., Brokaar, E., Collins, S., Connett, G. J., Daniels, T. W. V., Davies, J., Declercq, D., Gartner, S., Gramegna, A., Hamilton, N., Hauser, J., Kashirskaya, N., … Duff, A. (2023). Standards of care for CFTR variant-specific therapy (including modulators) for people with cystic fibrosis. *Journal of Cystic Fibrosis: Official Journal of the European Cystic Fibrosis Society*, *22*(1), 17–30. https://doi.org/10.1016/j.jcf.2022.10.002

Sridharan, S., Hernandez-Armendariz, A., Kurzawa, N., Potel, C. M., Memon, D., Beltrao, P., Bantscheff, M., Huber, W., Cuylen-Haering, S., & Savitski, M. M. (2022). Systematic discovery of biomolecular condensate-specific protein phosphorylation. *Nature Chemical Biology*, *18*(10), 1104–1114.

https://doi.org/10.1038/s41589-022-01062-y
Sridharan, S., Kurzawa, N., Werner, T., Günthner, I., Helm, D., Huber, W., Bantscheff, M., & Savitski, M. M. (2019). Proteome-wide solubility and thermal stability profiling reveals distinct regulatory roles for ATP. *Nature Communications*, *10*(1), 1155. https://doi.org/10.1038/s41467-019-09107-y
Starita, L. M., Ahituv, N., Dunham, M. J., Kitzman, J. O., Roth, F. P., Seelig, G., Shendure, J., & Fowler, D. M. (2017). Variant Interpretation: Functional Assays to the Rescue. *American Journal of Human Genetics*, *101*(3), 315–325. https://doi.org/10.1016/j.ajhg.2017.07.014
Starita, L. M., Young, D. L., Islam, M., Kitzman, J. O., Gullingsrud, J., Hause, R. J., Fowler, D. M., Parvin, J. D., Shendure, J., & Fields, S. (2015). Massively parallel functional analysis of BRCA1 RING domain variants. *Genetics*, *200*(2), 413–422. https://doi.org/10.1534/genetics.115.175802
Sun, B., Liu, Z., Liu, J., Zhao, S., Wang, L., & Wang, F. (2022). The utility of proteases in proteomics, from sequence profiling to structure and function analysis. *Proteomics*, e2200132. https://doi.org/10.1002/pmic.202200132
Swaney, D. L., Wenger, C. D., & Coon, J. J. (2010). Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *Journal of Proteome Research*, *9*(3), 1323–1329. https://doi.org/10.1021/pr900863u
Tape, C. J., Worboys, J. D., Sinclair, J., Gourlay, R., Vogt, J., McMahon, K. M., Trost, M., Lauffenburger, D. A., Lamont, D. J., & Jørgensen, C. (2014). Reproducible automated phosphopeptide enrichment using magnetic TiO2 and Ti-IMAC. *Analytical Chemistry*, *86*(20), 10296–10302. https://doi.org/10.1021/ac5025842
Tsai, C.-F., Hsu, C.-C., Hung, J.-N., Wang, Y.-T., Choong, W.-K., Zeng, M.-Y., Lin, P.-Y., Hong, R.-W., Sung, T.-Y., & Chen, Y.-J. (2014). Sequential phosphoproteomic enrichment through complementary metal-directed immobilized metal ion affinity chromatography. *Analytical Chemistry*, *86*(1), 685–693. https://doi.org/10.1021/ac4031175
Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., & Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, *13*(9), 731–740. https://doi.org/10.1038/nmeth.3901
Vertii, A., Hakim, C., Kotlyarov, A., & Gaestel, M. (2006). Analysis of properties of small heat shock protein Hsp25 in MAPK-activated protein kinase 2 (MK2)-deficient cells: MK2-dependent insolubilization of Hsp25 oligomers correlates with susceptibility to stress: MK2-DEPENDENT INSOLUBILIZATION OF Hsp25 OLIGOMERS CORRELATES WITH SUSCEPTIBILITY TO STRESS. *The Journal of Biological Chemistry*, *281*(37), 26966–26975. https://doi.org/10.1074/jbc.M602134200
Wagih, O., Galardini, M., Busby, B. P., Memon, D., Typas, A., & Beltrao, P. (2018). A resource of variant effect predictions of single nucleotide variants in model organisms. *Molecular Systems Biology*, *14*(12), e8430. https://doi.org/10.15252/msb.20188430
Wee, P., & Wang, Z. (2017). Epidermal growth factor receptor cell proliferation signaling pathways. *Cancers*, *9*(5), 52. https://doi.org/10.3390/cancers9050052
Weidenauer, L., & Quadroni, M. (2021). Phosphorylation in the charged linker modulates interactions and secretion of Hsp90β. *Cells (Basel, Switzerland)*, *10*(7), 1701. https://doi.org/10.3390/cells10071701
Wippich, F., Bodenmiller, B., Trajkovska, M. G., Wanka, S., Aebersold, R., & Pelkmans, L. (2013). Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. *Cell*, *152*(4), 791–805. https://doi.org/10.1016/j.cell.2013.01.033
Wiredja, D. D., Koyutürk, M., & Chance, M. R. (2017). The KSEA App: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics* , *33*(21), 3489–3491. https://doi.org/10.1093/bioinformatics/btx415
Wong, S. Y.-W., Beamer, L. J., Gadomski, T., Honzik, T., Mohamed, M., Wortmann, S. B., Brocke Holmefjord, K. S., Mork, M., Bowling, F., Sykut-Cegielska, J., Koch, D., Ackermann, A., Stanley, C. A., Rymen, D., Zeharia, A., Al-Sayed, M., Marquardt, T., Jaeken, J., Lefeber, D., … Morava, E. (2016). Defining the Phenotype and Assessing Severity in Phosphoglucomutase-1 Deficiency. *The Journal of Pediatrics*, *175*, 130–136.e8. https://doi.org/10.1016/j.jpeds.2016.04.021

