THE UNIVERSITY OF CHICAGO

STRESS -ASSOCIATED INTRACELLULAR ACIDIFICATION IS ADAPTIVE

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES

AND

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BIOPHYSICAL SCIENCES

BY

CATHERINE TRIANDAFILLOU

CHICAGO, ILLINOIS

AUGUST 2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I first want to acknowledge the tireless support and enthusiasm of my advisors Aaron Dinner and Allan Drummond. They have been generous with their time and invested in my growth as a scientist and as a person. Thanks to both of you for being on my team, and I look forward to having both of you as part of my lifelong support system.

I particularly want to thank a few former members of the Drummond Lab: Edward Wallace for formative discussions and infinite suggestions for excellent papers to read, and Josh Riback for setting the bar incredibly high. I also want to acknowledge the amazing current group (Haneul Yoo, Jared Bard, Samantha Keyport, Hendrik Glauninger, Caitlin Wong Hickernell, and Rosalind Pan). I frequently say how important your engagement and enthusiasm are – I would work directly with and trust without question each and every one of you; you are amazing colleagues and friends.

I also want to acknowledge several members (current and former) of the Dinner group: Bodhi Vani, Adam Antoszewski, Elizabeth White, Steven Redford, Zhiyue Lu, and Lu Hong for keeping me sane; Alan Hutchison for convincing me to join the group and always brightening my day; and Monika Scholz for being my role model in pretty much all things. Glen Hocky and Simon Freedman both provided both technical help and really wonderful conversations. Martin Falk was incredibly patient in teaching me, and is a joy to work with.

It's been such an honor to be part of the Biophysical Sciences Graduate program. It's a really special community of scientists, and provides students with so many opportunities for leadership and growth. I've met some of my favorite people and closest friends within and through this group. The success of the program is due in large part to the hard work and creativity of the leadership – Michele Wittels and Julie Feder in particular have been so encouraging, helpful, and supportive, and I can't thank them enough for their investment in me.

The friends and colleagues I met at the Marine Biological Laboratory in Woods Hole have changed my life for the better in so many ways – the atmosphere and connections created there have nurtured my curiosity like nothing else.

And finally, to my incredibly supportive parents and sister, and to all the amazing friends I've made in Chicago who have become my family – thank you for always being there for me. I never could have done this without you.

ABSTRACT

Exposure to high temperatures induces many changes in cells, including the induction of a transcriptional program regulated by the highly-conserved transcription factor Heat Shock Factor 1 (Hsf1). In a wide range of eukaryotes, stress also triggers transient intracellular acidification which, by unknown mechanisms, is associated with increased survival. Activation of the Hsf1-mediated heat shock response can be triggered by misfolding of newly synthesized polypeptides, and so has been thought to depend on ongoing protein synthesis. I have discovered that even in the absence of ongoing translation, heat-shock-associated cytosolic acidification specifically promotes the activation of Hsf1 in budding yeast. Additionally, preventing cells from transiently acidifying during heat shock compromises fitness, demonstrating that intracellular acidification is adaptive. These results imply the existence of non-canonical triggers for the response and additional contributions to cellular fitness beyond the repair of stress-induced damage.

CHAPTER 1
INTRODUCTION

It is remarkable that individual eukaryotic cells, microscopic agglomerates of water, salts, lipid, protein, and nucleic acid, can accomplish something that all complex, multicellular eukaryotes can: sense and respond to changes in the environment. Cells do so without obvious sensory or central processing systems, and yet they are extremely effective, weathering changes in environment that would decimate some higher organisms. How do cells accomplish this task, and importantly – what exactly are they responding to?

An ideal model system for studying the cellular response to environmental change is the heat shock response (HSR). [71, 82, 130] The HSR consists of an array of evolutionarily-conserved cellular changes that are enacted in a coordinated fashion when cells encounter temperatures (generally 5-30°C) above their optimal growth temperature. [71] The HSR includes changes in which genes are transcribed and at what levels: metabolic proteins continue to be produced at high levels, as do molecular chaperones; [17, 16] a specific set of stress-responsive molecular chaperones, largely under control of the "master regulator" of the heat shock response, the transcription factor Hsf1 are strongly upregulated; [71, 96] ribosomal and cell-cycle-related proteins stop being produced; [45, 17] and nuclear organization changes dramatically. [20, 21] The effects extend beyond the nucleus: cellular growth is halted [6, 107, 124] (depending on the temperature either temporarily or indefinitely), metabolism slows, the ATP/ADP ratio decreases, [135] translation is reduced, [3, 18] and the type and volume of secondary metabolites produced shifts dramatically. [50, 112] In the last decade or so, an array of *physical* changes within cells have also come to light (all of the following observations were made using yeast cells as a model): diffusion slows and the entire cytoplasm shifts from being a viscous liquid to a gel-like solid, [84, 61] the solubility of the entire proteome shifts as proteins and mRNAs assemble into dense intracellular aggregates, [131] and ionic gradients across cell membranes are reduced as the plasma membrane becomes more permeable and the interior of cells acidifies to match the exterior as a result. [91, 84]

What, if any, of these changes are cells reacting to when they encounter elevated temperature, and how are they able to sense them? At its heart, a change in temperature is a physical signal, and yet how this physical signal is interpreted by cells remains an open question. Classical models of the heat shock response hold that increases in temperature cause *damage* to cells, and that the cleanup of this damage is the primary goal of the heat

shock response. [71, 127, 75] In many ways this simple explanation is intuitively satisfying: it is well-known that temperature influences protein folding and stability, [35] and induction of the heat shock response has been shown to protect cells from conditions that would otherwise be lethal. [70, 96] However, the response has been studied for decades and not all evidence that has been generated supports the conclusion that the purpose of the heat shock response is to repair damage incurred by stress. These findings include the following observations:

1. Cells activate the HSR during other stresses where protein misfolding is not a clear outcome of the stressor, including starvation, [144, 138] viral infection, [87] and during lifestyle changes such as the switch that dimorphic fungi make to pathogenic forms. [13, 88]

2. The heat shock response has a sharp temperature onset, [71] implying that whatever sets off the response does so in a roughly trigger-like manner – general misfolding of proteins does not seem to be sufficient to explain the shape of the onset of the response. Furthermore, this onset point is highly divergent in different species, discussed in the following point.

3. Cells, even closely-related strains of yeast, can have very different temperature niches and onset temperatures for the HSR – in order to evolve or adapt to these new niches, a non-specific mechanism such as protein misfolding would require the unfolding temperature of the proteome as a whole to be shifted. While we acknowledge that a few particularly sensitive species may trigger the response and that the shifting of their thermal stability might change temperature niches, we take pains to note that this is not in line with classical misfolding models, and indeed under some interpretations these proteins themselves would in some way be 'sensors' of heat shock, implying that this is an evolved and conserved role of the protein, not a general malfunction to which cells are reacting.

4. It has recently been shown that the entire proteome reorganizes during heat stress [131, 65] and that many proteins assemble reversibly into protein-and-RNA-rich deposits that can be detected biochemically and via microscopy. Although the exact correlation between these bodies remains indistinct, the two species share some components and

it is largely accepted that these granules form reversibly and without damage to the components. [131, 65, 100]

While it is clear that protein folding and stability are a central part of the heat shock response – indeed, elements of the HSR are activated when gross misfolding within cells is triggered chemically [75] – the above observations led me to investigate, from a different perspective, how the physical environment of the cell leads to activation of the HSR. Specifically I focused on the phenomenon of intracellular acidification, [135, 24] a conserved [91, 12, 36, 139, 60, 123, 84, 102] physical change that co-occurs with heat stress. The goal was to understand whether this change, which has strong implications for both the folding and structure of individual proteins [63, 28, 69] and the physical properties of the cytoplasm as a whole, [84, 61, 95] was part of the adaptive response to heat shock. Although this sounds like a complicated question to answer, at its heart is a simple test – elements of cellular behaviors that are adaptive promote fitness, therefore by altering whether or not and to what degree cells acidify during stress and measuring the resulting changes in cellular fitness, we can determine whether heat shock-associated intracellular acidification is adaptive, a passive and neutral consequence of heat shock, or even an additional stressor.

The idea that the HSR could act beyond damage cleanup is not revolutionary – in a review penned over 30 years ago, Lindquist noted that HSPs were induced during development in *Drosophila* and other organisms. [71] Temperature can also act as a developmental cue for single-celled organisms – *Candida albicans*, a pathogenic yeast, uses temperature changes to sense when cells have entered a host and must proceed with genetic and cellular changes that promote infection and evasion of the host immune system; these changes are dependent on Hsf1 activity. [13, 88] There is also evidence that the regulon of Hsf1 changes and adapts to promote invasive growth and malignancy of cancerous tissue. [26, 76] Thus viewing the heat shock response not as damage-cleanup, but as a transition to a new growth program in response to environmental change has substantial backing from the literature, and motivates our search for the triggers of the response beyond misfolded proteins.

I addressed this question in two distinct ways: first, I directly examined the role of acidification during the heat shock response by manipulating intracellular pH during heat shock and assaying the induction of molecular chaperones, a hallmark of the HSR and of Hsf1 activity specifically. I found that when translation (the production of proteins) is inhibited during heat shock, stress-associated acidification promotes both activation of the HSR and

cellular fitness. This finding stands in contrast to classic models of the induction of the heat shock response that predict that ongoing translation is required for HSR activation. I link this acidification to the activity of Hsf1, highly conserved in eukaryotes, which has a specific mechanism of activation and has been implicated in human disease. These results demonstrate that stress-associated acidification is adaptive and promotes cellular fitness, and suggest that cells use cues beyond protein misfolding to sense changes in the environment and specifically elevated temperature.

Second, I delved further into the effect of pH on molecular structure by examining charge patterning, which pH is known to alter, [28, 69] in proteins. I performed simulations of intrinsically disordered proteins and determined whether existing measures of charge patterning developed for polyelectrolytes are appropriate for real biological sequences. I demonstrated that a region of a highly-conserved fungal protein shows signatures of selection on charge patterning, and link this to measures of the physical properties of the region. Finally, I integrated the molecular simulations with a toy model of evolution to demonstrate that for highly-charged regions, selection for a phyiscal property can yield apparent selection on charge patterning, further supporting the evidence of conservation of this trait in extant species. These results could help clarify the role that charge patterning and pH-induced changes in local charge can contribute to the physical structure and self-association of proteins involved in the heat shock response and beyond.

CHAPTER 2
TRANSIENT INTRACELLULAR ACIDIFICATION REGULATES THE CORE TRANSCRIPTIONAL HEAT SHOCK RESPONSE

2.1 Abstract

Heat shock induces a conserved transcriptional program regulated by heat shock factor 1 (Hsf1) in eukaryotic cells. Activation of this heat-shock response is triggered by heat-induced misfolding of newly synthesized polypeptides, and so has been thought to depend on ongoing protein synthesis. Here, using the budding yeast *Saccharomyces cerevisiae*, we report the discovery that Hsf1 can be robustly activated when protein synthesis is inhibited, so long as cells undergo cytosolic acidification. Heat shock has long been known to cause transient intracellular acidification which, for reasons which have remained unclear, is associated with increased stress resistance in eukaryotes. We demonstrate that acidification is required for heat shock response induction in translationally inhibited cells, and specifically affects Hsf1 activation. Physiological heat-triggered acidification also increases population fitness and promotes cell cycle reentry following heat shock. Our results uncover a previously unknown adaptive dimension of the well-studied eukaryotic heat shock response.

2.2 Introduction

To survive and thrive, organisms must rapidly respond when their environments turn harsh. Cells across the tree of life possess the capacity to adaptively respond to primordial stresses—heat, starvation, hypoxia, exposure to noxious compounds—in a conserved program involving the production of so-called heat shock proteins, many of which act as molecular chaperones [71]. Transcription of heat shock proteins surges at the onset of stress, reaching as much as a thousand fold during thermal stress, with more modest induction accompanying nutrient withdrawal and diverse other stresses [71, 144, 82, 48]. In eukaryotes, the transcriptional stress response is controlled by multiple factors, with the heat shock transcription factor Hsf1 regulating induction of a core group of chaperones [116, 98]. Basal levels of chaperones repress Hsf1 by direct binding [113, 142, 64], and removal of this repression in the absence of stress suffices to activate transcription [142, 64]. Induced chaperones, in turn, assist with

protein folding, as well as preventing and dispersing stress-induced molecular aggregates [127, 104, 18, 132, 65].

The mechanism by which the Hsf1-mediated transcriptional response is induced following physiological heat shock is incomplete and has remained so since the response's discovery nearly 60 years ago [105]. In the currently accepted model for heat-triggered Hsf1 activation, events proceed as follows. Hsf1 is constitutively bound and repressed by the molecular chaperone Hsp70 before stress [64, 142]. Heat stress is thought to cause deleterious protein unfolding [104] (misfolding) which exposes hydrophobic regions [127] for which Hsp70 has high affinity [108]. Titration of Hsp70 away from Hsf1 suffices to induce Hsf1 [142, 64]. Despite the crucial role misfolded proteins play in this model, no specific endogenous eukaryotic protein has been demonstrated to misfold in vivo in response to a sublethal heat shock. Instead, newly synthesized polypeptides (which include nascent chains still being synthesized and complete polypeptides which have yet to reach their native structure) are thought to serve as Hsf1 inducers during heat stress [5, 121, 68]. Suppression of newly synthesized polypeptides by translation inhibition suppresses the heat-induced transcription of genes regulated by Hsf1 [5, 7, 2, 75]. Consequently, ongoing translation has been deemed a requirement for Hsf1 activation [75].

Notably, the same diverse environmental changes which stimulate the transcriptional response are also accompanied by intracellular acidification—a drop in cytosolic pH [135, 12, 84, 65]. Like the transcriptional response, stress-induced acidification is broadly conserved in eukaryotes, including mammals [12, 139, 123, 32], insects [36, 143], plants [60], and fungi [135, 65]. Although acidification has sometimes been viewed as a toxic consequence of stress, particularly in studies of hypoxia and ischemia-associated acidosis [60, 123], the cytoprotective effects of short-term acidification were identified decades ago [123]. Recent work has shown that interfering with energy-depletion-induced acidification in budding yeast and in fission yeast, which diverged from budding yeast more than half a billion years ago [54], compromises the fitness of both species [84, 61]. Furthermore, many mature proteins associated with stress-induced condensation show a strong dependence on pH for their self-association, whether by polymerization or phase separation [97, 65, 84, 103, 43].

What role does stress-induced cellular acidification play in the transcriptional response to heat shock? Early work in *Drosophila melanogaster* produced mixed results: one study indicated that acidification had little impact on the production of heat shock proteins [36], while later work showed that Hsf1 trimerization, a key activation step, could be induced

by acidification in vitro [143]. Acidification during stress influences cell signaling [30, 51] and appears to be cytoprotective [84, 61, 24, 94]. To what extent this adaptive effect of pH depends on the core transcriptional stress response remains unknown. Furthermore, it has been shown that cell cycle reentry after heat shock necessarily follows the dissolution of stress granules, and that this dissolution depends on molecular chaperones [65]. These data suggest a clear link between transcription and growth. Exactly how do intracellular acidification, transcriptional induction, chaperone production, and cellular growth interrelate following heat shock?

To answer this question, we developed a single-cell system to both monitor and manipulate cytosolic pH while tracking the induction of molecular chaperones in budding yeast. We find that acidification universally promotes the heat shock response, and that when canonical triggers for the response—the newly synthesized polypeptides—are suppressed, acidification is required for cells to respond to heat shock. Acidification alone, however, is insufficient to induce a response. We measure fitness on both the population and single-cell level and find that in both cases, the physiological stress-associated drop in pH promotes fitness.

Finally, global measurement of transcript levels as a function of intracellular pH during heat shock reveals specific suppression of core Hsf1 target genes when intracellular acidification is prevented. All of our results are consistent with a previous hypothesis positing a role for temperature- and pH-dependent phase separation in sensing heat stress [103], leading us to predict a specific mechanism for induction of the heat-shock transcriptional response in which elevated pH suppresses a temperature-sensitive phase separation process. Our results link cytosolic acidification to the regulation of the canonical transcriptional heat shock response and subsequent stress adaptation in single cells, indicating that pH regulation plays a central role in the Hsf1-mediated stress response.

2.3 Results

A high-throughput assay allows quantification of pH-dependent, single-cell responses to heat shock

Yeast thrive in acidic environments, and spend a significant quantity of cellular resources on the activity of membrane-associated proton pumps which keep the cytoplasm at a resting pH of around 7.5 [91]. The resulting electrochemical gradient is used to drive transport and

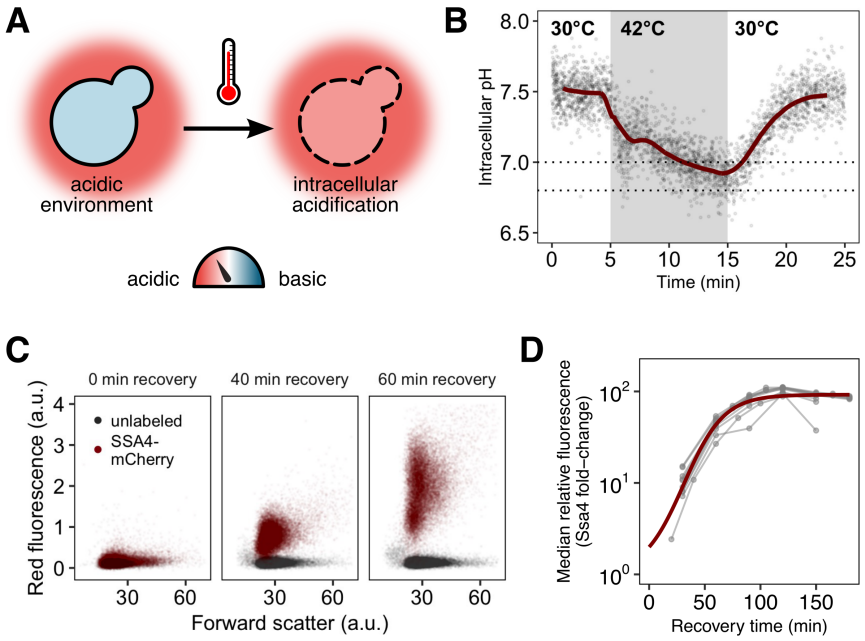

Figure 2.1: Yeast cells respond to heat shock with intracellular pH changes and production of heat-shock proteins that can be tracked at the single-cell level.
(A) *S. cerevisiae* cells live in acidic environments but maintain intracellular pH at neutral or slightly basic. During heat stress the cell membrane becomes more permeable, leading to intracellular acidification. **(B)** Intracellular pH changes during stress measured with continuous flow cytometry; each point is an individual cell. The gray region is the period during which cells were exposed to elevated temperature. A solid line shows a sliding-window average over all data; for visual clarity, only 2% of points are shown. Dashed lines represent the range we subsequently use as representative of the physiological pH drop. **(C)** Induction of labeled Hsp70 (Ssa4-mCherry) after heat shock. Each plot is a timepoint during recovery from 42°C, 20-minute heat shock showing forward scatter pulse area, which correlates roughly with size, versus red fluorescence. Unlabeled cells are shown in black for comparison. **(D)** Summary of induction of Ssa4-mCherry after heat shock; each point represents the fold change, relative to unstressed cells, of the median fluorescence of > 5000 cells expressed as a ratio to forward scatter; each gray line is an experiment ($n = 6$). Thick red curve is a sigmoid fit (see Methods).

other crucial cellular processes, but is disrupted during stress, causing cells to acidify (Figure 2.1A). While the mechanism of proton influx remains somewhat murky, elevated temperature increases membrane permeability [25] and other stresses have been shown to reduce proton pump activity [91, 93, 30]. We first sought to precisely measure the intracellular pH changes associated with heat stress.

To track intracellular pH during stress and recovery, we engineered yeast cells to constitutively express pHluorin, a pH-sensitive green fluorescent protein derivative used to measure intracellular pH, in the cytoplasm [78]. The probe was calibrated to known pH values in vivo (Figure 2.8 and Methods). We used this strain to characterize intracellular pH changes occurring during heat stress and recovery. After a 42° C, 10-minute heat stress in acidic

media (pH 4) we find that cells exposed to elevated temperature rapidly and reproducibly acidify from a resting pH of approximately 7.5 to a range of slightly acidic pH values: 6.8 to 7.0 (Figure 2.1B, Figure 2.8C), in agreement with previous results [135]. When returned to normal growth temperature (30°C), cells restore the resting pH in approximately ten minutes. The minimum pH reached is similar for cells stressed at 42° C for 20 minutes (Figure 2.8D).

The hallmark of the heat shock response is the production of molecular chaperones [71, 127, 82]. To assess the effects of acidification on this response, we measured chaperone induction by engineering a pHluorin-labeled yeast strain to express a red-fluorescent-protein-tagged version of Ssa4 (Ssa4-mCherry) from its endogenous locus. Ssa4 is a strongly temperature-responsive Hsp70 chaperone, and its encoding gene is a specific Hsf1 target [58, 98, 82]. This two-color reporter strain allowed us to simultaneously track intracellular pH and the stress response at the single-cell level.

We heat shocked cells at 42°C for 20 minutes and then returned them to 30° C to recover. Samples were collected at 15- to 30-minute intervals during recovery and analyzed by flow cytometry to monitor Ssa4-mCherry production. An example of the raw data, showing an increase in fluorescence in the mCherry channel over time, is shown in Figure 2.1C. Although the maturation time of the fluorophore, mCherry, confounds determination of the absolute timing of the response, this delay is shared across experiments, allowing for direct comparison between conditions and replicates. For each independent experiment, we tracked the median relative change in red fluorescence over time, creating induction curves which characterize the response, as in Figure 2.1D.

Intracellular acidification during heat shock promotes rapid heat-shock protein production

With the tools in hand to quantify intracellular pH and induction of stress proteins, we set out to first determine whether the ability to acidify during stress affected the cellular response. Evidence from the literature [91] strongly suggests that acidification results primarily from influx of environmental protons, rather than (for example) the release of protons from internal stores such as the vacuole. We confirmed a dependence on external protons by heat-stressing cells in normal, acidic media (pH 4), or in media where the pH had been adjusted to the cellular resting pH (7.5). Stressing cells in non-acidic media prevented their ability to acidify

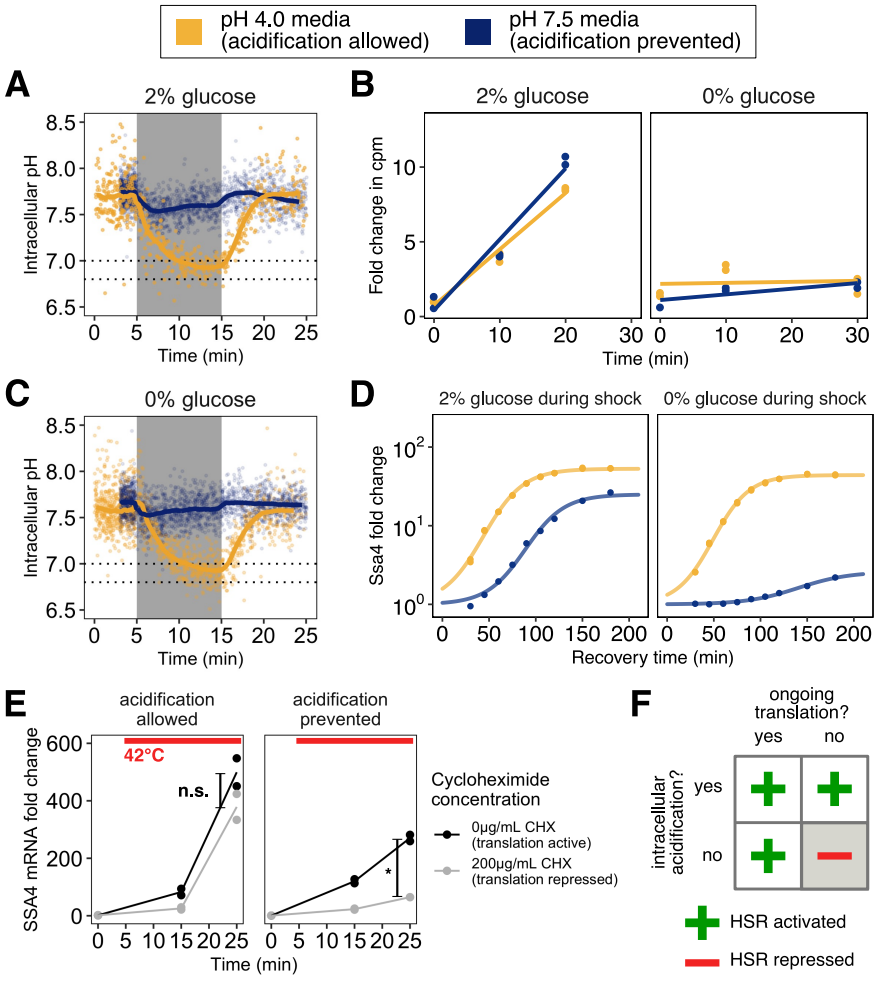

Figure 2.2: Preventing stress-associated acidification delays or impairs the heat shock response when cells are translationally inactive

(A) Intracellular pH change as a function of the pH of the environment. Cells stressed in acidic media (pH 4.0) acidify (yellow), while cells stressed in media at the resting pH (7.5) do not (blue). **(B)** Induction of Ssa4 in cells able (yellow) or unable (blue) to acidify during heat shock. Under normal exponential growth conditions (left, high glucose) failure to acidify delays and reduces the response to elevated temperature. However, if global translation is suppressed with brief glucose withdrawal during heat shock (right), acidification is required for induction of the response. All measurements are of cells recovering in media containing 2% glucose. **(C)** Same as **A**, but in media without glucose. The same trend is observed; an acidic exterior leads to acidification, but cells in media at the resting pH do not acidify. **(D)** Inhibition of translation by glucose withdrawal does not depend on environmental pH. Incorporation of radiolabeled amino acids into total cellular proteins as a function of time after a switch from medium at pH 4 with 2 % glucose to the indicated media. **(E)** Acidification promotes the transcriptional heat shock response (production of SSA4 mRNA) when cells are treated with the translation inhibitor cycloheximide prior to heat stress. *, $P = 0.011$; n.s., $P > 0.1$, Welch two-sample t-test. **(F)** Acidification promotes the heat shock response, and is required when cells are translationally inactive.

(Figure 2.2A). Cells that could not acidify during stress delayed and reduced the induction of Ssa4 (Figure 2.2B, left hand side).

Misfolding of newly synthesized polypeptides is thought to provide the primary trigger for Hsf1 activation, as described in the Introduction. To test whether acidification still promoted the stress response even under conditions where the concentration of newly synthesized polypeptides would be sharply limited, we used brief glucose withdrawal, which is known to rapidly and reversibly inhibit translation of most cellular mRNAs [3]. We heat-stressed cells, then returned them to favorable growth conditions (2% glucose, 30° C) to recover. Strikingly, we found that even in the absence of translation and presumably misfolded newly synthesized polypeptides, cells that could acidify during stress responded almost identically to cells stressed while global translation was unperturbed. However, cells that were not actively translating and also were unable to acidify during stress almost completely failed to respond (Figure 2.2B, right hand side).

We confirmed that this sharp dependence of heat-shock protein production on intracellular pH was not due to differences in intracellular acidification with and without translation (Figure 2.2C, compare to A), due to differences in fluorophore maturation (Figure 2.13A), and that translational suppression was not dependent on the pH of the media (Figure 2.2D). To ensure that the differences we saw were due to translation state and not nutrient withdrawal, we performed the same set of experiments with cells grown in maltose, a sugar that does not cause translational suppression when acutely withdrawn for short periods of time [3]. Cells stressed in the presence of maltose were able to respond to heat shock regardless of whether they were able to acidify. Crucially, cells that experienced brief maltose withdrawal prior to stress, but remained actively translating (Figure 2.9A), were also able to produce SSA4 during recovery regardless of whether they acidified (Figure 2.9B). This finding demonstrates that the translational state of the cell does affect the heat shock response, but that even in the absence of ongoing translation cells can still respond to heat shock, provided that they experience intracellular acidification.

We further verified these results by inhibiting translation with cycloheximide, an elongation inhibitor, followed by heat stress where acidification was either allowed or prevented. Because we could not track production of Ssa4 protein after this treatment, we measured induction of the *SSA4* transcript using qPCR. We found that when cells are able to acidify, they robustly respond to heat shock regardless of whether they can translate, but that if acidification is prevented, translational repression sharply inhibits the transcriptional response

(Figure 2.2E).

From these data we conclude that rapid, robust chaperone expression following heat shock depends either on ongoing translation, as previous studies have found, or on intracellular acidification (Figure 2.2F), an effect which has not been reported before. We therefore set out to further characterize the translation-independent, pH-dependent arm of the stress response, beginning with the quantitative effects of pH on the induction of heat-shock proteins.

Failure to acidify during heat shock impairs the core transcriptional stress response regulated by Hsf1

Our results thus far link pH regulation to the translation of a limited number of heat shock proteins (Figure 2.13A). Since the heat shock response is characterized by conserved changes in transcription of multiple regulons, we used mRNA-Seq to characterize the stress response with and without acidification and under various translation conditions, using the pH of the media to prevent or allow acidification and treatment with cycloheximide or acute glucose withdrawal to prevent translation. We also assayed cells where the equivalent treatments were performed without heat shock in order to account for transcriptional changes due to changes in the pH of the media and translation. Although the results for translation arrest with both cycloheximide and glucose withdrawal are often similar, the glucose withdrawal results are more complicated, and for clarity only the cycloheximide data are shown in this figure. Equivalent versions of all figures for glucose withdrawal are shown in Figure 2.10 and 2.11, and instances where the two differ substantially are noted in the text.

The transcriptome-wide response to elevated temperature as a function of acidification and translation state are shown in Figure 2.3A (mean of two biological replicates, see Figure 2.10 for correlation between replicates and data quality). Cells stressed while synthesizing proteins responded by strongly upregulating heat shock proteins, regardless of whether or not they acidified during stress (Figure 2.3A, upper panels). However, cells which were heat shocked after translation was blocked showed a decrease in the induction of heat shock proteins when acidification was also blocked (Figure 2.3A, lower right hand plot).

In order to further examine this effect, we plotted the abundance of the heat-shock transcriptomes with and without acidification against one another for both translating and non-translating populations (Figure 2.3B). Genes that lie above the diagonal are preferentially induced in cells that acidify, while those that lie on the diagonal are equally well regardless

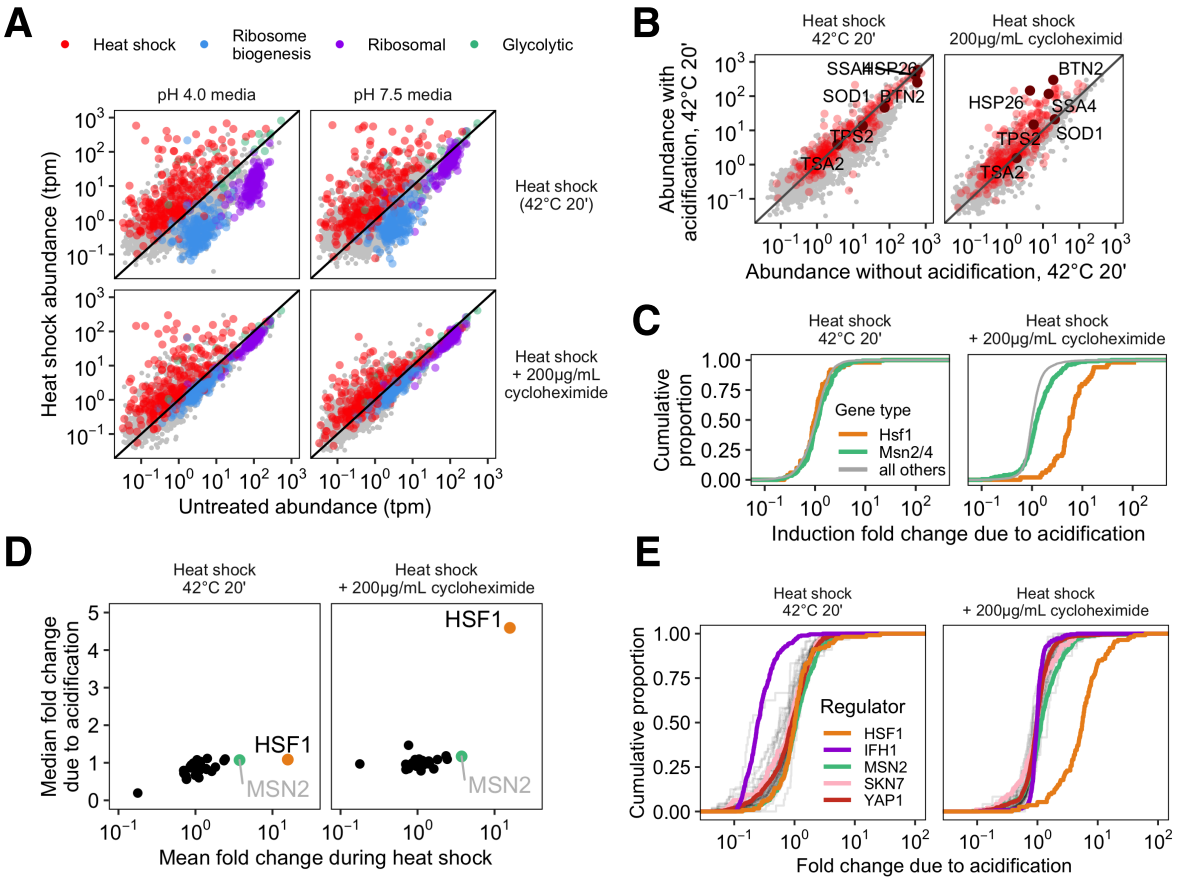

Figure 2.3: Failure to acidify during heat shock specifically represses Hsf1-activated genes **(A)** Transcript abundance (transcripts per million, tpm) in stressed versus unstressed samples. Colors correspond to gene type, gray points fall into neither category. **(B)** Cumulative distribution function (CDF) of per-gene transcript fold change to unstressed for each stress condition. Colors correspond to the transcription factor responsible for the induction of the gene; see Methods for details on categorization. **(C)** CDF of per-gene transcript abundance in cells heat shocked at pH 6.8 relative to cells shocked at pH 7.4 (induction ratio). The red line shows all heat shock proteins; this group is further divided into genes regulated by Msn2/4 (green) which show similar behavior to all detected transcripts (gray; $P = .402$, Wilcoxon rank sum test), and those regulated by Hsf1 (orange), which are significantly higher in acidified cells ($P < 0.01$, Wilcoxon rank sum test). **(D)** Induction ratio (determined by qPCR) of reporter genes for Hsf1 and Msn2/4 in cells shocked in media (without ionophore) where the external pH dictated whether or not cells could acidify during heat shock (see Figure 2.2C).

of pH. The vast majority of the heat shock genes lie along the diagonal in cells stress while translation is ongoing (Figure 2.3B, left hand side), and while this is still true for many heat shock genes during heat shock in the absence of translation, we noticed that a subset of heat shock genes, including SSA4, BTN2, and HSP26 were particularly pH sensitive. To further differentiate between heat shock proteins and characterize the pH-sensitivity of the response, we examined the transcription factors responsible for the induction of different heat shock proteins.

Although many transcription factors are involved in stress responses, three are thought to play a major role in the heat shock response: Hsf1, which regulates chaperone-centric stress responses in all eukaryotes [82, 76], and Msn2/4, a pair of paralogous factors limited to fungi [45, 39, 89]. Recent work has used multiple methods to clearly define the regulons of both [116, 98], and the genes for which induction is specifically due to one or the other. Both regulons are upregulated when cells are stressed without translation, however, the Hsf1 genes are much less induced than expected when cells are stressed without concomitant intracellular acidification (Figure 2.11A). To quantify the sensitivity to acidification we calculated the abundance after heat shock with acidification over abundance after heat shock without acidification. The distribution of values for all genes are shown in Figure 2.3C. Distributions with more values greater than one (more genes preferentially induced when cells acidify) will lie further to the right on such a plot. The pH sensitivity of the Hsf1 regulon is apparent when cells are translationally inhibited (Figure 2.3C).

Is the Hsf1 regulon the only regulon that is preferentially induced when cells acidify during heat shock? To address this question, we calculated the pH sensitivity of the regulons of all transcription factors for which data are available under heat shock conditions, using the YeTFaSCo database [29] to generate a list of proteins that have been shown to bind genomic DNA, and genes annotated to be regulated by each in the YeastMine database. Hsf1 was a clear outlier, showing strong pH sensitivity when translation is repressed (Figure 2.3D). This was also true when translation was modulated using glucose withdrawal, although in this case the cellular response was much more complicated, as evidenced by the induction of heat shock genes even in the absence of heat shock (Figure 2.11D, E). Under these conditions Hsf1 is still preferentially active in cells that acidify in a way that cannot be explained by the mock treatment (Figure 2.11E). From these data we conclude that while acidification affects the Hsf1 regulon specifically during heat shock, acidification, which is known to occur during other stresses, could affect the activity of other transcription factors as well during

other physiologically relevant stresses.

Finally, we widened our search to include other transcription factors which may change their regulation during heat shock, but have only been annotated under non-stress conditions; the full distributions for each regulon are shown in Figure 2.3E. When translation is arrested during heat shock, Hsf1 is the only transcription factor that shows significant pH sensitivity. Interestingly, another transcription factor, Ifh1, was pH-sensitive only under conditions where cells were translationally active (Figure 2.3E, left hand side). The repression of Ifh1 is known to be dependent on the synthesis of ribosomal proteins; [2] we have now demonstrated that it also depends on intracellular acidification, perhaps due to an effect on TOR or Ifh1 nuclear body formation. In line with this result, Ifh1 shows no pH-dependence during cycloheximide treatment; since translation is arrested there is no neosynthesis of RPs, and thus Ifh1 will not be deactivated.

From these results, we conclude that intracellular acidification differentially affects the regulons of several transcription factors. Most strikingly, when cellular translation is halted, conditions under which classical models predict no Hsf1 activation, we find that intracellular acidification promotes induction of genes specifically under control of Hsf1. This highlights a previously unknown facet of the regulation of this important TF; we consider potential mechanisms for this pH-sensitivity in the Discussion.

Manipulating intracellular pH during heat shock reveals the precise relationship between pH and the stress response

To determine the quantitative relationship between intracellular pH during heat shock and chaperone production during recovery, we sought a means to manipulate intracellular pH which would circumvent cellular regulation of the proton gradient. To accomplish this, we chemically manipulated intracellular pH using an ionophore, nigericin, modifying a published protocol [128]. Ionophores allow ions to penetrate cell membranes, temporarily destroying the electrochemical gradient. Nigericin is a K^+ / H^+ antiporter [44] which has been used in a variety of biological systems to equilibrate intracellular and extracellular pH [79, 85, 122, 22]. We verified that by placing cells in buffers at different pHs and treating with ionophore we were able to accurately manipulate intracellular pH, and that this control did not depend on temperature (Figure 2.4B). Manipulating intracellular pH independently of temperature allowed us to determine that acidification alone was not sufficient to produce a stress response

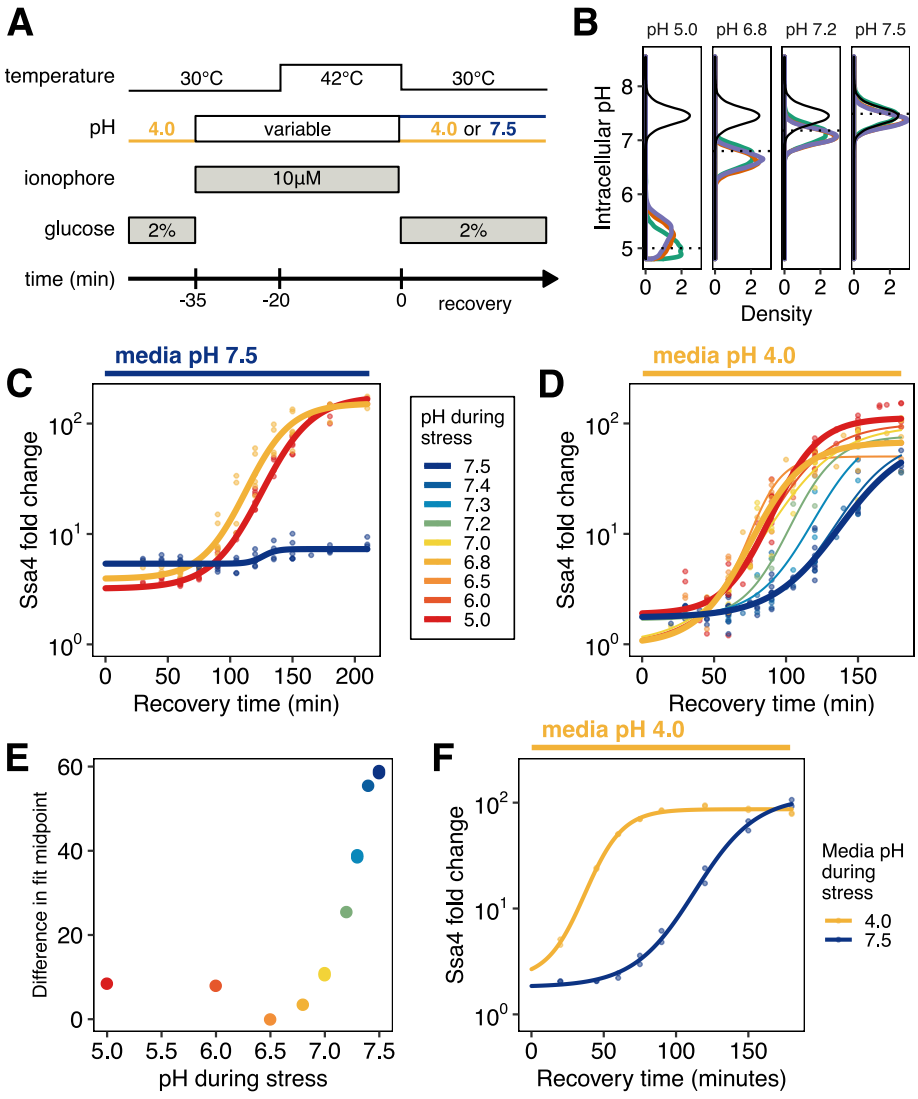

Figure 2.4: Quantitative control of intracellular pH reveals that, in the absence of translation, acidification is required for Ssa4 induction

(A) Schematic of intracellular pH manipulation experiments. **(B)** Intracellular pH is accurately manipulated during stress. Intracellular pH distributions were measured to determine the efficacy of pH manipulation before (green), during (red), and after (purple) 42° C heat stress. Dashed lines indicate buffer pH, and the black distribution shows unmanipulated cells for comparison. **(C)** Manipulation of intracellular pH with ionophore reproduces the acidification-dependent induction of Ssa4. Compare to **2.2B**, right hand side. **(D)** Fold change in Ssa4 expression following stress at different intracellular pHs and recovery in acidic media. Points represent the median of individual measurements; at least three biological replicates were performed for each condition (see Methods). Lines are sigmoid fits (see Methods for fitting details). **(E)** pH-dependence of the induction delay; points are the midpoint of the sigmoidal fits in **D**. **(F)** Dependence of the stress response on media pH, followed by recovery in acidic media, recapitulates the pH-dependence of the stress response when ionophore treatment is used; compare to **D**.

(Figure 2.13D, right hand side), with the exception of the lowest pH examined, pH 5.0, which is substantially below the range of physiologically realized pH values during short-term heat shock (cf. Figure 2.1B). We also verified that ionophore treatment did not have long-term fitness consequences by measuring the relative growth rate of treated and untreated cells (Figure 2.13C). Finally, we performed mRNA-seq on cells stressed in the presence of ionophore at both the stress-associated pH (6.8) and resting pH (7.4) (Figure 2.1B) and observed the same acidification-dependence of the Hsf1 regulon under these conditions (Figure 2.12).

Exposing ionophore-treated cells to heat shock (42° C for 20 minutes; Figure 2.4A) at a range of buffer-controlled pH levels permitted us to monitor the effect of intracellular pH on the subsequent response. After heat shock with pH control, we returned cells to ionophore-free media at 30° C and monitored Ssa4 induction by flow cytometry. Treatment with buffer and ionophore delayed the chaperone production in all samples relative to untreated cells, but did so consistently and did not affect the ultimate induction level (Figure 2.13B), supporting the assumption that pH-dependent differences between treatments can be appropriately interpreted.

The range of pH values investigated, from 7.5 to 5.0, reflected three main pH regimes. Cells held at or near pH 7.5, their resting pH, experienced little or no acidification during stress. Cells moved to pH 6.8 to 7.0 experienced an approximately physiological level of acidification (cf. Fig. 2.1B). Cells adjusted to below pH 6.8 during stress experienced a larger than normal pH change.

Using the ionophore to manipulate intracellular pH, we were able to reproduce the same phenotype we observed in cells stressed in media with and without the ability to acidify—populations stressed at an acidic intracellular pH were able to respond, and those stressed at the resting pH were not (Figure 2.4C). Furthermore, we found that additional acidification during stress, as low as pH 5, did not increase or decrease the response compared to physiological acidification.

Our initial experiments involved allowing cells to recover in media buffered to the resting pH, ensuring that the differences in the stress response were due to pH during stress. However, we noticed that proton availability after stress seemed to influence the response. Remarkably, when we pH-treated cells and allowed them to recover in acidic media, these cells were able to induce Ssa4, where cells recovering in buffered medium were not (Figure 2.4D, compare to Figure 2.4C). This recovery-media-pH-dependent induction occurred with a pH-dependent delay (Figure 2.4E) which was maximal when cells did not experience acid-

ification during stress. To ensure that this was not due to treatment with the ionophore, we performed the same experiment without ionophore, stressing cells in media that was acidic or at the resting pH, and recovering in acidic media. The same pattern of induction was observed: cells recovering in acidic media induced Ssa4, but with a substantial delay (Figure 2.4F). What could explain the dependence on media pH during recovery for induction of the stress response? One possibility is that acidification occurs after stress and enables the induction of the stress response; we test this proposal in the following section.

We draw several conclusions from these data. The physiologically observed acidification of the cytosol is necessary for rapid heat shock protein production when translation is repressed. Physiological levels of acidification alone do not activate the response. Depriving translationally inactive cells of the opportunity to acidify virtually silences chaperone production after heat shock, an effect which is mostly transcriptional. Cells offered the chance to acidify after heat shock are still capable of mounting a response albeit with a substantial delay. All this suggests that intracellular pH during recovery plays a significant role in the production of heat shock proteins, and we turned our attention to that possibility.

Reversal of stress-induced acidification during recovery promotes heat shock protein production in single cells

How does intracellular pH during recovery influence heat shock protein production? In acidic media, without pH manipulation, intracellular pH rapidly returns to pre-stress (resting) levels after return to ambient temperature [84] (cf. Figure 2.1A). We therefore wondered whether this intracellular pH recovery depended on the pH experienced during stress, and if it affected the response to heat shock. We examined intracellular pH restoration in cell populations heat shocked at different ionophore-enforced pHs and allowed to recover in acidic media. Populations stressed under acidic conditions rapidly restored intracellular pH during recovery (Figure 2.5A and 2.14A). In contrast, cells stressed at pH values above 7.0 took longer on average to restore intracellular pH to resting levels, and in some cases failed to do so even after two hours (Figure 2.5A). This effect was not due to ionophore treatment; when we examined cells stressed in acidic media versus media at the resting pH, we observed the same pattern (Figure 2.5B).

These results confirm the hypothesis proposed in the previous section: cell populations held at the pre-stress pH during stress acidified during recovery. These populations—which

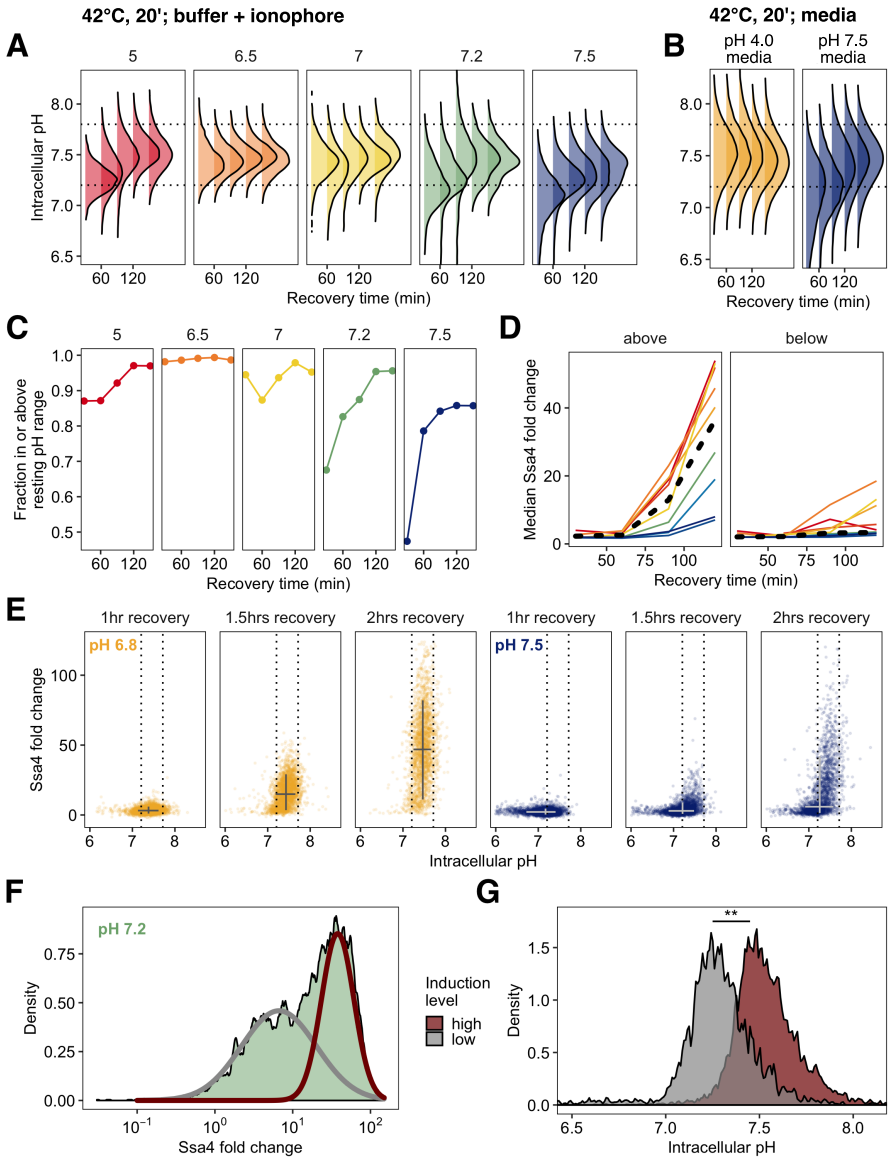

Figure 2.5: Preventing acidification during heat shock dysregulates the return to resting pH during recovery, suppressing chaperone production

(A) Intracellular pH changes as a function of environmental pH. Cells stressed in acidic media (pH 4.0, yellow) acidify, whereas cells stressed in media at the resting pH (7.5, blue) do not. **(B)** Inhibition of translation by glucose withdrawal does not depend on environmental pH. Incorporation of radiolabeled amino acids into total cellular proteins in counts per minute (cpm) as a function of time after a switch from medium at pH 4 with 2% glucose to the indicated media. **(C)** Same as **A**, but in glucose-free medium. **(D)** Induction of Ssa4 in cells able (yellow) or unable (blue) to acidify during heat shock. All measurements are of cells recovering in media containing 2% glucose. **(E)** Acidification promotes the transcriptional heat shock response (production of *SSA4* mRNA) when cells are treated with the translation inhibitor cycloheximide (200 μg/mL) prior to heat stress. *, $p = 0.017$; n.s., $p = 0.11$, Welch two-sample t-test. **(F)** Acidification promotes the heat shock response, and is required when cells are translationally inactive.

also showed pH-dependent delays in heat shock protein production—consistently had a larger proportion of cells outside the resting pH range (2.5C). We noted that on average, cells that had failed to return to the resting pH range also failed to induce Ssa4 (Figure 2.5D). This led us to investigate the connections between intracellular pH recovery and chaperone production on the single cell level.