Xu, W., Doshi, A., Lei, M., Eck, M. J., & Harrison, S. C. (1999). Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Molecular Cell*, *3*(5), 629–638. https://doi.org/10.1016/s1097-2765(00)80356-1

Yamazaki, H., Takagi, M., Kosako, H., Hirano, T., & Yoshimura, S. H. (2022). Cell cycle-specific phase separation regulated by protein charge blockiness. *Nature Cell Biology*, *24*(5), 625–632. https://doi.org/10.1038/s41556-022-00903-1

Yarden, Y., & Schlessinger, J. (1987). Self-phosphorylation of epidermal growth factor receptor: evidence for a model of intermolecular allosteric activation. *Biochemistry*, *26*(5), 1434–1442. https://doi.org/10.1021/bi00379a034

Young, I. S., & Woodside, J. V. (2001). Antioxidants in health and disease. *Journal of Clinical Pathology*, *54*(3), 176–186. https://doi.org/10.1136/jcp.54.3.176

Youn, J.-Y., Dunham, W. H., Hong, S. J., Knight, J. D. R., Bashkurov, M., Chen, G. I., Bagci, H., Rathod, B., MacLeod, G., Eng, S. W. M., Angers, S., Morris, Q., Fabian, M., Côté, J.-F., & Gingras, A.-C. (2018). High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. *Molecular Cell*, *69*(3), 517–532.e11. https://doi.org/10.1016/j.molcel.2017.12.020

Youn, J.-Y., Dyakov, B. J. A., Zhang, J., Knight, J. D. R., Vernon, R. M., Forman-Kay, J. D., & Gingras, A.-C. (2019). Properties of Stress Granule and P-Body Proteomes. *Molecular Cell*, *76*(2), 286–294. https://doi.org/10.1016/j.molcel.2019.09.014

Yu, Q., Liu, X., Keller, M. P., Navarrete-Perea, J., Zhang, T., Fu, S., Vaites, L. P., Shuken, S. R., Schmid, E., Keele, G. R., Li, J., Huttlin, E. L., Rashan, E. H., Simcox, J., Churchill, G. A., Schweppe, D. K., Attie, A. D., Paulo, J. A., & Gygi, S. P. (2023). Sample multiplexing-based targeted pathway proteomics with real-time analytics reveals the impact of genetic variation on protein expression. *Nature Communications*, *14*(1), 555. https://doi.org/10.1038/s41467-023-36269-7

Zhou, H., Ye, M., Dong, J., Corradini, E., Cristobal, A., Heck, A. J. R., Zou, H., & Mohammed, S. (2013). Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nature Protocols*, *8*(3), 461–480. https://doi.org/10.1038/nprot.2013.010

Zimmerman, S. M., Kon, Y., Hauke, A. C., Ruiz, B. Y., Fields, S., & Phizicky, E. M. (2018). Conditional accumulation of toxic tRNAs to cause amino acid misincorporation. *Nucleic Acids Research*, *46*(15), 7831–7843. https://doi.org/10.1093/nar/gky623

Zuniga, N. R., Frost, D. C., Kuhn, K., Shin, M., Whitehouse, R. L., Wei, T.-Y., He, Y., Dawson, S. L., Pike, I., Bomgarden, R. D., Gygi, S. P., & Paulo, J. A. (2024). Achieving a 35-plex tandem mass tag reagent set through deuterium incorporation. *Journal of Proteome Research*, *23*(11), 5153–5165. https://doi.org/10.1021/acs.jproteome.4c00668