Examination of the relationship between intracellular pH variation and production of Ssa4 in single cells revealed a clear pattern: virtually all cells that produced high levels of Ssa4 had returned to the resting pH (Figure 2.5E), and cells which did not return to the resting pH showed low levels of Ssa4 for up to three hours (Figure 2.14B). The vast majority of cells which had restored the resting pH after 105 minutes of recovery went on to robustly induce Ssa4 (Figure 2.14B). Cells far from the observed pre-stress resting pH induced less protein.

We further noticed that some populations showed a bimodal distribution of Ssa4 induction. In particular, we observed this behavior in populations stressed between pH 7.5 and pH 7.0. Figure 2.5F shows this distribution for cells stressed at pH 7.2; all distributions are shown in Figure 2.14C. The existence of subpopulations within identically treated samples which show different Ssa4 induction created a natural experiment, permitting us to test a strong version of the hypothesis that pH recovery is required for chaperone induction. We predicted that cells showing lower Ssa4 expression would have a lower intracellular pH compared to those with higher expression.

To test this prediction, we assigned cells to low- and high-expression categories by fitting the data with a mixture of two Gaussian distributions [8] at each timepoint (Figure 2.5F). We found that the lower-expressing subpopulation had a distinctly acid-shifted intracellular pH compared to the high-expressing cells (Figure 2.5G), confirming our prediction. Particularly at 120 minutes of recovery, when we see strong bimodality (Figure 2.14C), we also see strong separation of the intracellular pH distributions, with the low-expressing cells displaying intracellular pH values that fall below the ordinary unstressed range.

These data demonstrate that although cells require acidification during stress to mount a rapid response, the response further depends on subsequent reversal of acidification. The return to the resting pH dictates the dynamics of chaperone production. Acidification, either simultaneous with or following heat stress, followed by return to the resting pH is required for robust induction of chaperones after heat stress.

Precisely tuned stress-associated acidification increases cellular fitness during recovery from heat shock

In light of the connections we have established between intracellular pH changes and the induction of heat shock proteins, we sought to determine whether these pH changes promoted fitness during recovery from heat stress.

In single-celled organisms such as *S. cerevisiae*, fitness differences can be quantified by measuring the instantaneous growth rate relative to a competitor. This growth rate difference can be accurately measured by quantifying the slope of the logarithm of the ratio of population sizes as a function of time [46]. The difference in instantaneous growth rate, also known as the selection coefficient, quantifies how much better (positive) or worse (negative) cells grow relative to this reference competitor. Growth differences from two strains can then be directly compared to assess growth differences between strains, independent of the reference.

To measure fitness differences due to acidification during stress, we heat-shocked pHluorin/ Ssa4-mCherry dual-labeled cells in the presence of ionophore with a range of extracellular pH levels, enforcing a range of intracellular pH values as before. We then mixed these cultures with exponentially growing wild-type cells as the competitive reference and monitored relative proportions of these populations during recovery (Figure 2.6A). We performed additional controls to correct for potential strain differences and for the fitness effect of ionophore (Methods and Figure 2.15).

Population growth rate during recovery depended strongly on intracellular pH during heat shock. As expected, all heat-shocked populations grew more slowly than the unshocked control, with a minimum growth rate defect of -0.0043/min (Figure 2.6B), equivalent to a nearly four-fold increase in instantaneous doubling time. Maximum fitness was achieved by populations with intracellular pH enforced to be close to its observed physiological stress-induced levels ($\sim$6.8) in unmanipulated cells (cf. Figure 2.1). We observed the largest fitness defects in populations with pH levels set at pre-stress levels ($\sim$7.5). Ionophore treatment alone had only a minor, pH-independent effect on cell growth (Figure 2.6B and Figure 2.13C). We conclude that mimicking physiological intracellular acidification during stress maximizes fitness during recovery, again consistent with acidification playing an adaptive role in the heat shock response.

We have established that differences in intracellular pH during heat shock cause differ-

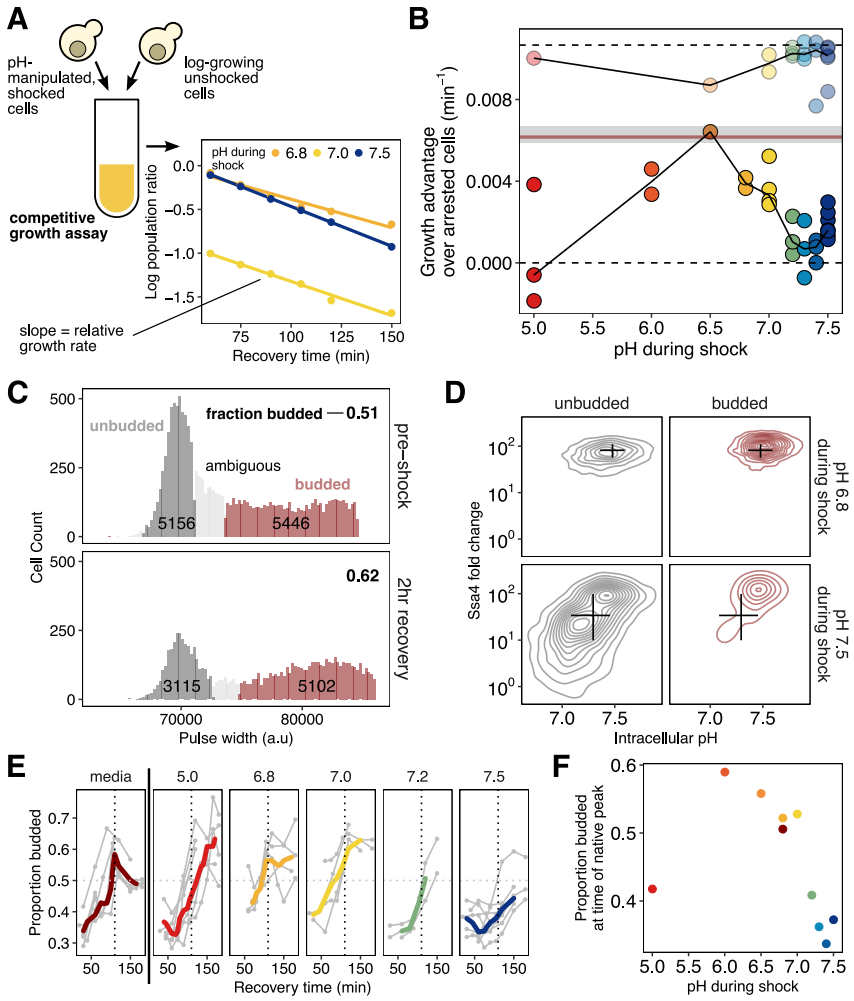

Figure 2.6: Intracellular acidification during heat shock promotes increased fitness during recovery on the population and single-cell levels

(A) Schematic of the competitive growth assay which measures population fitness. Heat-shocked Ssa4-mCherry/pHluorin cells and exponentially-growing unlabeled cells are mixed and allowed to grow at 30° C. Fitness is measured by fitting the log-ratio of the population sizes as a function of time to a line; the slope of the line is the difference in exponential growth rates (see Methods). **(B)** Intracellular pH during heat shock vs. relative growth rate expressed as the difference from the theoretical minimum for completely arrested cells. Each point is an independent experiment; opaque points are heat-shocked populations, transparent are control populations treated with ionophore at room temperature. Gray bar is the equivalent fitness loss for cells shocked without pH manipulation. See Methods and 2.15F for details and all fits. **(C)** Classification of cells: large/budded (red) and small/unbudded (dark gray). Classification was performed by fitting the forward-scatter pulse width to a two-component Gaussian mixture model and using a 90% confidence cutoff to classify cells into each category; cells that did not meet this criterion (shown in light gray) are not included in the analysis. Numeric labels show the number of cells in each category. **(D)** Ssa4 fold-change versus intracellular pH for budded and unbudded cells during recovery at three hours post-shock. Black lines are summary statistics of the entire population (budded and unbudded) and span the middle 50% of the data, crossing at the median. **(E)** Proportion of cells budded as a function of time during recovery. The characteristic shape of the curve derived from cells stressed without pH manipulation is shown in the left-most panel. The proportion budded peaks at approximately two hours of recovery (vertical dashed line). **(F)** Summary of **E**, showing the average proportion of cells budded between 90 and 120 minutes after heat shock.

ences in heat-shock protein expression at the population and single-cell levels, and that these same pH differences also cause differences in population fitness. To what extent do differences in expression cause observed differences in fitness? A causal model is motivated by the repeated observation that chaperone expression and subsequent dispersal of stress-induced aggregation precedes resumption of normal translation and progression through the cell cycle [18, 65].

A causal, directed link from pH to chaperone expression to growth predicts that 1) cell-to-cell variation in pH will predict cell-to-cell variation in both chaperone expression and growth, and 2) cells which have resumed progression through the cell cycle will have high levels of chaperones. To test these predictions, we monitored variation in cell cycle progression and chaperone expression as a function of intracellular pH in single cells within identically treated populations.

Cellular growth and division reflect progression through the cell cycle. In budding yeast, this progress can be tracked morphologically, because emergence of a bud signals that cells have exited the gap 1 phase (G_1) and have passed through the START cell-cycle checkpoint [53]. Heat stress causes yeast cells to arrest in G_1, an unbudded state [107]. The presence of a bud following stress indicates that the cell has re-entered the cell cycle and begun reproducing. Cells without a bud cannot be confidently assigned to a growth phase because they may either be arrested in G_1 or merely passing through this phase during active growth. To classify budded cells, we adapted a microscopy-based assay [107] to work with flow cytometry, using the pulse width of the forward scatter channel to measure cell (or cell plus bud) length. From the resulting data we scored cells as budded, unbudded, or ambiguous (Figure 2.6C, [126, 57, 56]; see Methods and Figure 2.15 for full details). Within these morphological categories, we then could assess the relationships between cell-cycle state, intracellular pH, and chaperone expression.

Cells heat-shocked at pH 6.8, mimicking normal acidification, showed robust chaperone expression during recovery. In contrast, cells shocked at pH 7.5, preventing acidification, created a large subpopulation of cells in which chaperone expression was suppressed (Figure 2.6D). Virtually all cells which could be confidently assigned to the budded state showed high chaperone expression, and nearly all cells showing low chaperone expression were found in the unbudded state. Low-expression, unbudded cells also showed near-uniform reduction in pH, consistent with the dysregulation of pH observed at the population level. These observations match both above predictions of a causal relationship between chaperone expression and cell-

cycle resumption, modulated by intracellular pH.

We also used the proportion of budded and unbudded cells as an orthogonal measure of population-level fitness, measuring resumption of the cell cycle as opposed to cell doubling. During recovery, cells are released from heat-induced G_1 arrest en masse, leading to a temporary synchronization of the population with a coordinated increase in the proportion of budded cells, which eventually returns to the steady-state value for exponentially growing cells [107] as seen in Figure 2.6E (left, dark red trace). After a 42° C, 20-minute heat shock without pH manipulation, the percentage of budded cells peaked just before two hours of recovery (dashed line in Figure 2.6E). In pH-manipulated cells, if the pH experienced with elevated temperature was close to the native stress-associated pH, this recovery peak occurred at approximately the same time as in unmanipulated cells. However, cells that experienced a more acidic or more basic pH during heat shock showed a delay in the occurrence of the budding peak (Figure 2.6E, summarized in 2.6F), in agreement with the difference in growth rates shown in Figure 2.6B.

By measuring growth in multiple ways, we have shown that post-stress resumption of growth is tuned to particular stress-associated cytoplasmic pH values. Moreover, fitness positively correlates both with increased chaperone production and with restoration of the pre-stress pH in populations and in individual cells. Resumption of growth is consistent, at the population and single-cell level, with induced chaperones contributing to release of stress-induced cell-cycle arrest, as others have observed [65].

2.4 Discussion

What is the physiological significance of the broadly conserved, transient intracellular acidification triggered by stress in eukaryotes? By decoupling changes in intracellular pH from heat shock in budding yeast, we have discovered that the canonical transcriptional stress response mediated by heat shock factor 1 (Hsf1) depends on cellular acidification. When cells are translationally suppressed, such as following glucose withdrawal, transient acidification becomes a requirement for achieving a robust transcriptional response. Even in translationally active cells, acidification promotes induction. Restoration of resting pH and chaperone protein expression increase competitive fitness by promoting reentry into the cell cycle and overall population growth rates, indicating that transient acidification is an adaptive component of the heat shock response.

Our initial results are consistent with the longstanding view that misfolding of newly synthesized polypeptides can serve as Hsf1 inducers [5, 75], presumably through recruitment of Hsp70 away from its repressive association with Hsf1 [142, 64, 68]. However, we have discovered an alternative pathway for Hsf1 activation under conditions when newly synthesized proteins are in short supply—when translational activity is low, such as following starvation or pharmacological inhibition. Here, intracellular pH plays a decisive causal role in Hsf1 activation following heat shock. Either ongoing translation or intracellular acidification is required, and the absence of either signal leads to suppression of the Hsf1-mediated transcriptional response during heat shock (Figure 2.7).

What is the source of the protons required for adaptive acidification? Our results strongly indicate that extracellular protons entering the cell following heat shock are the dominant cause of acidification. Simply placing translationally inactive cells in medium buffered to the resting cellular pH is sufficient to suppress the heat shock response during an otherwise robust heat shock, suggesting that no intracellular store of protons is liberated to cause acidification. Membrane permeability to small molecules increases with temperature in *S. cerevisiae* [25], and proton permeability specifically has been shown to increase with temperature [129], providing a likely mechanism for temperature-dependent acidification when a plasma-membrane-spanning proton gradient is present. (We discuss below certain physiological scenarios in which an ample source of extracellular protons and heat shock will reliably co-occur.)

Our results indicate a close causal connection between intracellular pH, chaperone production, and cellular growth. A surprising yet consistent detail is that cells must restore their resting pH before producing high levels of molecular chaperones. Previous work has demonstrated that heat shock causes changes in intracellular pH [1] and that intracellular pH controls growth rate [93]. Our results are consistent with these findings, while adding critical steps, such as demonstrating that chaperone production sits between pH and growth in the causal chain, and that these dynamics can be seen at the single-cell level.

The pH dependence of Hsf1 points to pH-dependent substrates of Hsf1's repressor, Hsp70. Why is acidification required to mount the transcriptional heat shock response under certain conditions? Acidification in the absence of heat shock, at least to pH levels which would normally follow heat shock, is insufficient to activate Hsf1, ruling out direct sensing of pH by Hsf1 or its repressor Hsp70. On the flip side, Hsf1 can be robustly

activated without a drop in pH, so long as cells are translationally active, indicating that acidification is not necessary for Hsf1 activation. Recent key studies have demonstrated that production of Hsp70 binding substrates that titrate Hsp70 away from Hsf1 suffices to induce Hsf1 in the absence of heat shock [142, 64]. All these results are consistent with the standard misfolding model: newly synthesized polypeptides misfold in response to heat shock, leading to recruitment of Hsp70, which causes Hsf1 activation. The pressing question is how Hsf1 is activated in the absence of newly synthesized polypeptides. Previous results argue against widespread heat-induced misfolding of mature endogenous proteins in vivo [131]. Here, consistent with those results, cells show marked repression of the heat shock response at 42° C when translation is attenuated and resting pH is maintained, suggesting that misfolding caused by temperature, if it occurs, is insufficient to trigger the Hsf1 response. The remaining possibility is that Hsp70 substrates can be produced without ongoing translation in an acidification-dependent manner.

That is, we seek a mature protein which, in a heat- *and* pH-dependent manner, changes its state in a way which recruits Hsp70. Remarkably, a profusion of such candidates exists. Nearly two hundred mature endogenous proteins in yeast form reversible condensates in response to heat shock which under extreme stress coalesce into stress granules [131]. Hsp70 colocalizes with stress granules [18, 132], and stress-granule dispersal depends on Hsp70 and other chaperones [18, 132, 65, 66]. Two individual protein components of stress granules, poly(A)-binding protein Pab1 and poly(U)-binding protein Pub1, condense by phase separation in vitro when exposed to heat in a pH-dependent manner [103, 66]. Pab1 condenses by phase separation which depends strongly on temperature and pH: a 42° C heat shock is insufficient to cause Pab1 condensation at pH 7.5, but sufficient at pH 6.6 [103]. In the case of Pub1, acidic pH promotes phase separation at non-shock temperatures, but these condensates spontaneously disperse when the resting pH is restored; only heat-triggered Pub1 condensates require dispersal by molecular chaperones.

Together, these studies indicate the existence of multiple proteins that undergo heat-triggered, pH-dependent condensation processes, producing assemblies which conditionally recruit Hsp70.

pH dependence constrains the search for temperature sensors in eukaryotes. How eukaryotic cells sense temperature remains unknown [140]. In the misfolding model for Hsf1 activation, misfolded proteins are the actors which convert an increase in temperature

into Hsp70 recruitment and thereby activate Hsf1 [82]; neither Hsf1 nor Hsp70 has temperature sensitivity in this model. Recent work has suggested that human Hsf1 possesses intrinsic thermal sensing ability regulating its trimerization [55]. While this is a tantalizing possibility, Hsf1 thermosensing has yet to be demonstrated in vivo or for yeast, whose Hsf1 is thought to be constitutively trimerized [82]. Our results also demonstrate that temperature alone is insufficient to activate Hsf1's response; a drop in intracellular pH is required.

We have previously proposed that heat-triggered protein condensation can take the place of misfolding-induced aggregation in the standard model for Hsf1 activation, with phase-separating proteins acting as the primary sensors of temperature [103]. Phase separation and other phase-transition behaviors provide a compelling solution to the tricky problem of sensing temperature, which typically involves only a few degrees' change: 30° C to 37° C for robust induction of yeast's heat shock response [45]. Phase transitions by definition are highly cooperative, amplifying tiny changes in individual molecules into massive system-level transformations [140]. Crucially, unlike misfolding of newly synthesized polypeptides, heat-shock-triggered condensation of mature proteins is not suppressed by translational inhibition [131].

Which proteins might serve at the front line of temperature sensing, transducing slight temperature shifts into a cellular signal capable of triggering the Hsf1-mediated heat shock response? We have previously identified more than a dozen proteins which form condensates in under two minutes in response to heat shock in vivo [131]. These so-called "superaggregators" all condense more quickly than Pab1 or Pub1; many of them reside in the nucleus; and most show substantial condensation at 37° C, unlike Pab1 or Pub1, but essential for any protein acting as an initial sensor of the Hsf1-mediated response. These thermosensitive proteins provide a compelling list of candidates for Hsf1 activators, and we predict they will have several characteristics shared by existing less-sensitive proteins: they will condense autonomously, recruit Hsp70 upon condensation, and shown condensation behavior that is suppressed at the resting intracellular pH.

Temperature acts as a physiological signal. Is Hsf1 activation a response to a heat-induced proteotoxic misfolding catastrophe, or something else? Heat-induced misfolding has long remained more a supposition than a result. While it is clear that artificially induced misfolded proteins can induce the heat shock response [46, 125], this does not constitute evidence that they serve as inducers under physiological conditions. As noted above, no

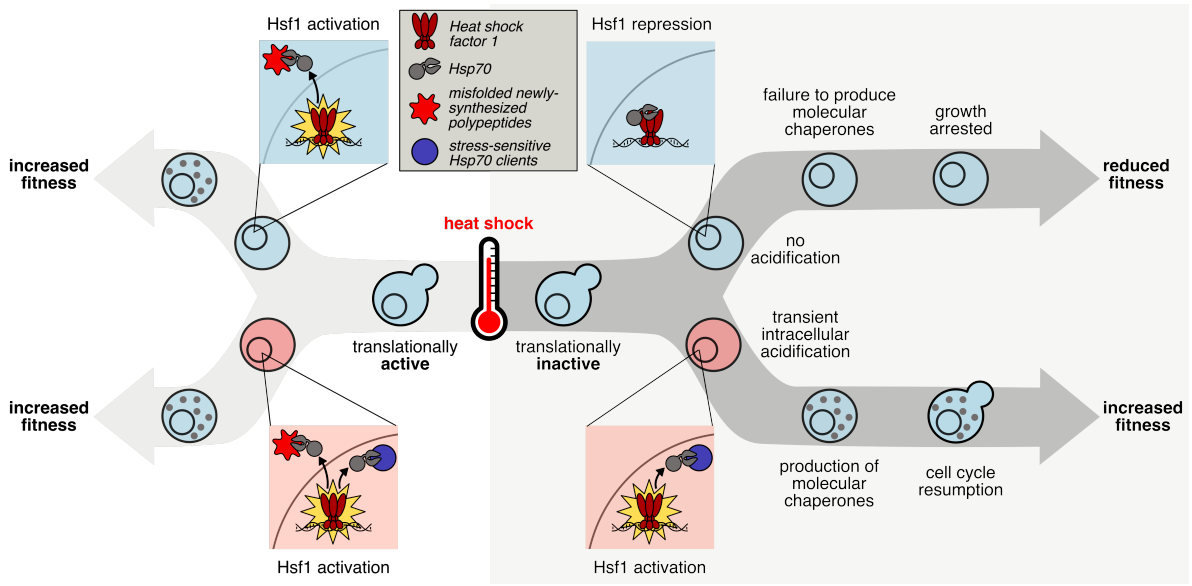

Figure 2.7: The transcriptional response to heat shock, chaperone production, and cellular fitness are promoted by intracellular acidification.
The key step in initiating the transcriptional heat shock response is the release of Hsf1 repression by Hsp70 through titration of the chaperone with clients. When cells are translationally active (left hand side), newly synthesized polypeptides that misfold in response to elevated temperature act as a trigger. However, cells that are not actively translating (right hand side) can still respond to heat shock, dependent on transient intracellular acidification, either during or after the temperature increase. We predict that pH-sensitive, stress-sensing proteins, similar to those already discovered, can act to titrate Hsp70, relieving Hsf1 repression and activating the transcriptional heat shock response.

specific endogenous protein has yet been identified which misfolds in response to a sublethal heat shock and thereby triggers the Hsf1 response. Here, we have shown that heat alone is insufficient to trigger the Hsf1 response, and that the newly synthesized polypeptides often cited as the primary inducers of Hsf1 are not required for Hsf1 activation.

An alternative to the misfolding model is that elevated temperature—within the physiological range to which organisms have adapted during their evolution—serves a signal, an environmental cue, which elicits an appropriate response.

Temperature acts as a physiological signal in other ascomycete fungi. For example, some dimorphic fungi live and grow in the environment as a mold, and convert into a yeast (a single-celled, reproducing fungus) in response to entering a mammalian host and detecting the resulting increase in temperature, the critical sensory cue [62]. The budding yeast and occasional human pathogen *Candida albicans* similarly requires a temperature increase to trigger the bud-to-hyphae transition critical for infection [13], which also induces chaperones in a classical Hsf1-mediated heat shock response [88].

The foregoing examples are pathogens. What physiological event would prompt the execution of such a heat-shock program in nonpathogenic *Saccharomyces cerevisiae*? *S. cerevisiae* does not produce fruiting bodies and depends upon animal hosts for dispersal [83]. This, along with other facts which we review here, suggests that a primary physiological heat shock for budding yeast is ingestion and dispersal by birds.

A survey of hundreds of migratory passerine (perching) birds [41] isolated yeast species from their cloacae, implying ingestion as the source; 14% of isolates were *Saccharomyces cerevisiae*. *S. cerevisiae* survived experimental passage through birds when inoculated in feed [41]. Passerine birds, the most numerous on earth, have an internal body temperature averaging 41.6° C (range 39° C to 44° C) when active, rising to an average of nearly 44° C (43.1 – 47.7) during high activity such as running and flight [99]. These temperatures correspond remarkably well to the upper bound of nonlethal temperatures for *S. cerevisiae* [109]. Ingestion will reliably induce a sudden thermal shift. The acidity of the stomach provides an ample source of protons to drive intracellular acidification.

A prominent ecological niche for *Saccharomyces cerevisiae* is the surface of fruits such as grapes [83], which birds eat—indeed, vineyard crop damage by passerine birds is a major challenge for the wine industry [117, 31]. Yeast proliferate to higher numbers on damaged fruit [83] which often results from bird pecking [42, 117]. Besides birds, other known dispersing hosts for the *Saccharomyces* genus include wasps, bees, ants, and fruit flies [27, 72, 83, 49, 23], all of which are preyed upon by birds, indicating that yeast may enter an avian carrier by multiple routes. Yeast cells that survive passage through a bird stand to benefit from broad geographic dispersal, an evolutionary advantage.

From these diverse and convergent lines of evidence, we conclude that ingestion and dispersal by birds is an ecologically established, physiologically relevant, and likely evolutionarily advantageous heat-shock condition for budding yeast. To obtain this advantage, yeast must travel through an acidic, low-nutrient environment averaging approximately 42° C.

Broader considerations. Recognition that a rise in temperature may represent a signal rather than merely a damaging agent alters how one thinks about the purpose of the response to temperature, the response's molecular triggers, and the physiological conditions under which the response would be deployed. Here, the suppression of the heat shock response by elevated pH suggests that acidification—and the capacity to acidify, which appears to be

determined in large part by extracellular pH—is a key part of the physiological context in which this thermal signal is received. This logic applies broadly. In humans, for example, a key physiological heat shock—fever—triggers the Hsf1-mediated heat shock response [115]. Perhaps fever causes new problems for cells, new self-inflicted damage to be cleaned up. More plausibly, however, fever acts as a systemic signal which activates a cellular program with key roles in modulating immune and inflammatory responses [115]. Indeed, the apoptotic response of human neutrophils to fever temperatures is sharply dependent on intracellular pH, with acidification promoting survival; local acidification is a hallmark of inflammatory conditions and promotes neutrophil activation [32].

We began by noting that the biological meaning of the longstanding association of cellular stress with cytosolic acidification, observed from fungal cells to vertebrate neurons, has remained unclear. Our results speak to a potentially broad effect: that this association is adaptive, and reflects, at least in part, the dependence of the core Hsf1-mediated transcriptional response on pH. Our work will focus a decades-long search for the specific eukaryotic sensors of heat shock on systems—likely, we argue, specific molecules—which depend on acidification for their sensory action.

2.5 Methods

Yeast strains

Scarless tagging of the Ssa4 protein with mCherry was accomplished in the BY4742 background via serial transformation and fluorophore exchange with the *URA3* gene such that no selection cassette remained in the genome. This was done by creating an intermediate strain with *URA3* at the C terminus of the *SSA4* locus; this sequence was replaced with mCherry and counterselection was done on 5-fluoro-orotic acid (5-FOA). The final strain has the *SSA4* gene in the native context with the native stop codon replaced by the mCherry sequence. In the BY4741 background, the coding sequence for pHluorin, under control of the constitutive *GPD1* promoter, was incorporated at the *LEU2* locus using Leu2 expression as a selectable marker. Strains were purified at least twice by streaking and picking single colonies, before being mated. The resulting strain, yCGT028 (MATa/α ura3Δ0/ura3Δ0 leu2Δ0/LEU2 pHluorin his3Δ0/his3Δ0 MET15/met15Δ0 lys2Δ0/LYS2 SSA4/SSA4-mCherry) was used for all experiments except those shown in Figure 2.13A, which uses strain yCGT032.

Strain yCGT032 was constructed in the same fashion, but with *SSA4* fused to a FLAG tag rather than mCherry.

Growth and stress conditions

Unless otherwise stated, yeast cells were grown at 30°C in synthetic complete media with 2% glucose (SCD) at pH 4. Under these conditions the doubling time of diploid cells was approximately 70 minutes. For all experiments, cultures were started from the same frozen stock, and grown so that the cell density was below optical density (OD) 0.1 for at least 12 hours before stress; a dilution of no more than 20-fold was performed at least 4 hours prior to stress. Cells were grown to between OD 0.05 and OD 0.1 (flow cytometry) or to OD 0.3-0.4 (mRNA-Seq) before being stressed.

All temperature stresses occurred at 42°C for 20 minutes, except for the data in Figure 2.1D and Figure 2.8C, which are 42°C for 10 minutes.

Measuring translation rate

Yeast cells were grown at 30°C with 250 rpm shaking in in synthetic complete media with 2% glucose (SCD) for gluocose withdrawal experiments or in YP + 2% maltose for maltose withdrawal experiments. Cells were grown to an OD600 of 0.2-0.3, then transferred to media adjusted to either acidic pH (4 for SC, 6.5 for YP) or at the resting pH (7.5), with or without 2% sugar, and containing ^{35}S-L-methionine and ^{35}S-L-cysteine at a final concentration of 1μCi/mL. Cells were grown at room temperature with no shaking (to emulate pre-stress conditions for all heat shock experiments), and aliquots were taken as a function of time. Proteins were precipitated by addition of 50% trichloroacetic acid (TCA) to a final concentration of 8.33%. Samples were placed on ice for 10 minutes, held at 70°C for 20 minutes, then returned to ice for another 10 minutes before being spotted onto glass microfiber filters. Samples were washed with 5% TCA, 95% ethanol, dried at room temperature for at least 24 hours, then placed in scintillation fluid. Radioactivity was quantified by liquid scintillation counting.

Flow cytometry

Technical information. Two cytometers were used to collect data: BD Biosciences LSR-Fortessa and BD Biosciences LSRFortessa-HTS. Both were equipped with 405, 488, 561, and 620 nm lasers. Cells were run on the lowest flowrate possible. Voltage and filter sets used were as follows (two filter sets were used on the HTS instrument):

Channel Name	Fluorophore	Fortessa HTS (1)	Fortessa	Fortessa HTS (2)
Forward Scatter (488)	NA	302	110	302
Side Scatter (488)	NA	242	236	236
PE Texas Red (561:610/20)	mCherry	550		
FITC (488:525/50)	pHluorin 488	450	422	422
BV421 (405:450/50)	NA	300	495	400
BV510 (405:525/50)	pHluorin 405	400	400	400
PEDazzle (561:610/20)	mCherry		625	625

All individual experiments were performed with the same voltage set, and the fluorescence values reported are normalized to a within-experiment fluorescence baseline (unstressed cells), allowing for direct comparison between experiments taken on different instruments or with different voltage sets.

Unstressed cells were used to determine manual gates on forward and side scatter to isolate cells. Growth conditions (see above section) were such that no significant populations of dead cells were expected. In some experiments a sub-population of cells became highly fluorescent in the BV421 channel. These cells were ambiguously bright in the FITC (488) channel, meaning that they could not be confidently assigned to either strain; although recorded, these cells were excluded from the analysis computationally by threshold gating in the BV421 channel. The percentage of these cells of the total, initially gated population was between 5 and 50%, and varied primarily with handling (no association with pH).

Dynamic intracellular pH measurements Cells constitutively expressing pHluorin in the cytoplasm (yCGT028) were grown as described in Growth Conditions above. A 400μL aliquot of cells was loaded onto the flow cytometer at room temperature and the instrument was run continuously for 5 minutes of equilibration. With the instrument still running, the sample tube was briefly removed and 1mL of media at 44° C was added (to account for heat loss in mixing); the tube was rapidly returned to the cytometer and held in a 42° C water

bath for 10 minutes, followed by 10 minutes at 30° C.

Sample size and reproducibility All flow cytometry stress experiments were performed at least in triplicate, with at least 10000 total events (cells) collected at each timepoint. Due to variation among partitioning between populations, the number of events for each sub-category varied, but was never below 1000 cells. All flow cytometry mock experiments were performed at least in duplicate, with the same standard for number of events as stress experiments.

pH manipulation

Calibration curve buffer. 50 mM NaCl, 50 mM KCl, 50 mM MES, 50 mM HEPES, 100 mM ammonium acetate, 10 mM 2-deoxyglucose; pH adjusted with HCl or KOH. 10 mM (1000x) nigericin in 95% EtOH was added just before buffer use to a final concentration of 10μM.

pHluorin calibration curve We used a protocol modified from [128]. Exponentially growing cells (OD 0.05-0.15) were spun out of SC media (3000g for 2–4 minutes) and resuspended in calibration curve buffer at 0.5 pH unit intervals between pH 4.5 and pH 8.5. Cells were equilibrated in buffer at room temperature for 15–30 minutes and then analyzed by flow cytometry. The calibration curve was generated by taking the median ratio of fluorescence in the 405:525/50 (BV510, pHluorin 405) channel to the 488:525/50 (FITC, pHluorin 488) channel, and fitting the resulting points to a sigmoid:

$$ratio_{405:488} \equiv R = \frac{a}{1 + \exp(-b(pH - c))} + d \tag{2.1}$$

where a, b, c, and d are fitting parameters. Ratios were corrected for background by subtracting the autofluorescence of unlabeled (wild type) cells in either media (for samples in media) or buffer (for the calibration curve). A new calibration curve was measured each time an experiment was performed. A representative calibration curve is shown in Figure 2.8B. A comparison between calibration curves in shown in Figure 2.8A: although the absolute value of the ratios may vary, the calculated effective pKa of the fluorophore is expected to

be consistent across experiments. The effective pKa was calculated using the formula: [4]

$$\log\left(\frac{R - R_{max}}{R_{min} - R}\right) = 0 \tag{2.2}$$

Determining ionophore efficacy at increased temperature To ensure that the ionophore treatment was effective at elevated temperature, the intracellular pH of cells in calibration curve buffer at 42° C was assessed. Cells were resuspended (at the same ratio of cells:buffer as used in pH manipulation experiments) in calibration curve buffer of varying pH and equilibrated at room temperature for 15 minutes. A small volume was used such that thermal changes rapidly equilibrated. A portion of the cells were analyzed by flow cytometry, and then the remaining samples were placed in a heat block at 42° C. Aliquots were taken at 10 and 20 minutes and analyzed by flow cytometry. The intracellular pH was calculated using a calibration curve generated at 30° C using different buffers. The close correspondence between the measured buffer pH and the calculated intracellular pH from the calibration curve is shown in Figure 2.4B.

Manipulating intracellular pH during stress Intracellular pH during stress was manipulated using calibration curve buffer. The concentration of the ionophore was low enough that any anti-biotic effects were negligible, as seen by the small fitness effect on pH-manipulated, unstressed cells (see Figure 2.15D, 'RT (mock)'.

1.2mL of cells grown as described in above 'Growth and stress conditions' section were spun out of media and resuspended in 60μL freshly prepared calibration curve buffer plus ionophore at the desired pH, equilibrated at room temperature for 15-30 minutes, and then either exposed to 42° C temperature ('heat shock') or room temperature ('mock') for 20 minutes. After stress, cells were recovered by removing the buffer and resuspending in 1.2 mL of fresh SC media and holding at 30° C with 250 rpm shaking. The fresh SC was either not pH adjusted (with a pH of approximately 4, data shown in Figure 2.4D, or was buffered to pH 7.4 using 0.1 M Na_2HPO_4 : NaH_2PO_4 buffer (data in Figure 2.4C).

Western blotting

yCGT032 was grown in 180mL SC media at 30° C shaking at 250 rpm for 12 hours then harvested at OD 0.026. Three aliquots of 50 mL cells were harvested by spinning at 3000g for 5 minutes. Each aliquot was washed with water and then resuspended in 1mL of a

different medium: SC, pH 6.8 calibration curve buffer with ionophore, or pH 7.4 calibration curve buffer with ionophore. Cells were equilibrated for 15 minutes at room temperature and then split into two samples, one for heat shock and one for mock treatment. Heat shock was performed by incubating cells in 42° C water bath for 20 minutes. Mock treatment was incubating cells at room temperature for 20 minutes. After treatment, cells were recovered for 60 minutes at 30° C. Protein was extracted by soaking in 0.1M NaOH followed by boiling in Laemmli buffer. Lysates were run on 4-20% SDS-PAGE stain-free gel, and imaged after UV activation to image total protein content. The gel was then transfered to nitrocellulose membrane. Dyed ladder was used as a guide to cut the membrane in half at approximately 50 kilodaltons (kDa). The lower part of the membrane was blotted for Hsp26 using a native antibody, a kind gift from Johannes Buchner. The upper half of the membrane was blotted for FLAG peptide with anti-FLAG (Proteintech 66008-2-ig). Western blots were performed using the 1-hour Western Kit from GeneScript (L00204 and L00205).

mRNA-Seq

Sample preparation (ionophore) (Data shown in 2.12) Cells were grown as described in 'Growth and stress conditions' section above, resuspended in 1 mL of freshly prepared calibration curve buffer plus ionophore, and equilibrated for 15 minutes before being heat stressed at 42°C for 20 minutes. Cells were resuspended in SC media and allowed to recover for 5 minutes before being harvested, resuspended and flash frozen in lysis buffer (20 mM Tris pH 8, 140 mM KCl, 1.5 mM $MgCl_2$, 1% Triton-X100). Two biological replicates were collected and averaged.

Sample preparation (media) (Data shown in 2.3, 2.10, and 2.11). Cells were grown as described in 'Growth and stress conditions' section above, resuspended in SC media with no pH adjustment (pH 4.0, acidification allowed), or adjusted to pH 7.5 using 2M KOH (acidification prevented). The following were then added to control translation state (all concentrations are final concentrations): 2% glucose (translation ongoing), 200 μg/mL cycloheximide (translation blocked), or nothing (0% glucose, translation blocked). Cells were heat stressed (42°C) or mock-treated (room temperature) for 20 minutes, spun down at 3000 g for 1 minute, and flash-frozen.

Library preparation (ionophore) Total cellular RNA was extracted using hot acid-phenol extraction and the resulting RNA was chemically fragmented. Samples were barcoded using a 3' adaptor with a unique sequence corresponding to each sample, and then pooled for downstream processing, as described in [114]. rRNA was depleted from the pooled samples using the Illumina Ribo-Zero Gold rRNA Removal Kit for Yeast (MRZY1306). Sequencing was performed at the Functional Genomics Core at the University of Chicago. Detailed protocol for library preparation is available; see [114].

Library preparation (media) Total cellular RNA was extracted from cells using the Zymo Direct-Zol kit (catalog number R2051). RNA was additionally treated with Turbo DNase (Invitrogen, catalog number AM2238), and libraries were made from the resulting material using the Illumina TruSeq Library Prep Kit **without** poly(A) selection.

Heat shock genes Genes upregulated during heat shock were curated by combining a list of Hsf1 targets from [98] and Hsf1 targets and Msn2/4 targets from [116].

Stress transcription factor determination Genes upregulated during stress were assigned to either Hsf1 or Msn2/4 as in [116, 98]. Briefly, the Msn2/4 genes were identified as genes that had a conserved Msn2/4 binding site in the upstream promoter and which were upregulated during heat stress in a strain of yeast where Hsf1 had been acutely deactivated. Hsf1 target genes were determined by differential expression after Hsf1 inactivation using a combination of transcript sequencing (RNA-seq), chromatin immunoprecipitation sequencing (ChIP-Seq), and native elongating transcript sequencing (NET-Seq). For 2.3D and E, transcription factors were identified using the YeTFaSCo database [29] to generate a list of proteins that have annotated DNA binding motifs (259 genes); the regulon of each transcription factor were determined by using the YeastMine database to generate a list of genes which had previously been shown to be regulated by each gene. The database includes interactions determined both during heat shock and non-heat shock conditions; 2.3D includes only transcription factors which had been assessed under heat shock conditions; 2.3E includes the regulons of other know stress-associated transcription factors which were determined under non-heat shock conditions. For both figures only regulons with 4 or more genes were considered (minimum 11 genes, maximum 1844 genes, median 72 genes), and the genes under control of Hsf1 or Msn2/4 were excluded from other regulons.

qPCR

Total cellular RNA was extracted from cells using the Zymo Direct-Zol kit (catalog number R2051). 100-200ng of RNA were reverse-transcribed (iScript cDNA synthesis kit; catalog number 1708891) using gene-specific primers. The resulting DNA was then used as a template for qPCR (idt PrimeTime Gene Expression Master Mix; catalog number 1055770). For *SSA4*, primers and probes against mCherry were used to detect the transcript; for all other genes assayed the native sequence was detected. All transcript abundances are either expressed as a ratio to a control gene (*TUB2*) in the same sample relative to the same value in unstressed cells (Figure 2.2E), or as the ratio to a control gene (*TUB2*) in acidified to non-acidified cells (induction ratio, Figure 2.3D).

Measuring fitness

Relative growth rate. Competitive growth assays rely on tracking the relative size of two populations of cells as a function of time, and differences in growth rate are inferred from these data. The ratio of two populations, for example pHluorin-expressing (pH) and wild-type (wt) populations, as a function of time is given by the following equation:

$$\log\left(\frac{n_{pH}(t)}{n_{wt}(t)}\frac{n_{wt}(0)}{n_{pH}(0)}\right) = (r_{pH} - r_{wt}t) \tag{2.3}$$

Where $n_x(t)$ is the number of cells of type x at time t, r_x is the instantaneous growth rate (in units of t^{-1}), and $\frac{n_{wt}(0)}{n_{pH}(0)}$ is the initial mixing fraction. This equality is true assuming constant exponential growth, which our data indicate is valid at least for the early stages of recovery; we only fit the linear portion of the data to ensure the validity of this assumption. For cells stressed without ionophore treatment, this was all timepoints less than 100 minutes, for cells stressed with ionophore this was all timepoints less than 160 minutes (this differents correlates roughly with the delay in induction we observe after ionophore treatment and is consistent across all pHs). See Figure 2.15F for all fits. We can use this equation to calculate the difference in growth rate, i.e. the fitness loss, for each population of cells having experienced stress at a different intracellular pH. This value is expressed as a difference to arrested growth (maximal fitness loss) in Figure 2.6B.

The reference population (subscript *wt* in the above equation) is wild-type cells growing exponentially ('spike' or 'spike-in'), which are distinguishable from the pHluorin-expressing

strains as they are are not significantly fluorescent in either pHluorin channel. Using a mixture of log-growing unlabeled and stressed labeled cells allows us to compare directly between the different pH and temperature combinations, as all the measured fitness loss values are relative to the same reference. It also implies that the difference $r_{pH} - r_{wt}$ will always be either 0 or negative, since the treatments being compared (pH manipulation either with or without heat shock) can only decrease the growth rate from maximal. To ensure that the pH manipulation itself was minimally stressful, the relative growth of pH-manipulated cells, which experienced 35 min at room temperature in calibration curve buffer with ionophore, was calculated and was found to be extremely close to 0 for all pH values considered (see figure 2.15D, 'RT (mock)' row).

To control for possible additional, strain-specific differences, we also calculated the relative growth rate when both the wild-type and yCGT028 cells were treated identically ('mix' or 'mix-in'); this value was also found to be nearly zero in every condition examined (see Figure 2.15D, 'Mix-in' column).

Determination of budded fraction We first computationally isolated the labeled, stressed cells, and then for this population looked at the distribution of values in the Forward Scatter Width channel. It has been shown that values in this channel correspond most closely to cellular volume and size [126, 57] because the measurement represents the amount of time spent passing in front of the interrogating laser. We note that there are two populations of cells, which we assign to budded (larger) and unbudded (smaller) cells (Figure 2.15A, density plot). This approach has been previously used to discriminate budded and unbudded cells [56]. Tracking the fraction of budded cells as a function of time gives information about cell cycle re-entry in a fashion analogous to the manual counting of budded and unbudded cells as previously performed [107].

To verify this labeling, we sorted cells into two populations based on the forward scatter pulse width into 95% ethanol to fix, and then visualized the fixed cells using light microscopy; Figure 2.15A shows sorting parameters and representative microscopy images. Cells from both populations were scored as either budded (containing an obvious bud that is at least 1/4 the size of the mother cell) or unbudded (having no bud). Full quantification is shown in Figure 2.15B. Fixed cells were then stained with Sytox to assess cell cycle position following a published protocol [106], and DNA content was analyzed by fluorescence intensity using flow cytometry. The 'budded' population contained more cells in the 2x DNA peak, indicating

that they were doubling their DNA and were thus actively growing; see Figure 2.15C.

Code and data analysis

Data analysis All data analysis was performed with R [101] using packages from the tidyverse [137]. Plots were made with ggplot2 [136]. Custom packages can be found on Github (repo: flownalysis). Raw data and scripts processing it to produce all figures that appear in this work are available online.

In general, summary lines on plots are fits of the log-transformed data with the form:

$$\text{fold change} = \frac{a}{1 + \exp(-b(\text{time} - c)} + d \tag{2.4}$$

where a, b, c, and d are fitting parameters, and d is constrained to be greater than or equal to 1. The exception to this are Figures 2.1B; 2.2A, C; 2.6E; and 2.14A, which are moving averages.

Statistical testing Statistical significance was determined with either the Welch two-sample t-test (Figure 2.2E) or the Mann-Whitney U test (Wilcoxon rank sum test) (Figure 2.5G and Figure 2.3C). All tests were performed using the stats package in the R programming language [101].

2.6 Supporting information

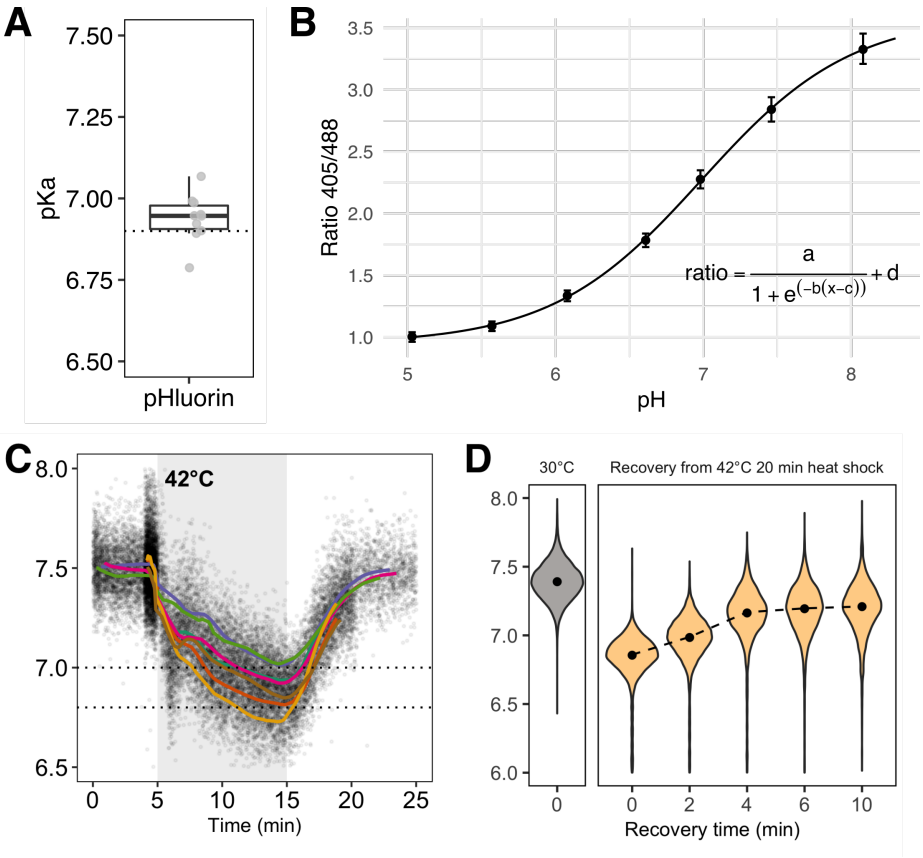

Figure 2.8: Measurement of intracellular pH during stress
(A) Comparison between calibration curves taken on different days. The curves are compared by solving for the apparent pK_a of pHluorin (see Methods section for equation); while absolute values of the ratio vary by day and instrument, the pK_a should be constant. The in vitro pK_a as calculated in [4] is shown with the dashed line. Each point is a separate experiment, n=10. See Methods for full details. **(B)** A representative pHluorin calibration curve showing the relationship between intracellular pH and fluorescence ratio. Error bars are the standard deviation of the population of cells measured. **(C)** Traces of intracellular pH as a function of time in cells expressing pHluorin and perturbed with a 42° C, 10 minute heat stress. Each point is an observation of a single cell; colored lines are the moving average of one experiment. Individual points are subsampled for clarity; each experiment has at least 10 000 cells. **(D)** Intracellular pH drops in response to a 42, 20 minute heat stress; the degree of acidification is the same as a 10 minute heat stress.

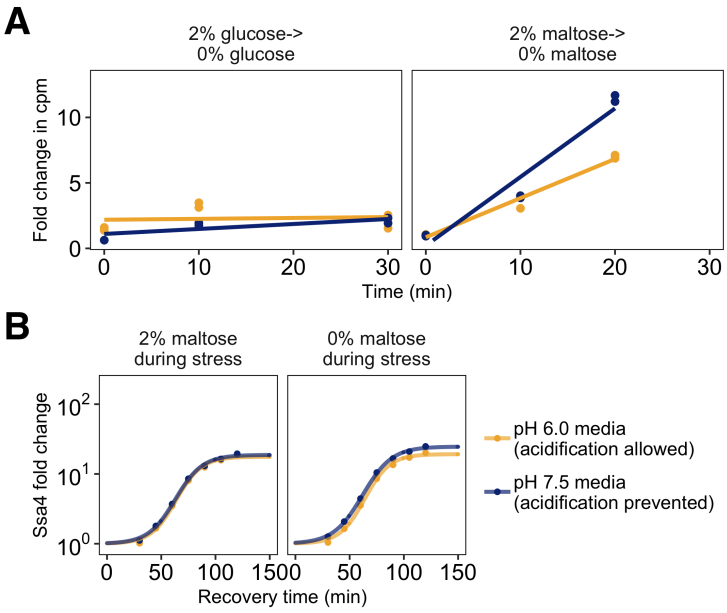

Figure 2.9: Preventing stress-associated acidification delays or impairs the heat shock response when translation is inhibited

(A) Measurement of the incorporation of radiolabeled amino acids into total cellular protein as a function of time after transfer to sugar-free medium. Translation abruptly ceases after withdrawal of glucose (left hand side), but continues after maltose withdrawal (right hand side). In both cases, acidification does not affect the translation rate. **(B)** Induction of Ssa4-mCherry for cells stressed after growth in maltose (left hand side) or growth in maltose followed by brief maltose withdrawal (right hand side). Yellow curves are data from cells in acidic media where acidification is prevented, blue are data from cells grown in media at the resting pH where acidification is prevented. Ssa4 induction after maltose withdrawal is pH-independent, demonstrating that translation attenuation rather than nutrient withdrawal explains the difference between induction in the right hand side of **B**.

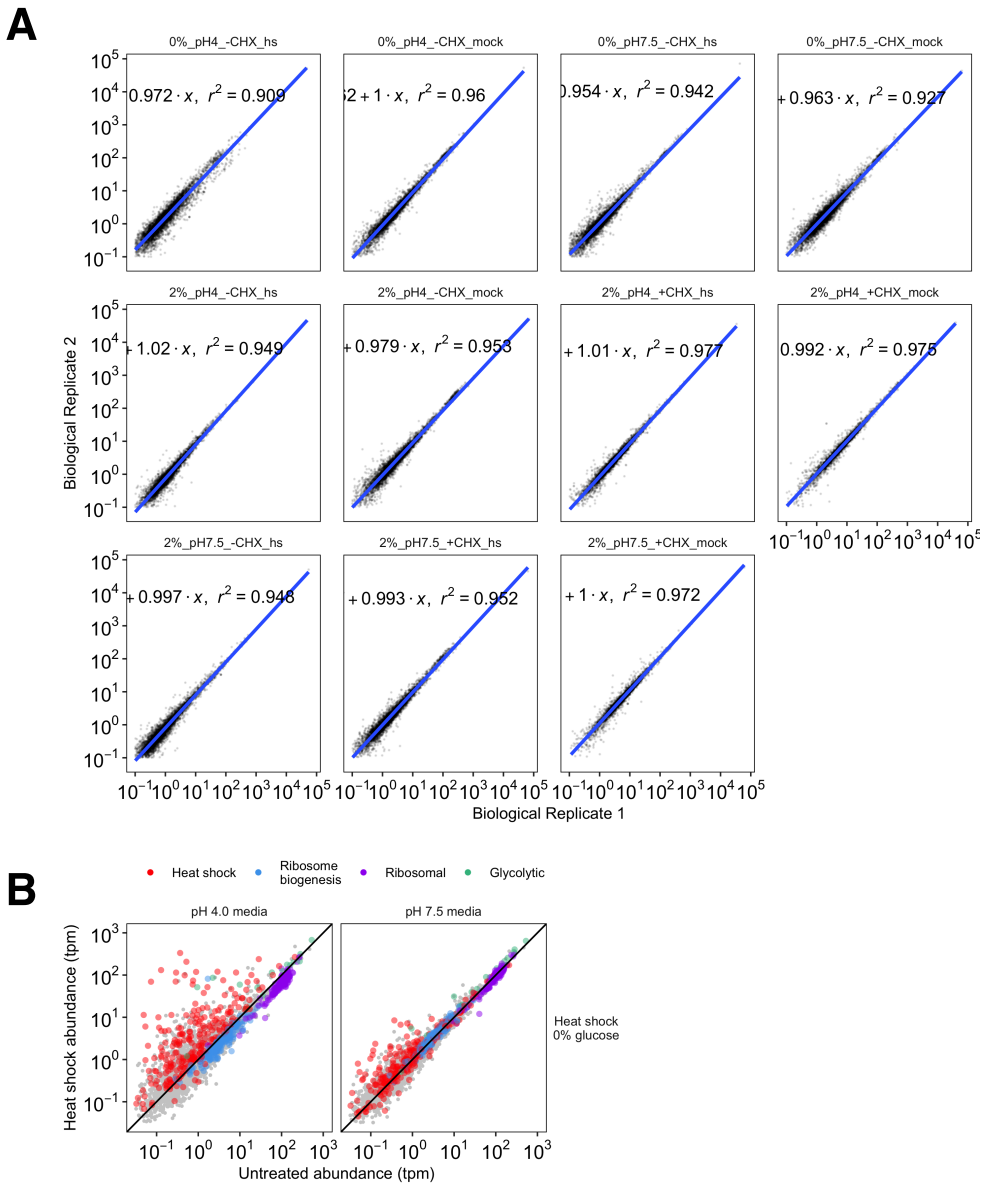

Figure 2.10: Failure to acidify during stress specifically represses Hsf1-activated genes
(A) Correlation between gene abundances in replicate samples (mRNA-seq data)
(B) Gene abundances in cells heat shocked after acute glucose withdrawal – unstressed abundances are shown on the x-axis, abundances after stress on the y. Color corresponds to gene type.

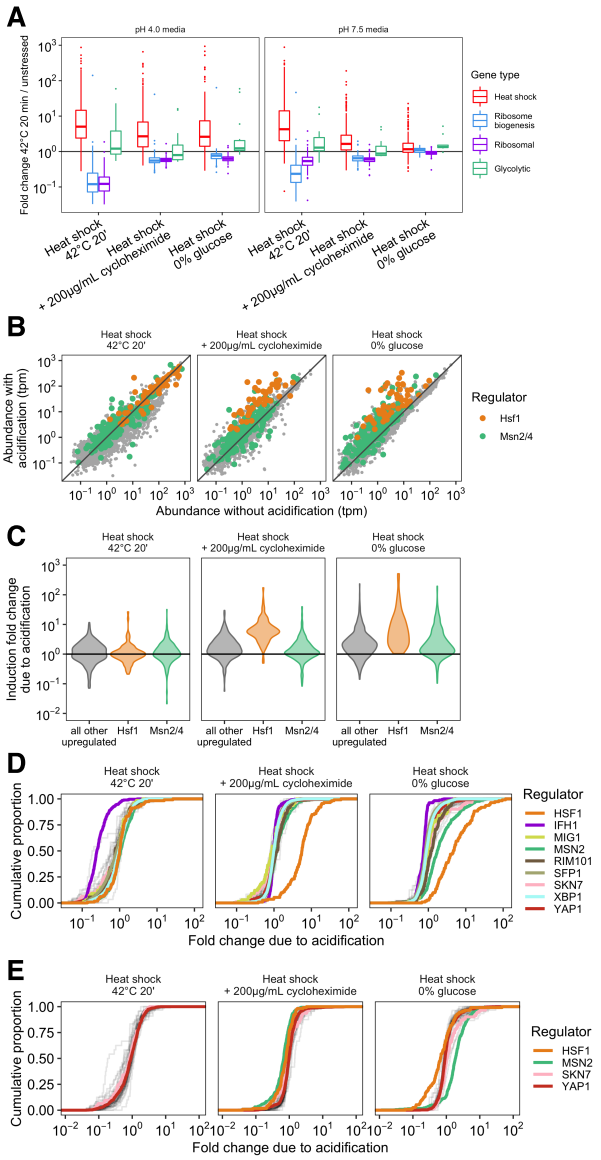

Figure 2.11: Failure to acidify during stress specifically represses Hsf1-activated genes
(A) Fold change distribution for groups of genes (colored by gene type) after stress. Populations that were allowed to acidify are on the left (pH 4.0 media), and those where acidification was prevented are shown on the right (pH 7.5 media). **(B)** Gene abundance after stress without (x-axis) or with (y-axis) acidification. Panels correspond to populations translating proteins (left) or where translation has been arrested (center and right). Color corresponds to the transcription factor responsible for the induction of that gene. **(C)** A global analysis of the pH-sensitivity of all genes induced during acute 42°C heat shock. The threshold for upregulation was set using a $>4\sigma$ cutoff for fold change in unstressed replicates ($\sim$2.5-fold). **(D)** Distribution of pH-sensitivity values for genes belonging to regulons of all annotated stress-responsive transcription factors (grey). TFs of particular interest are shown in color. **(E)** Distribution of pH-sensitivity values for the same set of TFs as (D), but shown after manipulation of translation state followed by mock treatment (30°C, 20').

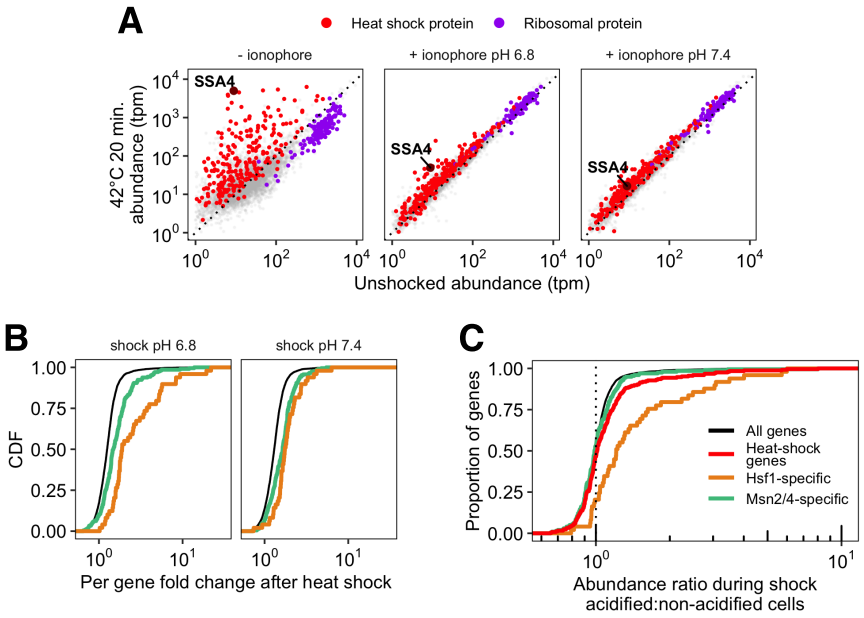

Figure 2.12: Failure to acidify during stress specifically represses Hsf1-activated genes
(A) Induction of heat shock genes in cells heat shocked without ionophore in acidic media (left panel), or in ionophore (center and right panels). **(B)** Induction (fold change in stressed cells relative to unstressed cells) for genes belonging to the Hsf1 (orange) and Msn2/4 (green) regulons. **(C)** pH sensitivity (abundance in heat shocked, acidified cells relative to heat shocked non-acidified cells) of genes belonging to the Hsf1 or Msn2/4 regulons, and the heat shock genes.

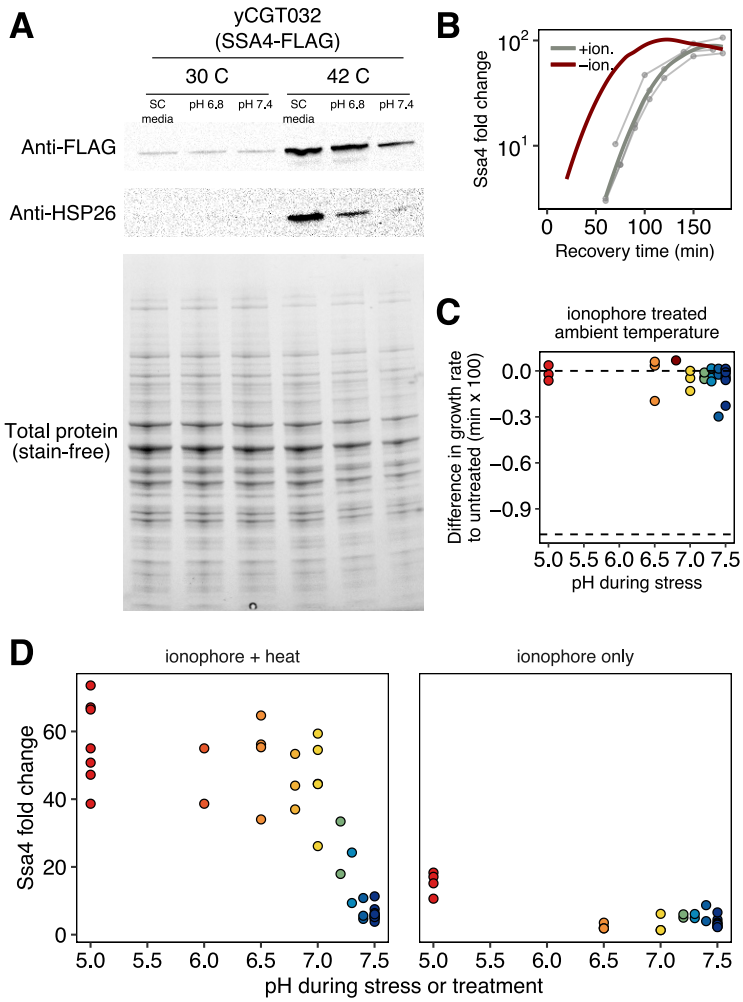

Figure 2.13: Quantitative control of intracellular pH using an ionophore
(A) Western blot and total protein gels for yeast carrying a the genomic copy of SSA4 tagged with a FLAG tag heat stressed with and without ionophore treatment, as described in Growth Conditions in the Methods section. Samples were taken 1 hour after stress. **(B)** Induction of Ssa4 during recovery from normal (red) or pH-manipulated (gray, pH 6.8) stress. Thin curves are individual experiments and thick curves are smoothed conditional means (see Methods for details). The red curve is the same data from Figure 2.1D for comparison. Although pH manipulation causes a delay in Ssa4 production, it does not affect the ultimate level of induction. **(C)** Growth rate difference in cells treated with ionophore for 35 minutes at room temperature followed by return to ambient growth conditions. Competitor was untreated cells. The values cluster around zero, indicating little to no loss of fitness due to ionophore treatment. Bottom dashed line shows theoretical minimum of the growth rate difference, which would result if cells completely arrested growth. These data are the same as those in Figure 2.6B, light-colored points. **(D)** Comparison of Ssa4-mCherry induction in ionophore treated cells that were either heat stressed (left) or held at room temperature (right). Acidification artificially induced by ionophore treatment does not cause appreciable accumulation of stress protein.

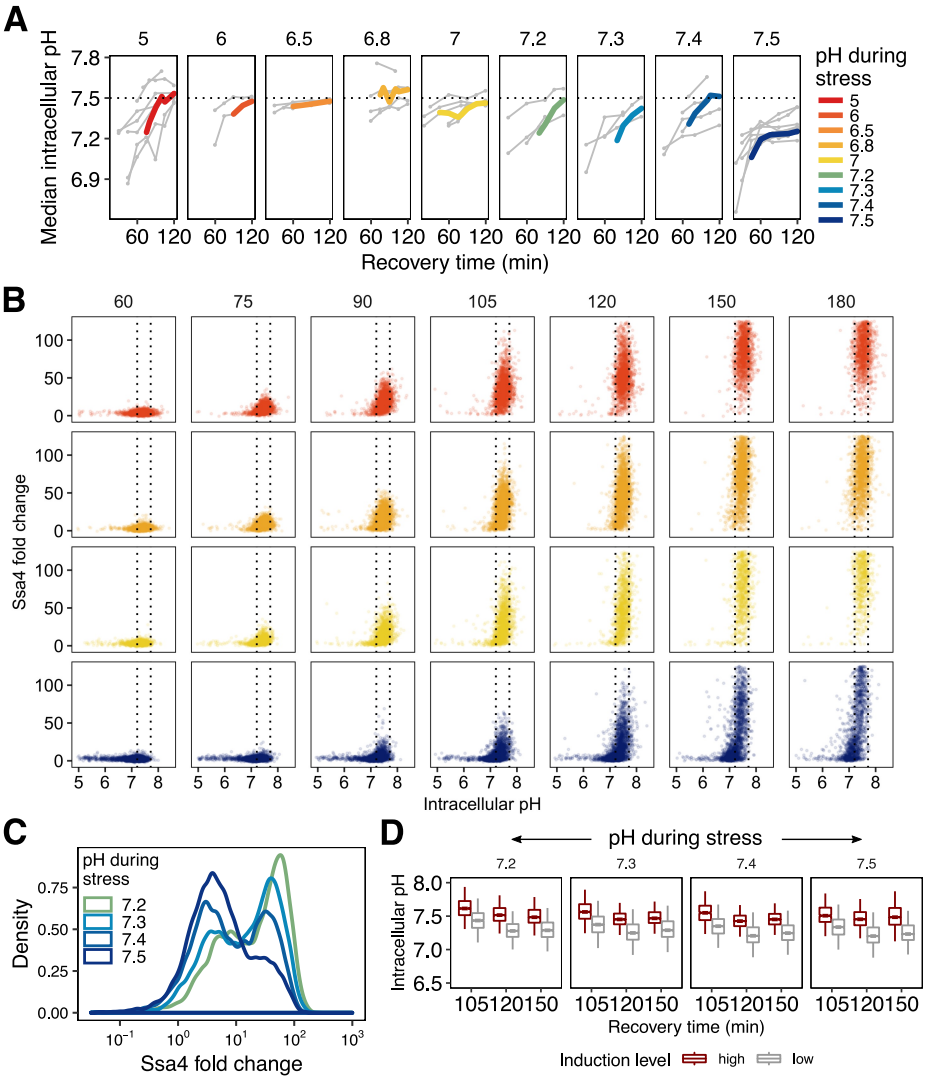

Figure 2.14: Post-stress acidification can rescue induction of the heat shock response
(A) Intracellular pH recovery after stress at different intracellular pHs. Thick, colored lines are the moving average of individual experiments (thin gray lines). **(B)** Recovery of intracellular pH is correlated with high Ssa4 levels on the single-cell level. Cells that are stressed at the resting pH have a large proportion of cells that do not recover intracellular pH and do not produce high levels of Ssa4. **(C)** Bimodal distribution of Ssa4 fold-change in cells stressed close to or at the resting pH. **(D)** Intracellular pH distributions for both high-expressing (red) and low-expressing (gray) cells for all conditions shown in C at multiple timepoints during recovery from heat shock.

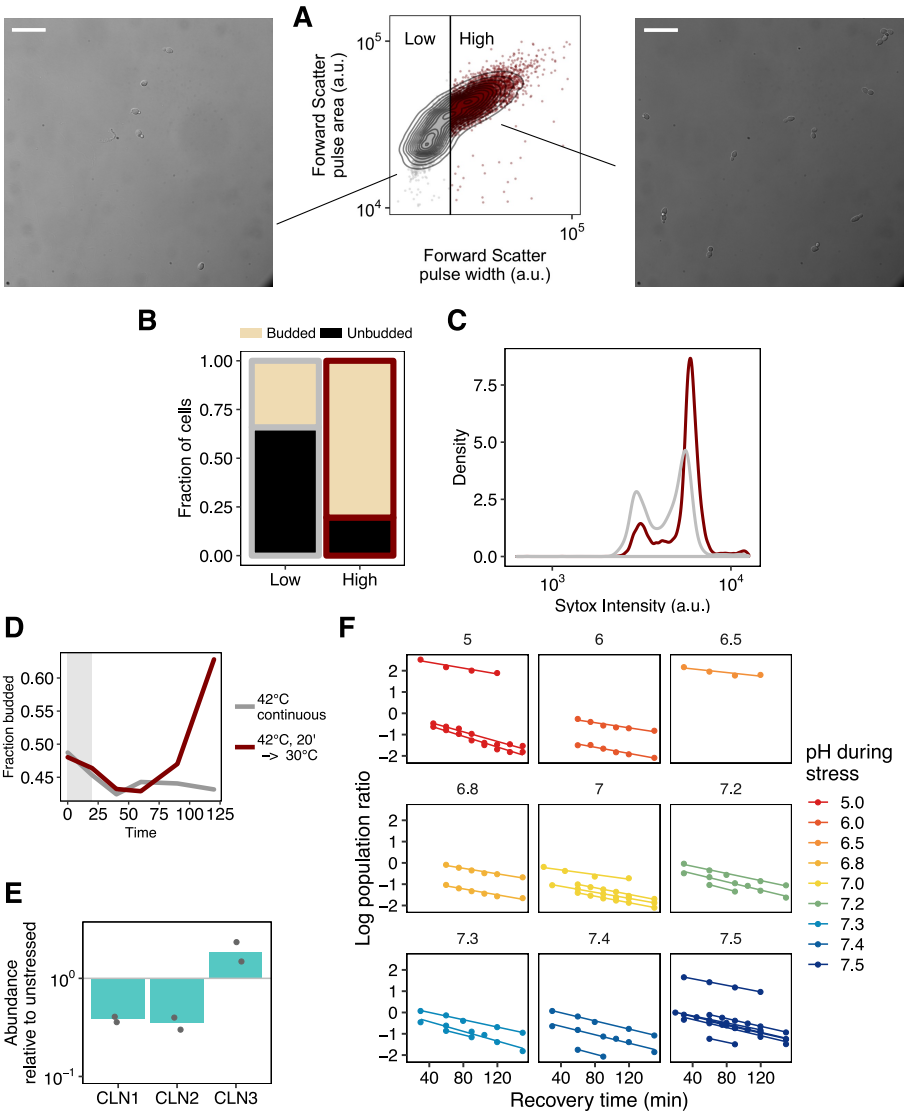

Figure 2.15: Fitness, intracellular pH, and heat shock protein production during recovery is correlated in single cells

(A) Cells were partitioned into two categories in the forward scatter width channel and were sorted based on this partitioning. Sorted cells were fixed and visualized by microscopy. Representative images for each population are shown. Scale bar is 25 μm. **(B)** Quantification of microscopy data; N = 217 cells scored. **(C)** Fixed cells were stained with Sytox and analyzed by flow cytometry to assess DNA content. The relative heights of the two peaks reflect the proportion of cells in each population that have doubled their DNA, and are thus actively growing. **(D)** Proportion of budded cells as a function of time for cells moved from 30 and held at 42 (gray line) or cells that experienced a 42 heat shock followed by recovery at 30 (red line). Both populations initially show a dip in the proportion of budded cells, but populations returned to ambient growth temperature then rapidly and synchronously re-enter the cell cycle, as evidenced by an increased in the proportion of budded cells. **(E)** Relative mRNA abundance for three cell cycle transcripts after a 42, 20 minute heat stress. Degradation of CLN1 and CLN2 is characteristic of heat shock [107]. **(F)** Fits used to determine the relative growth rate for all data shown in Figure 2.6B. The log of the population ratio as a function of time was fit with a line using linear least squares.

2.7 Appendices

2.7.1 Appendix I: Live Cell Measurement of the Intracellular pH of Yeast by Flow Cytometry Using a Genetically-Encoded Fluorescent Reporter

Abstract

The intracellular pH of yeast is a tightly regulated physiological cue that changes in response to growth state and environmental conditions. Fluorescent reporters, which have altered fluorescence in response to local pH changes, can be used to measure intracellular pH. While microscopy is often used to make such measurements, it is relatively low-throughput such that collecting enough data to fully characterize populations of cells is challenging. Flow cytometry avoids this drawback, and is a powerful tool that allows for rapid, high-throughput measurement of fluorescent readouts in individual cells. When combined with pH-sensitive fluorescent reporters, it can be used to characterize the intracellular pH of large populations of cells at the single-cell level. We adapted microscopy and flow-cytometry based methods to measure the intracellular pH of yeast. Cells can be grown under near-native conditions up until the point of measurement, and the protocol can be adapted to single-point or dynamic (time-resolved) measurements during changing environmental conditions.

Background

The intracellular pH of yeast is correlated with characteristics like viability and growth rate, and the regulation of intracellular pH consumes a large proportion of cellular energetic resources. [91] However, intracellular pH can change rapidly and is highly environmentally sensitive, so it is crucial to have a fast, minimally perturbative method of measurement for this important aspect of cell physiology.

Genetically-encoded biosensors that convert local concentrations of a compound of interest into fluorescent readouts have revolutionized our ability to characterize the intracellular environment. For some sensors, the absolute fluorescence intensity is correlated with the readout. This can be a problem when performing in-cell measurements, since the fluorescence depends both on the sensor expression level, which varies cell to cell, and the characteristic of interest. Ratiometric sensors, which depend on the ratio of fluorescence in two

different parts of the spectrum in the same fluorophore, do not suffer from this drawback. One such sensor, pHluorin [78], is a pH-sensitive fluorescent biosensor based on GFP that can be used to measure intracellular pH; the emission intensity (measured around 520 nm) after excitation in the near-UV and blue (405 and 488 nm typically) varies with pH such that the ratio of emission intensities can be related to pH.

In this protocol, we outline how to use pHluorin to measure intracellular pH in living budding yeast cells using flow cytometry. Flow cytometry is a method that combines microfluidic focusing and optical interrogation to measure the fluorescence of single cells in liquid culture with little to no special sample preparation required. Others have also used this method to measure intracellular pH in yeast cells. [133, 128] The advantage of flow cytometry is that it is much higher throughput than microscopy-based methods, [4, 92] while still being able to characterize individual cells and measure the fluorescence of multiple fluorophores. This access to both single-cell measurements and enough data to generate population-level statistics with a great deal of confidence is highly valuable. One potential downside to using flow cytometry to analyze the fluorescence of biosensors is that the spatial distribution of the fluorophore within the cell is not accessible (at least with traditional flow cytometry), as only a single, average fluorescence value in each channel is reported for each event (cell). However, for intracellular pH measurements in particular, the variation in any compartment (here, the cytosol, although pHluorin can be targeted to other organelles) [92] is expected to be minimal due to the unique properties of proton exchange in buffered aqueous solutions such as the cellular interior. [9]

The outline of a typical experiment is illustrated in Figure 1. Cells expressing pHluorin are suspended in buffer of known pH and an ionophore, in this case nigericin, is added. This addition makes the cell membrane permeable to protons and thus equilibrates the intracellular and extracellular pH. The ratiometric fluorescence of pHluorin is measured for these cells with known intracellular pH, and then these data are used to construct a calibration curve that can be used to convert measured fluorescence ratios in other cells to real intracellular pH values.

Materials and Reagents

Required reagents

1. Strain of *Saccharomyces cerevisiae* from the S288C background, such as BY4741/2/3.

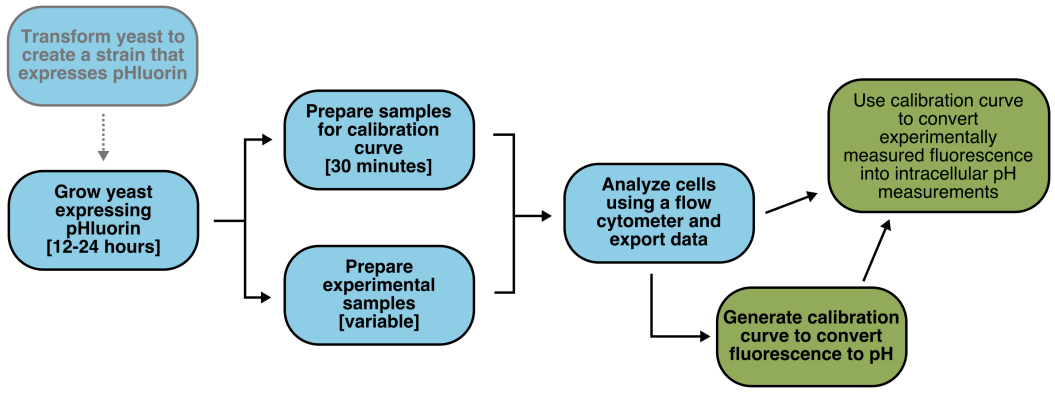

Figure 2.16: Overview of protocol for measuring yeast intracellular pH by flow cytometry
Experimental steps are colored blue, analysis steps are in green. Optional steps are in gray text

Alternatively, another strain of yeast may be used, as long as both non-fluorescent and pHluorin-expressing strains are available or can be made by the investigator (see optional reagents below).

2. 50 ml conical flasks (Olympus plastics, catalog number: 28-108)
3. 0.22 m filter (Fisher, catalog number: CLS 431118)
4. 5 ml polystyrene round-bottom tubes, 12 x 75 mm style (Corning Falcon, catalog number: 358058)
5. 2-Deoxy-D-glucose (Sigma, Sigma Life Sciences, catalog number: D8375-5G)
6. Nigericin (Adipogen, Adipogen Life Sciences, catalog number: AG-CN2-0020), prepared as a 10 mM stock in 100
7. Yeast Nitrogen Base (YNB) + Nitrogen (Sunrise Scientific, catalog number: 1501-250)
8. Synthetic Complete (SC) dropout mix (Sunrise Scientific, catalog number: 1300-030)
9. D-(+)-Glucose (Research Products International, catalog number: G32040-5000.0)
10. MES (Fisher, catalog number: BP300-100)
11. HEPES (GoldBio, catalog number: H-400-1)
12. KCl (potassium chloride) (Fisher, catalog number: P217-3)

13. NaCl (sodium chloride) (Fisher, catalog number: BP358-212)
14. 2 M HCl (hydrochloric acid) (Fisher, catalog number: A144-212)
15. 2 M KOH (potassium hydroxide) (Sigma, catalog number: 484016-1KG)
16. Ammonium acetate (Sigma, catalog number A1542-250G)
17. 2X Calibration Curve Buffer (see Recipes)
18. Yeast growth media (see Recipes)

Optional Reagents

1. pCGT05 or pHluorin expression vector (see Notes and attached plasmid map for more details; if the strain expressing pHluorin does not already exist and must be made by the investigator, then this reagent is not optional)
2. PmeI (New England Biolabs, catalog number: R0560S; comes with CutSmart Buffer)
3. Salmon testes DNA (Millipore Sigma, catalog number: D1626-250MG) prepared as a 2 mg/ml solution in TE Buffer (see Recipes)
4. Molecular Biology grade agarose (Apex Bioresearch Products, catalog number: 20-102GP)
5. 10x TBE Buffer (Bio-Rad, catalog number: 161-0733), diluted to 1x in ultrapure water
6. SCD (2%) plates without leucine, for selection
7. Tris-HCl (Fisher, catalog number: BP153-1)
8. Tetrasodium EDTA (Fisher, catalog number: S311-100) prepared as a 1 M pH 8.0 stock
9. Lithium Acetate (Sigma L6883-1KG)
10. Polyethylene glycol (PEG) 3350 (Sigma P4338-500G)

Equipment

1. Microcentrifuge capable of spinning 1.5 ml tubes at 3,000 x g

Reagent (Concentration)	**Amount (μL)**
pCGT05 (200 ng/μL)	5
Cutsmart Buffer (10X)	5
PmeI (10,000 units/mL)	1
ultrapure water	39

Table 2.1: Reagents for PmeI digestion

2. BD Biosciences LSR Fortessa flow cytometer (see note for instrument settings), or any flow cytometer capable of exciting between 380 and 410 nm and between 470-490 nm, and measuring emission (for both excitations) between 500 and 550 nm (see note). Other laser configurations could also be used; see Miesenbck et al. 1998 for full characterization of pHluorin

3. Electronic pH meter such as the Mettler Toledo SevenCompact pH meter S220 (Mettler Toledo, model: 30019032)

4. (optional, for yeast transformation)

   (a) Heat block/dry bath/water bath

   (b) Mold for casting agarose gels

   (c) Standard electrophoresis power supply and setup for running agarose gels

Procedure

1. (optional) Create pHluorin-expressing yeast strain

   (a) To integrate at the *leu2* locus (S288C backgrounds), cut vector pCGT05 with PmeI by mixing the reagents shown in Table 2.1 and incubating at 37 C for 60 min

      i. Alternatively, a plasmid with a selectable marker may be used rather than integration; see Notes. Use this as a template in the transformation in Step (c) iii and plate on appropriate selection media.

      ii. If the *leu2* locus is not available in the desired strain, integration at another locus is also possible.

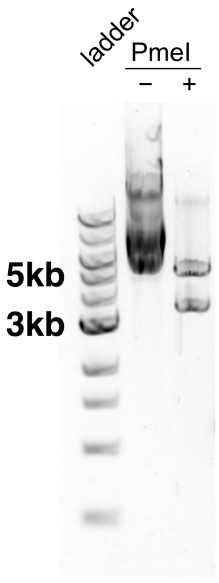

Figure 2.17: Vector pCGT05 before and after cutting with PmeI
The first lane is a ladder for size reference, the second lane is the vector without incubation with the restriction enzyme PmeI, the third lane is the vector after incubation with PmeI (there are two cut sites such that enzymatic digestion yields two fragments).

(b) (optional) Check digestion efficacy To check digestion, run 2-5 l of product on a 1% agarose gel, along with uncut vector and a ladder. The cut product should run at two bands, one at 5.2 kb and the other at 3.5 kb; see Figure 2.17.

(c) Transform into desired strain of yeast

i. To integrate into the leu2 locus, follow a standard lithium acetate yeast transformation protocol using the product of the above digestion reaction as the DNA template.

ii. Screen colonies on SC plates without leucine (30 C, allow 2-3 days for growth), and further colony purify on selection plates followed by non-selective plates to ensure that no wild-type cells remain.

iii. Cells in media may be mixed with glycerol to form a 20% final solution and stored at -80°C for years.

2. Grow cells (night before experiment)

(a) Inoculate starter cultures (at least 5 ml) of each strain of yeast youd like to measure; include wild-type yeast not expressing pHluorin (for background subtraction) and a sufficient volume of pHluorin-expressing cells to both generate

experimental samples and make a calibration curve to convert fluorescence to pH (see below for volume estimates).

i. Aim for cells to be at around OD_{600} 0.05-0.1 by the following morning (about 16 h later) so that they can either be analyzed right away, or mildly diluted in order to be analyzed later the same day.

ii. The calibration curve will require approximately 5 ml of pHluorin-expressing yeast at OD_{600} 0.05-0.2.

iii. The target OD_{600} for experimental samples is 0.05 to 0.1, lower than what is typically used for microscopy or biochemistry-based experiments. Calibration curve samples can be slightly higher, up to OD_{600} 0.2.

(b) Prepare 2x Calibration Curve Buffer. It can be prepared ahead of time and stored at 4°C.

3. Prepare calibration curve (morning of experiment)

(a) Add 5 ml of 2x Calibration curve buffer to eight 50 ml conical tubes. Add 3 ml of ultrapure water (or sufficient volume such that the electrode of the pH meter can be fully submerged in the liquid, keeping in mind that variable amounts of acid and base will need to be added to adjust the pH).

(b) Using concentrated HCl and KOH, adjust the pH of each tube to half pH units ranging from 5.0 to 8.5 (8 total; this spans the range of pH values to which the fluorophore is sensitive; measurement outside this range is not accurate). The initial pH should be in the range of 5.5-6.5. Make sure that the buffer is at room temperature before adjusting pH, as the pH can be temperature-sensitive.

(c) Once the pH is in the correct range, add ultrapure water until the final volume is 10 ml (calibration curve buffer is now at 1X). Note the final pH and record the exact value on the outside of the tube. pH-adjusted 1x calibration curve buffer (without ionophore) can be stored at 4°C for around a month. When re-using, bring to room temperature, re-measure the pH, and note any changes on the outside of the tube. Small changes ($\sim$ 0.05 pH units) are tolerable, but large changes indicate that the buffer should be discarded and re-made from the 2x stock.

4. Prepare the calibration curve (∼30 min before experiment)

   (a) Aliquot 1 ml of calibration curve buffer at each pH into labeled 1.5 ml tubes and to each tube add 1 μl of 10 mM nigericin in 100% ethanol to a final concentration of 10 μM. This step should be done as close to the time of measurement as possible, as the ionophore degrades more rapidly at room temperature.

   (b) Label additional 1.5 ml tubes for each pH that will be measured in the calibration curve (for cells), plus one additional tube for unlabeled cells.

   (c) Once cells are between OD_{600} 0.05 and 0.2, aliquot ∼500μl of pHluorin-expressing cells into each labeled tube. Aliquot one additional tube of unlabeled cells, which will be used for background subtraction.

   (d) Spin cells at 3,000 x g for 3 min. Decant supernatant (media) with a pipette, being careful to remove as much of the liquid as possible without disturbing the cell pellet.

   (e) Briefly wash cells by resuspending in 100 l of deionized water and spinning at 3,000 x g for 3 min. Decant supernatant.

   (f) Resuspend cells in 400μl of the appropriate pH prepared calibration curve buffer with nigericin.

      i. If many cells were lost during decanting (for example if the pellet was disturbed by pipetting) or if the concentration is low to begin with, a smaller volume of buffer may be used, but at least 200 l of volume is needed to make the measurement (this may vary with the model of cytometer used to make the measurements).

      ii. The unlabeled cells may be resuspended in any pH, as the fluorescence background does not vary with the pH of the buffer.

   (g) Allow cells to incubate in buffer for 15 min to equilibrate pH. After this, proceed to step 5 as rapidly as possible. It may be possible to hold the cells in buffer for longer without loss of integrity or a change in the fluorescence values, but this should be tested empirically by running full curves multiple times at different time intervals.

5. Analyze samples on the flow cytometer

(a) Briefly vortex tube or mix by inversion before loading each sample.

(b) Add 200μl of wild type (non-pHluorin-expressing) cells, in media, into a 5 ml polystyrene flow tube. Place the tube on the flow cytometer and begin acquiring data. While running the sample on the lowest possible flow rate, isolate the events associated with cells by drawing a gate in the forward scatter vs. side scatter plot to collect the vast majority of cell events, see Figure 2.18 for typical scattering profile.

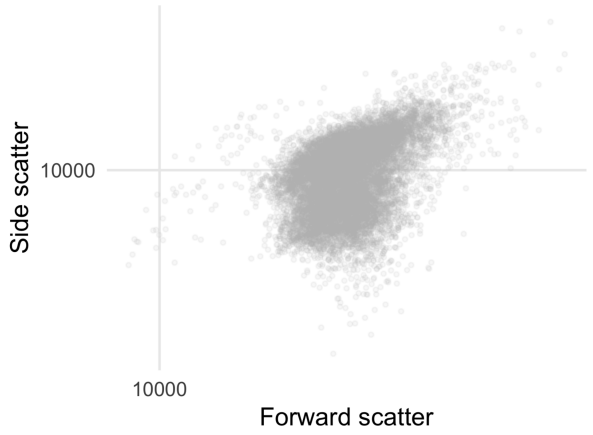

Figure 2.18: Typical scattering profile of cells
Data from a BD Biosciences LSR Fortessa with voltages as listed in Table 2.2.

i. Collect at least 10,000, but preferably 20,000+ events for each sample.

ii. Voltage settings for a BD LSR Fortessa instrument are given in Table 2.1. Other instruments and settings can be used; see Notes for more details.

iii. **It is essential that the same voltages are used within the same experiment, and preferably should not be changed even between experiments. Data taken with different voltage settings cannot be directly compared.**

(c) Using the same settings and volumes, run and collect data for pHluorin-expressing cells in media, and wild type cells in calibration curve buffer. All of these samples are essential for background subtraction.

(d) Using the same settings and volumes, run and collect data for all the samples in calibration curve buffer prepared in Step 4.

(e) Run any additional experimental samples using the same settings.
Note: In the case of experimental samples in different medium, a separate sample of unlabeled cells in the corresponding medium must also be analyzed in order to do proper background subtraction; see Data Analysis section below.

6. Export Data
Export all data. The standard format for flow cytometry data is the .fcs format; data analysis below assumes data in this format.

Data Analysis

The goal of the analysis is to first use data from cells in buffer to construct a calibration curve, which can then be used to convert the fluorescence readings of cells in media to intracellular pH. In the example script, the two green channels used are called FITC, corresponding to excitation at 488 nm and emission at 525/50 nm, and BV510, corresponding to excitation at 405 nm and emission at 525/50 nm. Below, they will be referred to below as F_{405} for emission intensity at 525 after excitation at 405, and F_{488} for emission intensity at 525 after excitation at 488. Superscripts will refer to the strain and background.
The following is a general description of the data analysis pipeline.

1. Load the data, contained in an FCS file, and convert it to human-readable data in table format

2. Calculate background readings
The median fluorescence of unlabeled cells (in both the FITC and the BV510 channels) will be subtracted as background. Unlabeled cells in buffer will be used to subtract background from labeled cells in buffer, and likewise for samples consisting of cells in media.

3. Make calibration curve

(a) For each cell (event), calculate the background-subtracted ratio of fluorescence that varies with pH.

$$R_{405/488} = \frac{F_{405}^{pH,buffer} - F_{405}^{wt,buffer}}{F_{488}^{pH,buffer} - F_{488}^{wt,buffer}} \tag{2.5}$$

(b) Bin/group events from each calibration curve sample, corresponding to each pH buffer. Calculate the median fluorescence ratio for each pH. It may be helpful to establish a fluorescence cutoff in order to account for any low, non-expressing, or damaged cells. However, this protocol should retain the majority of the events, and may not be necessary for strains with pHluorin genomically integrated.

(c) To make the calibration curve, first plot the known pH of each buffer against the median R405/488 value, see Figure 2.19 for an example calibration curve.

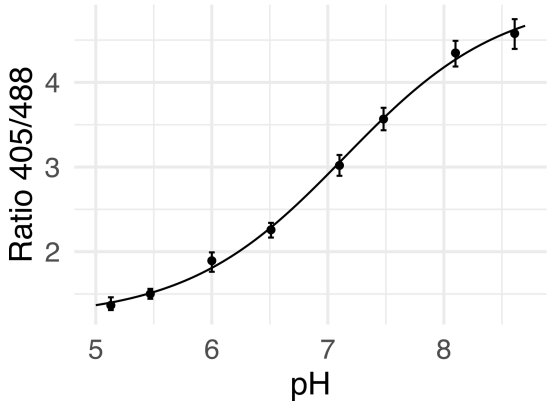

Figure 2.19: pHluorin calibration curve
Points are medians of a population of at least 10,000 cells, error bars are the 25th and 75th percentile.

(d) Fit the points to a sigmoid of the form

$$R_{405/488} = \frac{a}{1 + \exp(-b(pH - c))} + d \tag{2.6}$$

where a, b, c, and d are fitting parameters (a corresponds to the max of the curve, b to the steepness of the rise, c to the midpoint, and d to the baseline); see Figure 2.19. Once these parameters have been fit, this function can be used to map between the fluorescence ratio for any individual cell and intracellular pH by calculating the background subtracted R405/488 value, and rearranging the

Channel Name	Excitation (nm)	Emission (nm)	Voltage
Forward Scatter	488	NA	110
Side Scatter	488	NA	236
FITC	488	525/50	422
BV510	405	525/50	400
BV421	405	450/50	495

Table 2.2: Instrument settings (BD LSR Fortessa)

above equation as follows

$$pH = \frac{\ln\left(\frac{a}{R_{405/488}-d}+1\right)}{-b}+c \tag{2.7}$$

Note: This function can only be used in the range of sensitivity of the fluorophore, pH 5.0 to 8.5. [78] In general, a calibration curve must extend to these limits in order to set the baselines correctly, and only ratios that fall within the minimum and maximum pHs used in the calibration curve can be confidently assigned to a pH. Ratios that fall outside this range are ambiguous and should not be analyzed.

(e) Analyze experimental samples

i. Subtract background from each experimental sample by first calculating the median fluorescence of unlabeled cells in the same medium as the sample, then calculating the $R_{405/488}$ as in Step C1 above (substituting F_{405} and F_{488} values for unlabeled cells in medium rather than buffer).

ii. Convert $R_{405/488}$ to pH using equation 2.3.

Notes

1. *On yeast strains and strain availability:* In the attached example script, data collected from strain yCGT028, a diploid (BY4743 background) yeast strain expressing pHluorin from the *leu2* locus, which is empty (leu2Δ0) in the parent strain, and SSA4::mCherry. However, any strain expressing pHluorin, either from the genome or from a plasmid maintained with a selectable marker could be used in this protocol. Because the probe is ratiometric, the expression level variation expected for plasmid expression should not significantly affect the data analysis, however care should be taken that any cells that

fail to express the fluorophore are not included; because expressing and non-expressing cells are easily distinguished (see Figure 2.20), non-expressing cells can be manually gated out on the instrument, or excluded during data analysis.

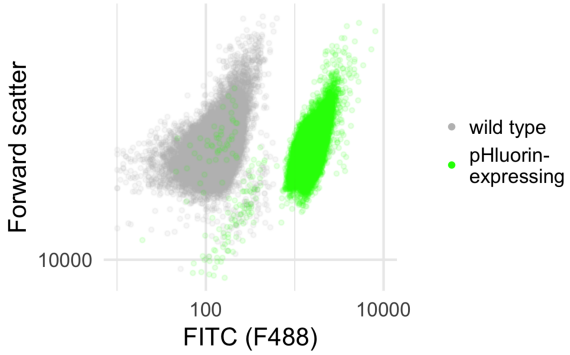

Figure 2.20: Separation of pHluorin-expressing and wild type populations by fluorescence
Data from a BD Biosciences LSR Fortessa with voltages as listed in Table 2.2.

2. *On model of flow cytometer:* In this protocol we give specific instructions for running samples on a BD Biosciences LSR Fortessa; but any cytometer capable of measuring excitation and emission in the correct range: excitation between 380 and 410 nm and between 470 and 490 nm and emission (for both excitations) between 500 and 550 nm. To set the voltages for these channels, it is helpful to run both cells in pH 5 and pH 8.5 calibration curve buffer, as these represent the maximum and minimum fluorescence intensities expected. [78] Both samples should be well-resolved in both channels (i.e., not saturating the detector at zero or the max value).

3. *On fluorescence data:* The data exported from the flow cytometer may contain pulse height and width information as well as the total fluorescence (area). Area measurements are used in all aspects of this protocol.

4. *On compensation:* Although compensation (a procedure used to calculate and correct for the overlap between fluorescence channels) is common in multicolor flow cytometry, in this protocol only one fluorophore is being analyzed and it is not necessary to perform compensation. pHluorin may be co-expressed with other fluorophores, such as mCherry, that are not excited in the same range, and analyzed in the same fashion.

5. *On frequency of measuring calibration curve:* Because instrument settings and mechanics can vary day-to-day, it is best practice to measure a new calibration curve every time

an experiment is performed. If ratio values seem to be very consistent day-to-day, then only the minimum and maximum pH buffers can be analyzed; if the previously generated calibration curve correctly predicts these samples, then that calibration curve can be used to analyze data taken that day. On reproducibility of calibration curves: Although the absolute value of the ratios may change on different days, the midpoint of the curve, which is the apparent pKa of pHluorin, should be consistent. To check whether a calibration curve accurately reproduces the known characteristics of the fluorophore, plot the pH of the buffer for each sample against the following quantity [4]:

$$\log 10 \left(\frac{R_{405/488} - R^{max}_{405/488}}{R^{min}_{405/488} - R_{405/488}} \right) \tag{2.8}$$

This equation can be fit by linear least squares, and where the resulting line crosses the y = 0 line is the apparent pKa of the fluorophore. The published in vitro pKa is 6.9. [4]

6. *On filtering events:* In some samples, depending on handling, there is a subpopulation of cells that both lose fluorescence in the pHluorin-associated channels, but also are much more (auto) fluorescent in another channel (ex 405, em 450/50). We interpret these cells as being dead or damaged in some way, and exclude them from analysis. They generally appear in experimental samples and not in calibration curve samples. See analysis script for an example of isolating and filtering out this subpopulation.

2.7.2 Appendix II: Ribosome profiling reveals that acidification-dependent transcriptional changes are matched by changes in translation

Introduction

The heat shock response is characterized by changes in both transcription and translation; [71] the abundance of heat shock protein transcripts increases by orders of magnitude; these transcripts are efficiently translated despite an overall reduction in the translational capacity of cells. [18] Although my results quantifying the transcriptome of cells heat shocked with and without acidification clearly established effects on the transcription of Hsf1 genes, it

remained possible that cellular translation was also affected by intracellular pH changes. My results showing differential production of SSA4-mCherry suggested that under the same conditions where the transcript was produced it was also translated, but was this true for all induced genes? And furthermore, did genes expected to be translationally repressed under these conditions actually show reduced translation? To answer these questions, I turned to ribosome profiling.

Ribosome profiling is an experimental technique to assay translation in vivo by next-generation sequencing of ribosome-protected fragments of mRNA. [59, 134, 80] Briefly, cell lysate is subjected to a mild RNase treatment to break up polysomes, and the resulting single ribosomes, still bound to the mRNA molecules that were being translated, are purified from the lysate by centrifugation through a sucrose gradient and fractionation. The fractions corresponding to fully assembled, single ribosomes containing an mRNA fragment are collected and the RNA extracted. The resulting material is then size-selected, depleted of rRNA, and a cDNA library is generated for sequencing. The end result is a pool of mRNA reads that correspond to the transcripts that were being translated when the cells were collected (Figure 2.21A).

I used this technique to assay protein translation in cells where I prevented or allowed acidification by heat shocking in buffers of different pH in the presence of an ionophore. mRNA-Seq was also performed on the same samples, such that both the amount of a transcript and it's interactions with the ribosome could be determined.

Results and Discussion

Figure 2.21A shows an overview of the experiment. Ribosome profiling provides information about the overall translational state of the cell as well as the translation of specific messages. Overall translation is summarized in a metagene profile, where the normalized, average read density across all genes is plotted as a function of position. A typical trace for log-growing yeast cells is shown in Figure 2.21B; density in the 5' and 3' ends is shown, with zero on the plot representing the Start and Stop codons respectively. The traces for unshocked and shocked are qualitatively very similar to those previous published. [47]

The samples treated with ionophore have a much different profile than those that are not ionophore treated; in particular there is strong depletion from the 5' end of the gene. This effect is likely due to the reduction in both initiation on transcripts and overall translation

that occurs upon the withdrawal of glucose, an effect characterized extensively in Section 2.3. In particular, this result mirrors the results in Figure 2.2E, providing additional evidence of pH-independent reduction in translation that occurs upon withdrawal of glucose. This effect complicates interpretation of the ribosome profiling data for these samples.

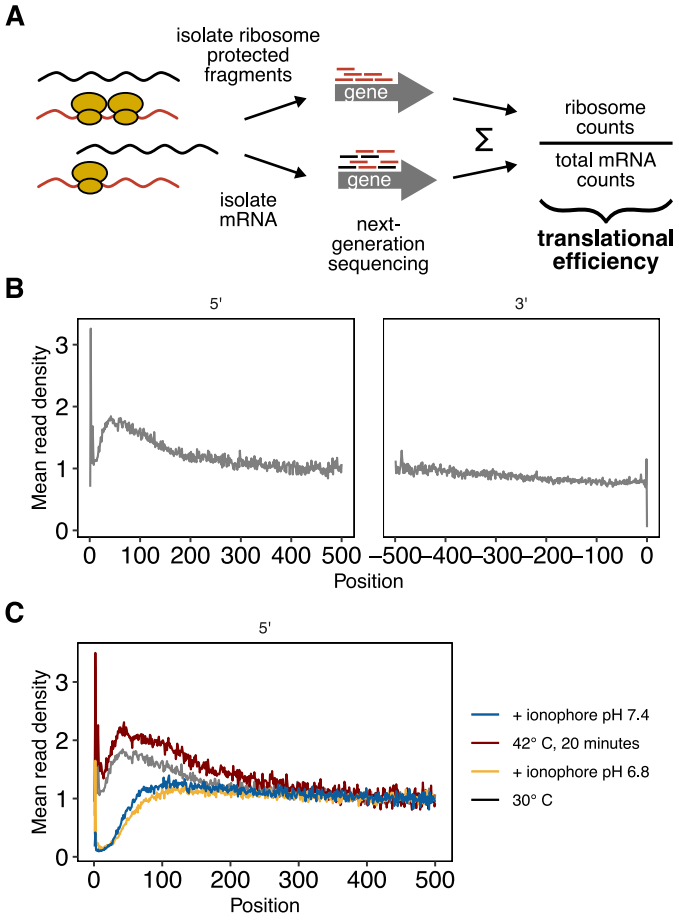

Figure 2.21: Ribosome profiling as a means to assay translation as a function fo pH during stress **A)** Overview of ribosome profiling experiment **B)** Metagene profile of unstressed cells **C)** Metagene profiles for cells stressed (42°C, 20 minutes) without ionophore (red), and cells stressed with ionophore treatment (6.8, yellow; 7.4, blue). A trace from unstressed cells (gray) is included for comparison.

Despite the difficulty in interpreting the metagene profiles, it is still possible to determine how well individual genes are translated under these conditions, and whether translation changed as a function of temperature and pH during heat shock. The "translational efficiency" ()TE) can be used to measure how well a transcript is translated — the TE is the ratio of the abundance of ribosome protected reads to the mRNA abundance (Figure 2.21A). We note that without internal controls (spike-ins of known concentration, for ex-

ample), TEs from different samples cannot be directly compared to one another; while the overall translational state can (and likely will) change as a function of the biological state of the sample, the amount of material analyzed is the same for each sample, meaning that the absolute abundance of transcripts loses meaning. However, we can still compare the *normalized* translational efficiencies, and in this way determine where a gene falls on the overall distribution of TEs, and how that value changes as a function of treatment, in this case heat shock and the pH during heat shock; an illustration of this concept is shown in Figure 2.22A.

The distribution of normalized TEs for each treatment condition is shown in Figure 2.22B. The unstressed, and stressed + ionophore-treated samples all look similar after normalization, while the distribution of TEs in the stressed cells is wider. It is likely that this result reflects the known changes in translation that accompany heat stress. [15, 18] Within each sample, examining the distribution of TE values for specific classes of genes gives further insight into the translational priorities of the cell. For example, during ambient growth conditions, heat shock messages are translated with approximately the same efficiency as all other genes, while transcripts encoding glycolytic and ribosomal proteins are more highly translated; this is evident from the shift to higher TE values in the green and purple curves in Figure 2.22C, left-most panel. During heat shock, however, heat shock messages pull away from the mean and become more well-translated relative to all genes (Figure 2.22C). However, this effect is not seen in the ionophore-treated samples; despite being heat shocked the by-gene-category distribution of TE values looks more similar to log-growing cells than to cells heat shocked without ionophore. This echoes the similarity in the overall distribution of TE values between the untreated and inophore-treated samples (Figure 2.22B).

To visualize the change in translation between different treatments, the difference in translational efficiency between either ambient (30°C) and heat-shocked (42°C, 20 minutes) samples (Figure 2.23A, C) or between heat-shocked and ionophore treated samples (Figure 2.23B, D) was calculated. Positive values reflect more efficient translation during heat stress versus ambient conditions, or in cells that acidified during stress (versus those that did not).

Figure 2.23A emphasizes the trend observed in Figure 2.22C; that during heat shock, heat shock messages are translated more efficiently than during ambient growth. Although this result seems intuitive, given the high degree of *transcriptional* upregulation that these genes experience when cells are subjected to high temperature, it is interesting to note that not only is the copy number of the gene increased, but the efficiency with which each of these molecules

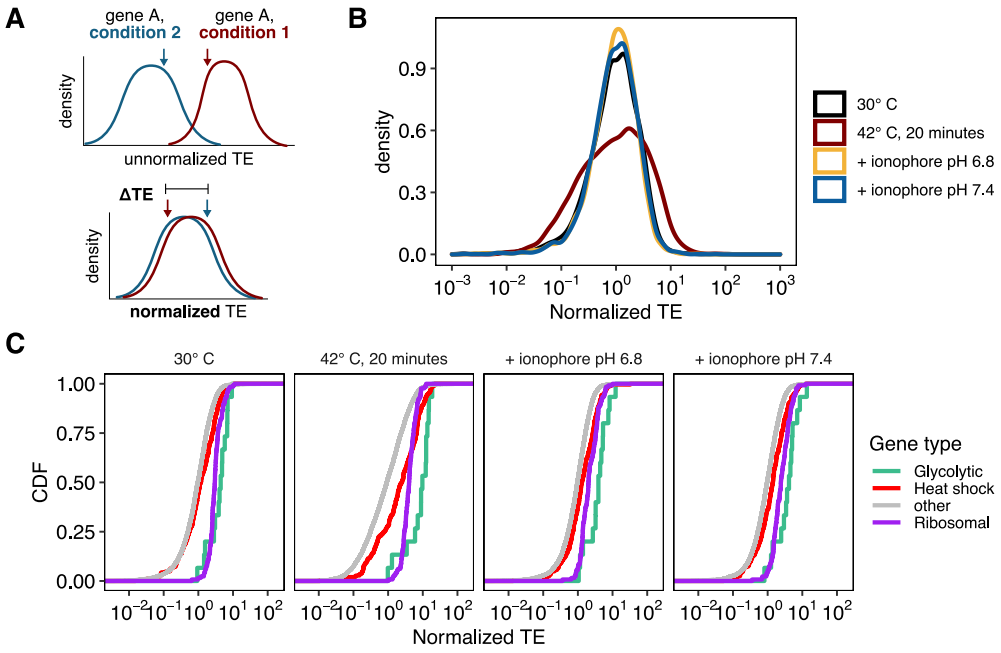

Figure 2.22: Distribution of within-sample translational efficiencies **A)** Illustration of normalization of TEs for calculating the relative change in TE (ΔTE). **B)** Distribution of all normalized TEs. **C)** Within-sample distributions of TEs as a function of gene type.

is translated also increases, reflecting a shift in cellular translation preferences. This same pattern is not observed when comparing cells stressed with and without acidification (Figure 2.23B). During ionophore treatment and heat shock, neither heat shock genes nor any class of genes we could identify are preferentially translated in acidified versus non-acidified cells. This could reflect unforseen interactions between ionophore treatment and the translational machinery.

It is interesting to note that another class of genes increases in translational efficiency during heat shock: the glycolytic enzymes (Figure 2.23A, green trace). This appears to be specific to energy production and not metabolism as a whole; ribosomal proteins are translated with the same efficiency. Given the decreased abundance of these transcripts (Figure 2.3A) we predict that under these stress conditions cells are producing fewer ribosomes.

When we assayed the pH-dependence of the transcriptional response, we found that Hsf1 targets were preferentially produced in cells that acidified during stress. However, it remained possible that there were additional differences in translation that amplified or suppressed this effect. To determine if this was the case, I plotted the distribution of translational efficiencies as a function of the transcription factor responsible for the induction of the gene.

When comparing stressed and unstressed cells (Figure 2.23C), both classes of genes, those induced by Hsf1 (orange) or Msn2/4 (green), had increased their TE on average. However, similar to what was observed for gene categories, these differences were not present when comparing TE as a function of pH during stress (Figure 2.23D).

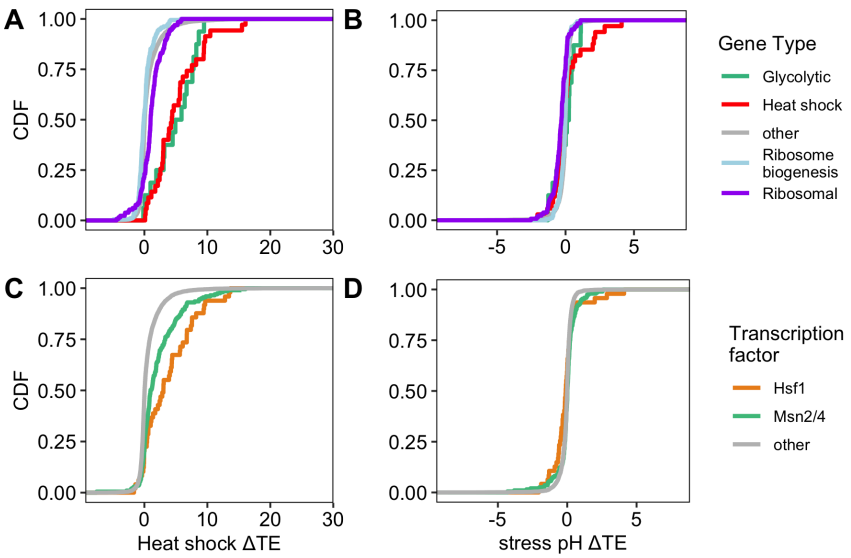

Figure 2.23: Change in translational efficiency during stress for several classes of genes. **A, B)** Distribution of the change in translational efficiency of genes as a function of the gene type due to heat shock and pH during heat shock respectively. **C, D)** Distribution of the change in translational efficiency of genes as a function of the transcription factor responsible for induction due to heat shock and pH during heat shock respectively.

Summary and Conclusions

While groups of genes with different biological function show interesting changes in translation efficiency when comparing between stressed and unstressed cells not treated with ionophore, there is a strong decrease in translation and polysome collapse associated with ionophore treatment (Figure 2.21C) and glucose withdrawal [3] that complicates the interpretation of these data. The relative translation of the heat shock messages is similar between acidified and non-acidified cells (ionophore treated), indicating that the phenotypic outcomes reported in the main text (Figure 2.2B and 2.4C, D) are largely due to transcriptional differences rather than differences in translation.

CHAPTER 3
SIMULATION AND EVOLUTION OF CHARGE PATTERNING IN PROTEIN INTRINSICALLY DISORDERED REGIONS

3.1 Introduction

I have demonstrated that intracellular pH change during the heat shock response can activate Heat Shock Factor 1 (Hsf1). Given the mechanism of Hsf1 activation (briefly, its repression by molecular chaperones is relieved by their engagement with other clients following the onset of stress), we know that molecular conformation plays a key role in activating this transcription factor. I have demonstrated that pH change can help trigger the response, implying that a pH-induced conformational change in a protein promotes interactions with molecular chaperones (specifically Hsp70 [64, 142, 75, 113]) to activate Hsf1. This observation led me to more deeply consider what chemical changes occur in proteins in response to pH change, and how such changes could affect protein structure.

Fundamentally, changes in pH alter the patterning, location, and interaction of charges, and these changes can alter peptide or protein conformation; [35] pH affects the charge state of amino acids in proteins and the attraction and repulsion of charges influences both local and global structure, [35, 63] as do interactions between charged and polar species. Furthermore, some proteins that phase-separate in response to stress-associated pH changes have been proposed to be pH sensors or have phase diagrams that depend on pH; [43, 103] however the molecular details of how such phase behavior is modulated is lacking. Finally, the fitness-promoting effects of intracellular acidification imply that pH-sensitivity is selected for. If structural sensitivity to pH is part of the evolved role of a protein, then we should be able to find signatures of selection in the sequence or properties of the protein that reflect this evolutionary pressure. This chain of logic led us to examine the evolution of charge patterning in intrinsically disordered sequences as a first step towards understanding evolved pH-sensitivity. Although this work is motivated by stress-associated acidification, it is relevant to more basic questions about protein structure and function and how these properties change in response to changing physicochemical environments.

The model system I chose to investigate is highly-charged intrinsically disordered regions (IDRs) within otherwise structured proteins. These regions, which often have a high overall fraction of charged residues (FCR) and a low overall charge asymmetry (σ, see Equation

3.1, [28]), present novel opportunities and challenges. Investigating IDR structure is experimentally difficult – IDRs sample many conformations, and quantifying the distribution and dynamics of such states, while required to characterize the proteins, is not accessible to many molecular techniques, and for those that are cannot be done in a high-throughput fashion. However, newly-developed theoretical methods have greatly facilitated molecular simulation of such sequences. [33, 34, 120, 19] There are two primary reasons why such sequences, of which there are hundreds in the yeast proteome, are of particular relevance. 1) IDRs are regions of proteins without inherent secondary or tertiary structure; these regions are often more expanded and solvent-exposed than folded proteins of similar length, leading to an increased interaction with and sensitivity to solvent. Changes in pH will likely have large effect on the distribution of charges within these sequences. Furthermore, simulations of synthetic sequences consisting only of charged residues have shown a strong connection between the arrangement of charges within a sequence and the resulting ensemble of structures that such a sequence takes on, both for individual chains [28] and assemblies composed of such sequences. [120, 52, 69] However, these simulations were performed on polyelectrolytes; whether charge patterning is still important for 'real' protein sequences that contain other residues, and if such properties are selected for, remain open questions.

Because of the previously-mentioned difficulties in experimental study of such sequences, I turned to coarse-grained simulations to understand the relationship between charge patterning and molecular structure. Although simulating the effect that pH changes has on the location and amount of charges in native sequences presents many challenges beyond the scope of this work, I began to address the issue by looking at the relationship between charge patterning and structure, with the implicit understanding that pH changes can change charge patterning, and thus that the findings presented here will be relevant to future studies that consider the effect that pH change has on the charge distribution within such sequences. By examining how this charge patterning, a factor that pH change can influence, affects the structure of this region, I explored the relationship between sequence characteristics such as FCR, σ, and the structural properties of the region. Importantly, I linked these findings to evolutionary constraints by examining tens of paralogous sequences and showed that a region of interest, despite having large changes in length and sequence, showed enrichment for short, regularly spaced highly charged units, suggesting that such sequences are selected for. Finally, I used a toy model of evolution to select for sequences with a particular physical characteristic, and showed that this leads to biases in patterning distribution, demonstrating

the feasibility of detecting selection on physical properties of IDRs by identifying selection on characteristics of the sequences themselves.

3.2 Results

A charge patterning metric designed for polyampholytes is appropriate for native sequences

To begin to analyze charge patterning, we must start with a metric that quantifies the distribution of charged residues in these regions. Several metrics have previously been developed; one in particular from Das et al [28] has many convenient and appropriate characteristics, described below, and I chose to use this as a starting point to characterize the sequences of interest. This metric, called κ, ranges from zero (entirely well-mixed charges) to one (entirely separated charges), as illustrated in Figure 3.1A. More precisely, it is an average (over window lengths) measure of the mean deviation in charge asymmetry of windows (of which there are g along the sequence) from the overall charge asymmetry:

$$\sigma = \frac{(f_+ - f_-)^2}{(f_+ + f_-)} \tag{3.1}$$

$$\delta = \frac{\sum_{g=1}^{N_g} \left(\sigma_g - \sigma\right)^2}{N_g} \tag{3.2}$$

The calculated δ values are then normalized, and averaged over two window lengths:

$$\kappa_i = \frac{\delta_i}{\delta_{i,max}} \tag{3.3}$$

$$\kappa = \frac{\kappa_5 + \kappa_6}{2} \tag{3.4}$$

In addition to providing a bounded, normalized quantification of charge patterning, this metric is also useful because it has been shown to correlate with the compaction/extension of the sequence, measured by the radius of gyration (R_g, see Methods for definition). This metric was designed to characterize IDRs, but it was developed and explored largely in the context of polyampholytes consisting of *only* charged residues. [28] Although initial applications to native sequences, which are heterogeneous and contain other species which

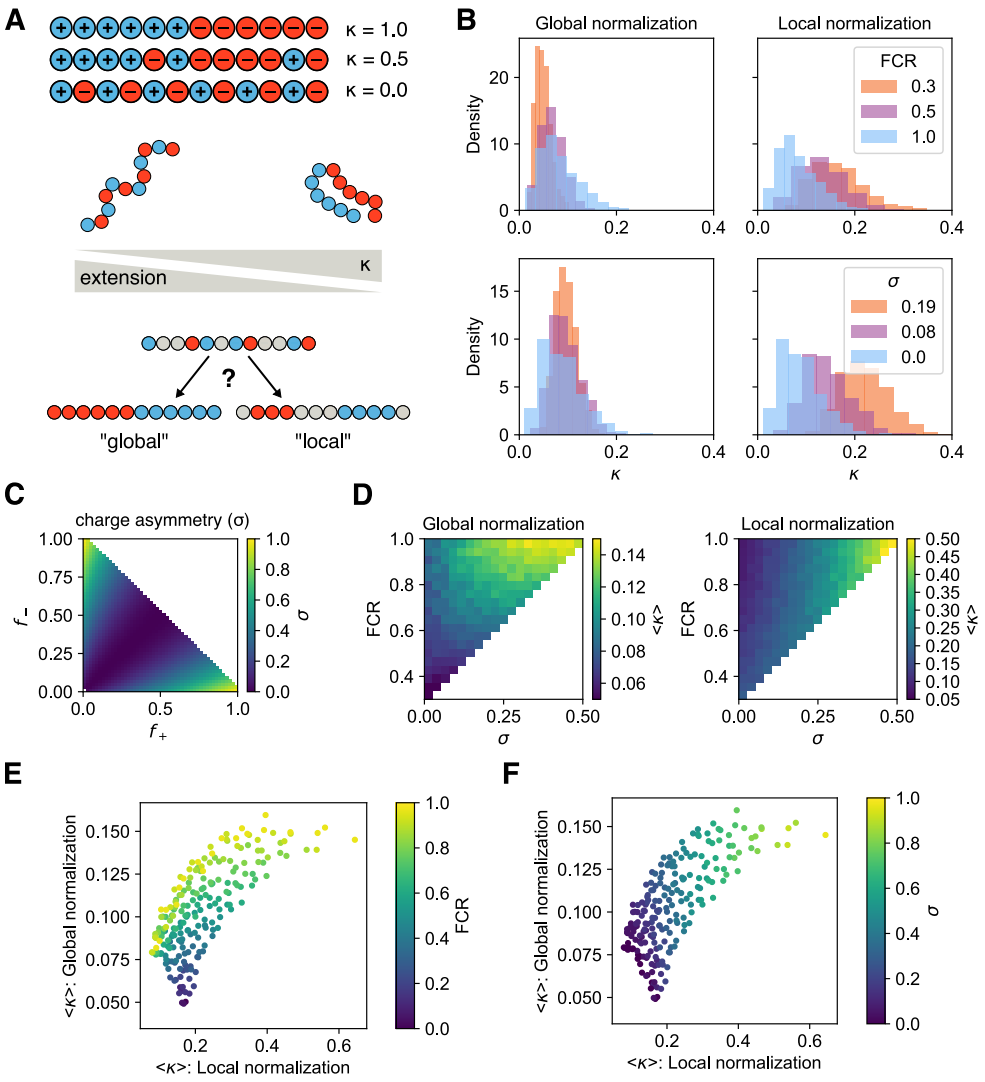

Figure 3.1: κ **as a metric to quantify charge patterning**
A) Illustration of κ, its relationship to charge distribution and structure, and two possible methods for normalizing the metric. **B)** Distribution of expected κ values as a function of FCR (top two panels) or σ (bottom two panels). Both global (left) and local (right) normalization is shown. Each plot shows the distribution of κ values for 1000 randomly-generated sequences for each value of FCR or σ. **C)** Heatmap of charge asymmetry (σ) as a function of the fraction of positive and negative residues in a sequence. Note that the distribution is non-uniform; that is there are many more sequences that produce a $\sigma < 0.5$ than $\sigma > 0.5$. **D)** Mean κ from randomly-generated sequences (100 per entry) of a given FCR and σ, using local (left) or global (right) normalization. **E)** Direct comparison of the mean κ values for both normalizations colored by FCR **F)** Direct comparison of the mean κ values for both normalizations colored by σ

are not charged, were promising, it remains an open question whether this metric, and especially its connections to physical characteristics apply to sequences with FCR < 1.0 and charge asymmetry that is not precisely 0.

In particular, the analysis of sequences with an FCR < 0 calls into question the nor-

malization of κ (δ_{max} in Equation 3.3) – should sequences be normalized to the overall most charged and most separated sequence of the same length ("global" in Figure 3.1A) or to the most separated sequence of the same length and FCR ("local" in Figure 3.1A)? One can imagine advantages and disadvantages for each: comparison to a global standard could be more appropriate for comparing sequences with different FCRs, but doing so might mask small differences when comparing sequences with the same or similar FCR. Even more importantly, do both normalizations still predict molecular structure, as was demonstrated previously (see Figure 3.1A and [28])?

To address these questions, I first examined the expected distribution of kappa values for sequences with a given FCR and σ. I generated random sequences of charged amino acids (specifically glutamate and lysine), varying both FCR and σ (see Figure 3.1C for a demonstration of the relationship between these two variables), and calculated the resulting κ value for each using both normalization schemes. The results are shown in Figure 3.1B. The most obvious difference between the two normalization schemes is that the local normalization generates higher values of κ, and that this effect is more pronounced as FCR decreases or σ increases. The global normalization, on the other hand, is much less sensitive to either parameter, particularly σ. It is important to note that in comparing naturally-occurring sequences with different FCRs, a normalization like the global normalization that puts all sequences on the same scale may be more appropriate than one which prioritizes within-sequence normalization ("local").

To more systematically address the relationship between FCR, σ, and κ, I expanded the sequences studied in FCR and σ space, and calculated the mean expected κ for each combination of FCR and σ (Figure 3.1D). Note differences in magnitude of the color scale; the right hand side goes to 0.5 while left hand side only goes to 0.14. Moreover, the shape of the distributions is different – for the local normalization κ is directly correlated with σ for a given FCR, whereas for the global normalization for a given σ, $< \kappa >$ corresponds more with the fraction of charged residues and has an inverse trend to the local normalization (see Figure 3.1B) – that is low FCR corresponds to lower κ values, and higher FCR corresponds to higher κ values. Since when comparing among homologous sequences the FCR often varies more than the sigma values ([28] and Figure 3.4C), it is possible that using the global normalization may be more appropriate than the global normalization.

Comparing the two metrics directly yields a rough positive correlation, which is stronger for high (relative) κ values and much lower for low-κ values (Figure 3.1E,F). Figure 3.1E

and F are colored by the fraction of charged residues and charge asymmetry respectively, and show that for sequences with FCR greater than ∼0.3 (which are reasonably the only sequences for which a measure of charge patterning would be relevant) we can expect some positive correlation between the metrics (at least on average). However, for sequences with a very low charge asymmetry (like the native sequences considered here) we can expect a negative correlation (as was observed in Figure 3.1B and D).

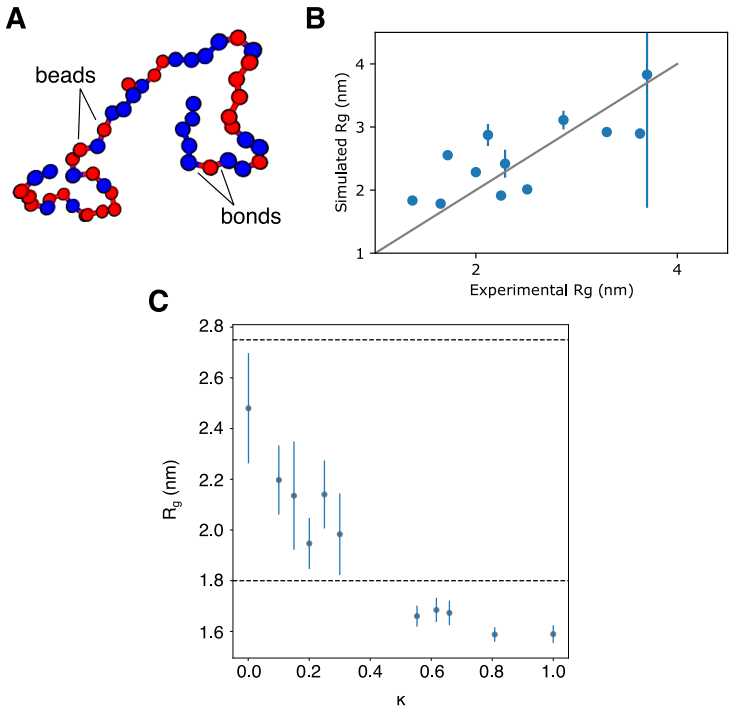

Figure 3.2: Overview of simulation method
A) Still from a movie generated from the coarse-grained simulations used in this work. **B)** Comparison between experimentally determined and simulated R_gs. Experimental values were compiled in [34] and references therein. **C)** Correlation between κ and the simulated R_g for polyampholytes of length N = 50, FCR = 1.0 and $\sigma = 0$. Dashed lines are the upper and lower limit of the same values calculated in [28].

The advantage of using κ to characterize these sequences is that we seek to gain insight into the physical properties of a region from sequence alone. Given the previously-demonstrated correlation between charge patterning and physical properties of polyampholytes, [28, 120, 69] I wanted to investigate whether the same correlation between κ and R_g held for heterogeneous peptide sequences with FCR < 0 (see Figure 3.2C for an demonstration of this relationship with sequences with FCR = 1.0).

To study the structure and dynamics of these sequences, I turned to coarse-grained simulations, previously shown to be effective for studying single chains and large assemblies

of intrinsicially disordered proteins. [33, 34, 19] Coarse-grained simulations greatly facilitate study of these systems, as the large number of particles present in assemblies causes them to be too computationally expensive for fully atomistic simulation. Each chain is modeled as a series of beads attached via bonds that are modeled as spring with a characteristic length and mean extension. Each bead has an amino acid identity, which gives it a size, charge, and hydrophobicity (Figure 3.2A). Solvent interactions are coarse-grained into the effective attraction between beads which consists of both Coulombic and Lennard-Jones type potentials. The potentials contain one free parameter, which was fit such that the R_gs of several experimentally well-characterized IDRs are reproduced with good correlation by the force field (Figure 3.2B). See Methods for the numerical form of the potentials and free parameters. The same set of force fields reproduces results previously generated using Monte-Carlo simulations where the R_g of polyampholytes was shown to correlate with κ (Figure 3.2C). Although the exact values are off by a few Angstrom (dashed lines in Figure 3.2C show upper and lower bound from [28]), the overall trend is extremely similar. Using this simulation method, we can interrogate the physical properties of sequences of interest, investigating such characteristics as the radius of gyration (R_g), preferred internal contacts, relative solvation etc.

Previous work (recapitulated using a different simulation method in this work, see Figure 3.2C) and [28] demonstrated a strong correlation between the charge patterning in a polyelectrolyte, quantified with κ, and its R_g. This finding implies that the physical properties of a region can be predicted as a function of sequence alone. However, it has not been rigorously tested whether such a correlation, demonstrated only for polyampholytes consisting of 100% charged residues (FCR = 1) holds either for random heterogeneous sequences with FCR < 1 or for native IDRs with the same properties.

Demonstration of the predictive power of κ for sequences beyond polyampholytes would provide motivation for using this metric to characterize native sequences. To address this outstanding question, I focused on sequences with low σ, as these most closely mimic the native sequences we are interested in; low-σ IDRs are very common. [28, 74] To determine whether kappa was more universally predictive of R_g, I generated random sequences in a range of FCR and κ values. Non-charged residues were assigned to be alanine. Figure 3.3A shows the mean R_g of a sequence of a given FCR and κ. I found that for both normalization methods, there was a correlation between κ and R_g for a given FCR, although the range of possible R_g values increased with increasing FCR – that is to say that highly charged, low-κ

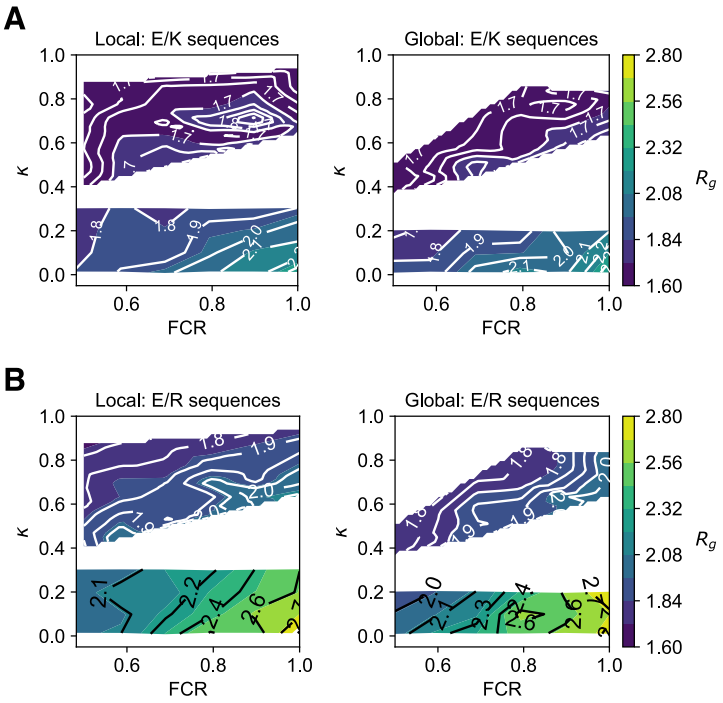

Figure 3.3: Relationship between κ and R_g as a function of FCR
A) R_g as a function of FCR and κ for sequences consisting of glutamate, lysine, and alanine – values were generated from simulation of a single sequence; 10-15 sequences were generated and interpolated across the span of FCR and κ values to generate density plots. **B)** Same as A) but for sequences containing glutamate, argenine, and alanine.

sequences are highly extended, while less highly-charged low-κ sequences are somewhat more collapsed. This is similar to what was demonstrated in a non-systematic way previously [28].

I was also interested in whether the choice of charged amino acid affected these findings. Although the two negatively-charged amino acids (D and E) are chemically similar, the two postively-charged species (K and R) are less similar, with R being bulkier and more extended. Moreover, for unknown reasons R is sometimes depleted or enriched relative to K in IDRs ([110, 10] and Figure 3.4D). I examined whether the same correlation between compaction and κ held when the positively charged residue was argenine rather than lysine (Figure 3.3B). This proved to be the case, although generally the sequences have higher R_gs (likely because R is bigger than K). Although this does not explain the preferential use of R or K in some sequences, it suggests that such preferences are not due to differences in structure that can be captured considering only the effects of charge on structure.

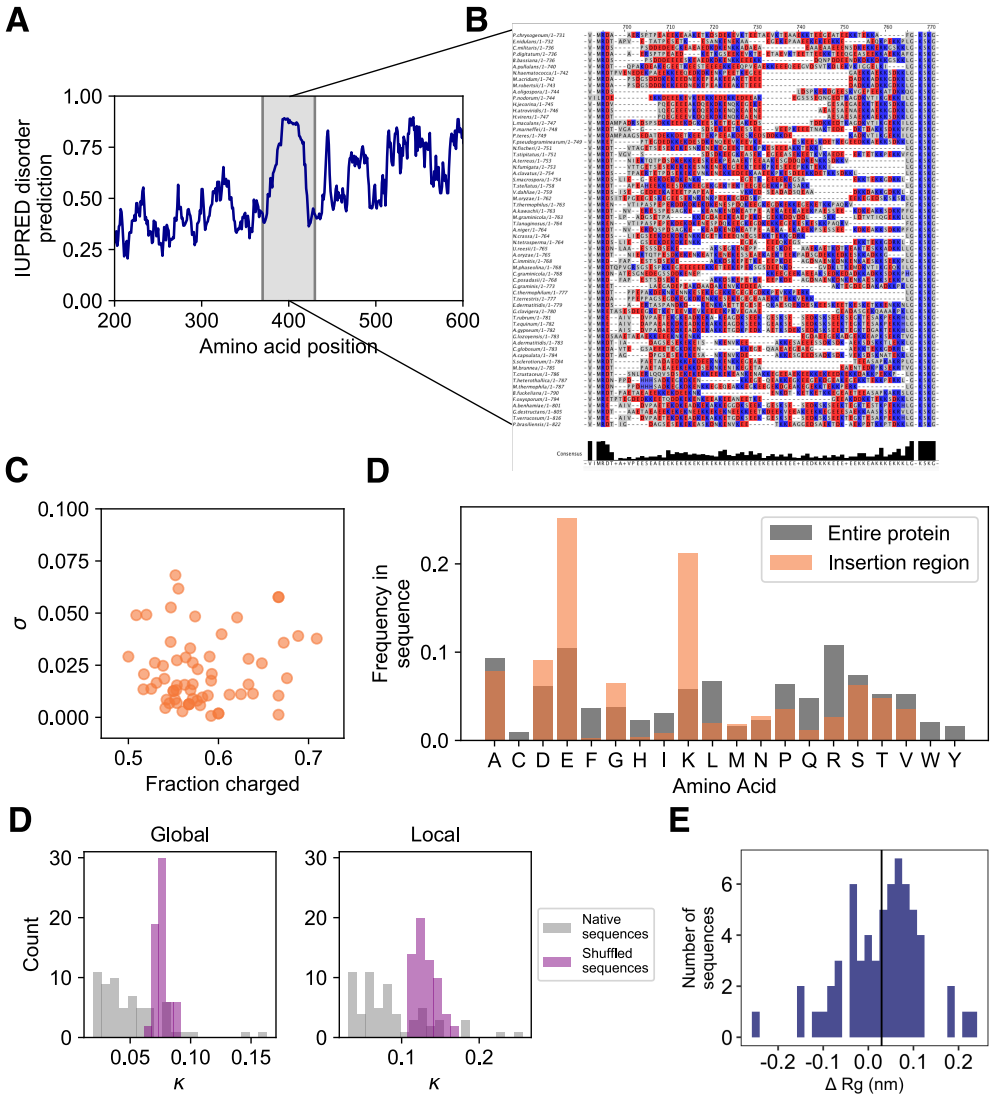

Figure 3.4: Characterization of the poly(A)-binding protein insertion region
A) Disorder prediction [77] for poly(A)-binding protein from *S. macrospora*. The fungal insertion region is highlighted. **B)** Alignment of analyzed sequences. Alignment was generated using the surrounding highly-conserved regions. Positively and negatively charged amino acids are shown in blue and red respectively; all other amino acids are in gray and gaps are in white. Conservation and consensus sequence are shown in black at the bottom. **C)** Distribution of σ and FCR values for the fungal insertion regions; each dot represents a different species. **D)** Average amino acids usage in the fungal insertion region (orange) versus the entire Pab1 sequence (grey). **E)** Distribution of κ values in the native fungal insertion regions (gray). Purple bars represent the distribution of mean κ values obtained after shuffling each sequence 1000 times, recalculating κ after each shuffle. **F)** Distribution of the difference in R_g between shuffled and native fungal insertion region sequences. The mean is indicated with a black vertical line.

κ is non-randomly distributed in naturally-occurring IDRs

The results thus far have shown that κ can be a meaningful metric to quantify charge patterning in sequences with varying FCR (heterogeneous polyampholytes) and, at least for

sequences composed of charged residues and alanine, is predictive of physical properties of the sequence (specifically R_g). Having established this, I turned my attention to a model system in naturally-occurring sequences that could be used to further explore the connections between κ and the physical properties of sequences.

I chose to start by characterizing a region of poly(A)-binding protein (PABP in humans, Pab1 in yeast), a conserved eukaryotic protein that has important roles in translation, [73] RNA interactions, [14] and phase-separates in response to temperature stress. [103] Pab1 is characterized by several highly-conserved RRMs, followed by a disordered proline - rich domain, and a C-terminal domain. In certain fungal species, Pab1 also contains an region we have named the "fungal insertion region", located in the middle of the last RRM. An alignment of this region is shown in Figure 3.4B. It is predicted to be highly disordered (Figure 3.4A), and contains a very high fraction of charged residues (Figure 3.4B, C). Compared to the rest of the protein, this region has a much greater frequency of charged amino acids, with the surprising exception of arginine (R, see discussion of this topic above), which is extremely depleted in these sequences (Figure 3.4D). The sequences also show strong depletion of bulky aromatic amino acids.

Beyond matching the characteristics of the IDRs of interest set out in the intro, this region is of particular interest because of its pattern of evolution. The region is maintained in some species but not in others. [103] Its preservation leads us to the hypothesis that its presence is selected for, and yet, similar to many other IDRs, [74] the length and sequence identity between the regions is highly variable. Given that by some estimates, one third of eukaryotic proteins are thought to contain IDRs, [67] the ability to track evolutionary patterns in such sequences is highly desirable. However, the lack of conservation at the sequence level presents challenges for evolutionary analyses, [141] which depends on good alignment between homologs. The demonstration of a conserved feature of the region that is a function of the primary sequence but does not depend on precise correlation between homologs would provide additional tools for identifying and analyzing evolutionary patterns of IDRs.

To address whether the distribution of κ values in this region was non-random, I first calculated κ values for the insertion regions (approximately 70 sequences) using both the local and global normalizations. The distribution of values, shown in grey in Figure 3.4E, are both highly skewed towards low-kappa values, and have a very different shape than the expected distribution of κ values for sequences with this composition (purple distributions in

Figure 3.1B, top). To further test whether the distributions were nonrandom, I shuffled each sequence 1000 times, and reported the mean κ value for each. The distribution of kappas for shuffled sequences is shown in Figure 3.4E in purple. The mean and variance of the native and shuffled sequences are remarkably different, lending evidence to the idea that κ is a selected for in this region, and specifically that low κ values are preferred.

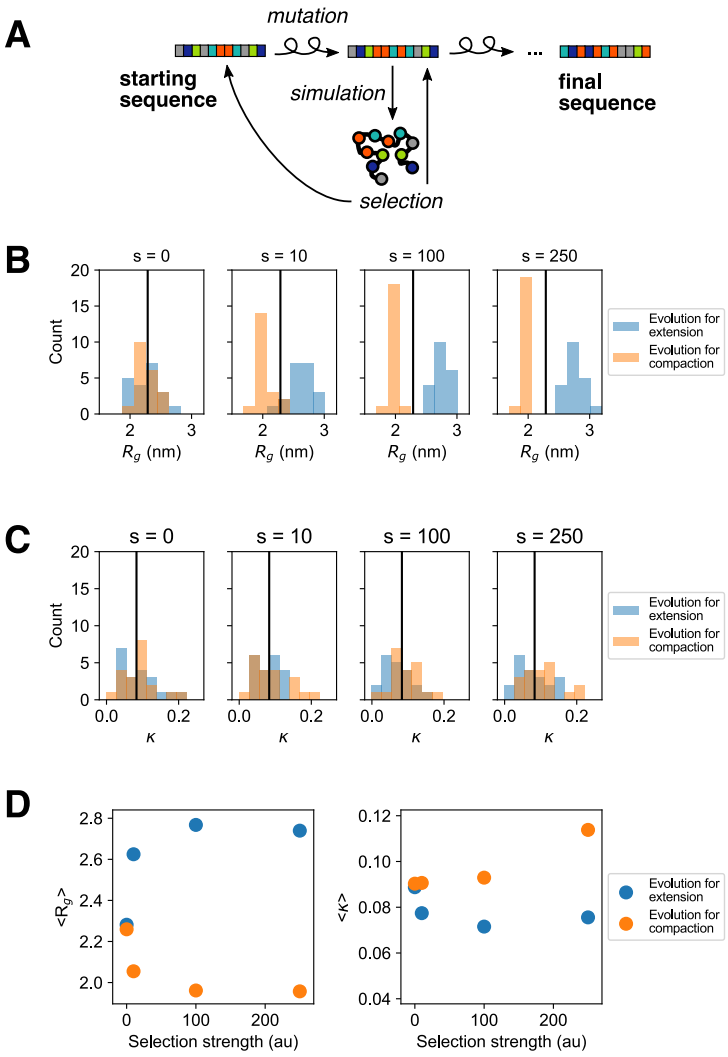

Figure 3.5: A toy model for evolution produces apparent selection for κ
A) Illustration of evolutionary algorithm **B)** Distribution of R_g values after artificial evolution for increasing (blue) or decreasing (orange) R_g. All runs were initialized with the same starting sequence. 3 mutations were generated between each simulation step, and 120 steps total were run for each sequence. **C)** Distribution of κ values for the same sequences as shown in B) **D)** Average R_g and κ shown as a function of the selection strength in the evolutionary algorithm.

κ not only describes the charge patterning of a sequence; it can also predict physical properties (Figure 3.3). Does the apparent selection for κ reflect selection for R_g? To answer

this question, I simulated each naturally-occurring sequence, calculated the R_g, and then shuffled the sequence and performed the same simulation. The distribution of the change in R_g is shown in Figure 3.4F. My prediction is that selection for κs closer to zero would yield sequences that are on average more extended than expected for their composition. Indeed, when sequences are shuffled the resulting difference in the differences between the native and shuffled R_gs is greater than zero, indicating that the native R_gs are more extended (Figure 3.4F) with statistical significance p=0.010 (paired t-test).

Figure 3.4 demonstrates preliminary evidence of selection for a physical property (R_gs) that can yield apparent selection on a sequence metric; however I wanted to demonstrate this principle in a more direct fashion. I created a toy model of evolution, illustrated in Figure 3.5A. Briefly, a polyampholytic sequence characterized by set of κ, FCR, and σ values is first generated at random. I focused on sequences that are similar to the native sequences – that is, that have an intermediate R_gs value (given the possible range, see Figure 3.2C), a κ value within the native range (Figure 3.4D), and low charge asymmetry. I started with sequences that have an FCR of 1.0 as κ has the largest effect on R_g when FCR = 1 (Figure 3.2A); note that under this condition the local and global κ metrics are identical (Figure 3.1B). Once the initial sequence is generated, it is is mutated at random – if a position with a positively charged residue is selected that residue is flipped to a negative, and vice-versa. I then simulated the resulting sequence to estimate the R_gs; the move in sequence space is accepted or rejected according to the following rule:

Selection for smaller R_g:

$$p_{acc} = \min\left[1, \exp(-s\beta(R_g^{new} - R_g^{old}))\right] \tag{3.5}$$

Selection for larger R_g:

$$p_{acc} = \min\left[1, \exp\left(-s\beta\left(R_g^{old} - R_g^{new}\right)\right)\right] \tag{3.6}$$

$s\beta$, the effective selection strength, sets the stringency of the selection.

The results of 20 independent replicates of the evolutionary process for selection for extended or compacted R_g with a variety of selection strengths is shown in Figure 3.5B-D. For both selection for more compact or more extended sequences, the final distribution of R_g values for the evolved sequences was skewed in the expected direction, indicating that

the evolutionary algorithm was successful in moving sequences through R_g space by making changes to the primary sequence (Figure 3.5B). When the κ values after evolution were examined, the resulting distribution skewed more towards high κ values when compaction was selected for, and towards low κ values when extension was selected for, in line with the correlation between κ and R_g (Figure 3.5C). This result demonstrates that when the only property that is evolutionary relevant is the compaction of an IDR, this pressure can generate a bias in the distribution of a parameter that quantifies the patterning of residues within the sequence.

How does this effect depend on the strength of selection? Although the metric for selection used here, s, is arbitrary, we can still comment on how the relative effect size of the evolutionary process scales with selection strength. By plotting the selection strength versus the average final R_g and κ values (Figure 3.5D), it can be seen that the relative effect on R_g saturates at selection strength of around 100. The effect on κ is similar, although when compaction is selected for there is a large increase in the mean κ value between s=100 and s=250; this result should be verified with additional evolution simulations starting from different starting sequences.

Much remains to be done even with this basic analysis; for example, we have not yet carefully determined whether the evolutionary process is equilibrated (which would involve tracking the acceptance probability of moves over time). The same process must also be repeated for other starting sequences to ensure it is not an artifact of the starting conditions. However, these initial findings are promising, and open the door to exciting additional analyses and simulations, outlined in the following section. The results presented here demonstrate that evolutionary pressure on a physical parameter of an IDR can yield predictable selection patterns in a metric calculated using only the primary sequence.

3.3 Conclusions and Future Directions

Here we have demonstrated that κ is a promising metric to characterize naturally-occurring polyampholytic sequences; it can be normalized several ways, and the different normalization schemes are likely more appropriate for certain types of sequences over others. This can be tested by applying the same analyses shown here to a variety of other native, highly-charged IDRs and comparing the distributions. We have demonstrated that for the region of Pab1 considered in this work, that the naturally-occuring sequences show apparent selection for

certain values of kappa; we hypothesize that this is due to underlying selection pressures on the physical properties of the sequences, namely the R_g.

It has been noted that highly-charged sequences are a common type of IDR in all parts of the tree of life; [28] it is easy to imagine how the analyses here could be repeated for IDRs outside of the fungal clade, with the exciting possibility of detecting selection signatures in such sequences for other proteins. Connection to widely-accepted models of evolutionary processes may also be possible; if so, this would allow mappings between the arbitrary selection strength we have defined here, and other measures of fitness tied to more explicit models of evolution. [111]

An exciting future direction lies in the prediction of higher-order physical properties from sequence. Intrinsically-disordered proteins such as those considered here have generated much attention over the last decade for their ability to self-associate and form distinct phases within cells. [10, 81, 74] For pure polyampholytes, self-association behavior has been linked to both charge patterning and the Rg of the individual chains. [69] These connections may allow detection of evolutionary selection for higher-order properties like critical temperature or concentration. Importantly, because the simulations employed here are coarse-grained, they can be effectively scaled up to simulate this higher-order behavior.

3.4 Methods

Coarse-Grained simulation All simulations were performed using OpenMM, [37] an open-source molecular dynamics API implemented using a wrapper developed by the Mirny lab. [40] Amino acids are modeled as beads connected by harmonic bonds, and each bead has properties (charge, size, and hydrophobicity) that are a function of its amino acid identity. The sizes and hydrophobicities of each bead are assigned as in [34]. Custom force-fields were developed, with the following functional forms:
Van der Waals:

$$\Theta(r) = \begin{cases} \Theta_{LJ} + (1-\lambda)\,\epsilon, & \text{if } r \leq 2^{1/6}\sigma \\ \lambda\Theta_{LJ}, & otherwise \end{cases} \tag{3.7}$$

where

$$\Theta_{LJ} = 4\epsilon\left[\left(\frac{\sigma}{r}^{12}\right) - \left(\frac{\sigma}{r}^{6}\right)\right] \tag{3.8}$$

here λ is a tuning parameter; the value that maximized the correlation with experimentally determined R_g values was 0.55.

Coulombic forces are:

$$E_{ij} = \frac{q_i q_j}{D r \epsilon_0} \exp\left(\frac{-r}{\kappa_D}\right) \tag{3.9}$$

where ϵ_0 is the permittivity of free space, κ_D is the Debye screening length, D is the dielectric constant (set to 80 for water) r is the distance between particles, and q_i, q_j are the charges on the ith and jth particles respectively. Note that although not considered here, κ_D can be tuned to simulate changes in ionic strength, and the charges on beads can are hand-assigned such that changes in charge, possibly due to solvent pH changes, could be simulated in the future.

Sequence R_gs were calculated according to the definition of R_g:

$$R_g^2 = \frac{1}{N} \sum_{k=1}^{N} \left(r_i - r_j\right)^2 \tag{3.10}$$

All R_g values reported were obtained in the following way: chains of length 68 (Figure 3.3, the median length of the naturally-occuring Pab1 sequences, 50, Figure 3.2A and 3.5, or the length of the naturally-occurring sequence, Figure 3.4E) were run for 2000 steps of burn-in, followed by 100,000 steps of production. The R_g was calculated every 100 steps and averaged to give the final R_g. The mean of 5 replicates was reported as the final value.

Analysis of Pab1 sequences

The Pab1 orthologs and alignment were obtained from previously published work. [103] Only Pab1 orthologs with the insertion region were analyzed. Alignment was performed with MUSCLE. [38] Analysis of sequences and simulation results was performed with custom scripts written in the python programming language.

CHAPTER 4
ADDITIONAL APPENDICES

4.1 Appendix III: Detection of proteins with constitutive chaperone-dependent solubility requirements

Introduction

Thus far, I have largely considered the role that Heat Shock Factor I (Hsf1) plays in the adaptive stress response. However in yeast [116, 118, 119] and in some human cancers [26, 76], Hsf1 is essential for survival. For decades, the molecular requirements for Hsf1 essentiality remained mysterious, but recently chemical genetic techniques have allowed researchers to define the essential Hsf1 regulon as well as the genes that absolutely require Hsf1 for induction during both heat shock and constitutive growth. [116, 98] Using chemical genetic tools, Solis *et al.* [116] showed that the production of two types of molecular chaperones, Hsp70 and Hsp90, was sufficient to rescue the lethality of Hsf1 knockout, demonstrating that a basal level of specific chaperone activity was required for cellular survival.

This work raises a fascinating question: why does the cell require basal levels of these two chaperone subtypes? We hypothesized that there is a subset of proteins, some of which are likely to be essential, that require chaperone activity to fold or function. We reasoned that if this were the case, following depletion of Hsf1 and the essential chaperones it regulates, we could detect such proteins in an aggregated state by mass spectrometry using the same techniques we previously used to detect proteins that aggregate during heat shock. [131] I performed this analysis on Hsf1-depleted cells and cells that had been treated identically but supplemented with Hsf1 via expression from a plasmid (previously shown to rescue viability [116]). I detected a set of proteins, shown in Table 4.2, that depend on the essential activity of Hsf1 for solubility. Although the group is diverse, it contains many essential genes (13/37), and is enriched for proteins found in the proteosome or that are classified as peptidases. Further analysis of the sequence and structure of these targets, as well as verification of aggregation through orthogonal means such as fluorescence microscopy or Western Blot, will establish the validity of these results and shed light on why they require chaperones for their solubility and potentially function.

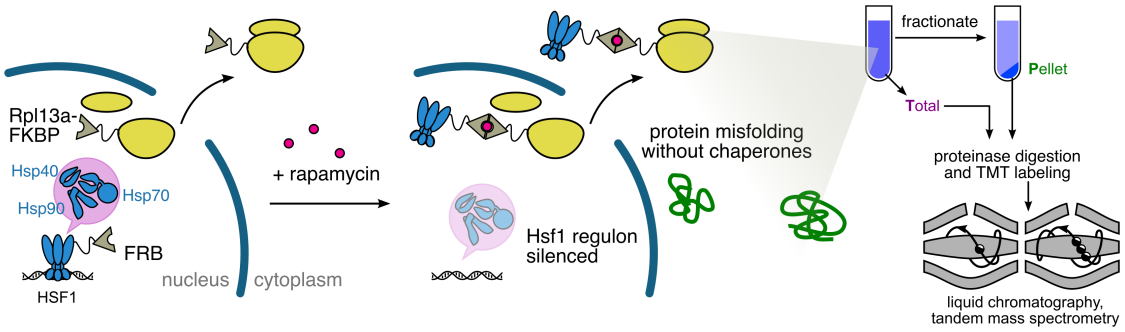

Figure 4.1: Schematic of Hsf1-AA mass spec screen

Using the same chemical genetic strategy employed by Solis et al., I removed Hsf1 from the nucleus to downregulate chaperones and collected cells after 1 and 6 hours of depletion (Figure 4.1). As a control for the effects of the genetic and chemical perturbations, I also collected equivalent samples where exogenous Hsf1 was expressed from a plasmid. For internal normalization, I collected samples where the equivalent treatment had been performed with carrier alone; below all protein abundances are expressed as a ratio to this value, allowing both correction for the effect of treatment with the carrier (DMSO), and internal normalization. On each sample I performed cryogenic lysis, separated soluble from insoluble proteins by ultracentrifugation, extracted proteinaceous material, and submitted total and insoluble fractions for TMT mass spectrometry analysis; this process is illustrated in Figure 4.1.

Using stringent filtering techniques for the quantified data, I was able to confidently (see Methods for details) detect over 200 proteins, with a median of 10 peptides per protein detected (Figure 4.2A). Due to limitations in TMT mass spec, only 10 samples could be run with simultaneous quantification; those samples for which biological replicates were analyzed are shown in Figure 4.2 B-D. The correlation between the Hsf1 depletion samples, both total and pellet, is quite good (Spearman correlation > 0.85). The correlation between the pellet samples from cells rescued from Hsf1 depletion by expression of exogenous Hsf1 had much greater variation; as neither sample could be identified as being of higher quality, the two replicates were averaged for all downstream analysis.

For further quality control, I checked for the reduction in abundance of Hsf1-dependent proteins from the total fraction in cells where Hsf1 had been depleted; these same targets should not decrease in those cells expressing the Hsf1-rescue plasmid. Although depletion of

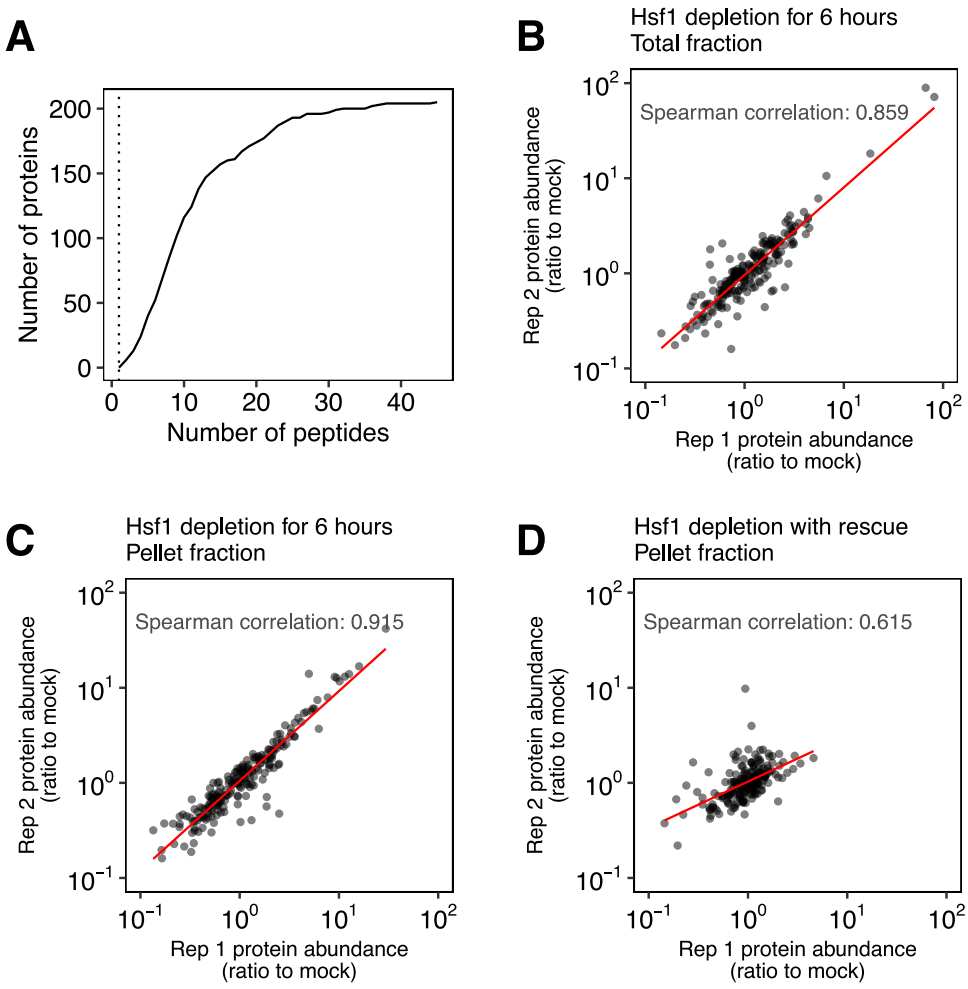

Figure 4.2: Quality control of Hsf1-AA mass spec screen
A) Empirical CDF of the number of proteins confidently detected and the number of peptides detected per protein. **B-D)** Correlation between biological replicates in the indicated samples; total and pellet fractions for Hsf1 depletion (6 hours), and pellet fractions for Hsf1 depletion with rescue (6 hours).

Hsf1 likely does not cause an increase in the degradation of its targets, without ongoing transcription and translation of the genes, protein turnover will eventually lead to a depression of the entire Hsf1 regulon. A time course of abundance for several key Hsf1-exclusive targets are shown in Figure 4.3A; SSA1 and SSA2 are members of the Hsp70 subfamily, and HSC82 is a member of the Hsp90 subfamily. While none of these proteins are individually essential, their activity can rescue viability of Hsf1 depleted or null cells. [116] Figure 4.3A clearly shows, as expected, that these targets decrease in abundance in cells where Hsf1 is depleted from the nucleus; in rescued cells this depletion is to a much lesser extent or nonexistent.

To analyze the effect of Hsf1 depletion proteome-wide, I compared the abundance of each

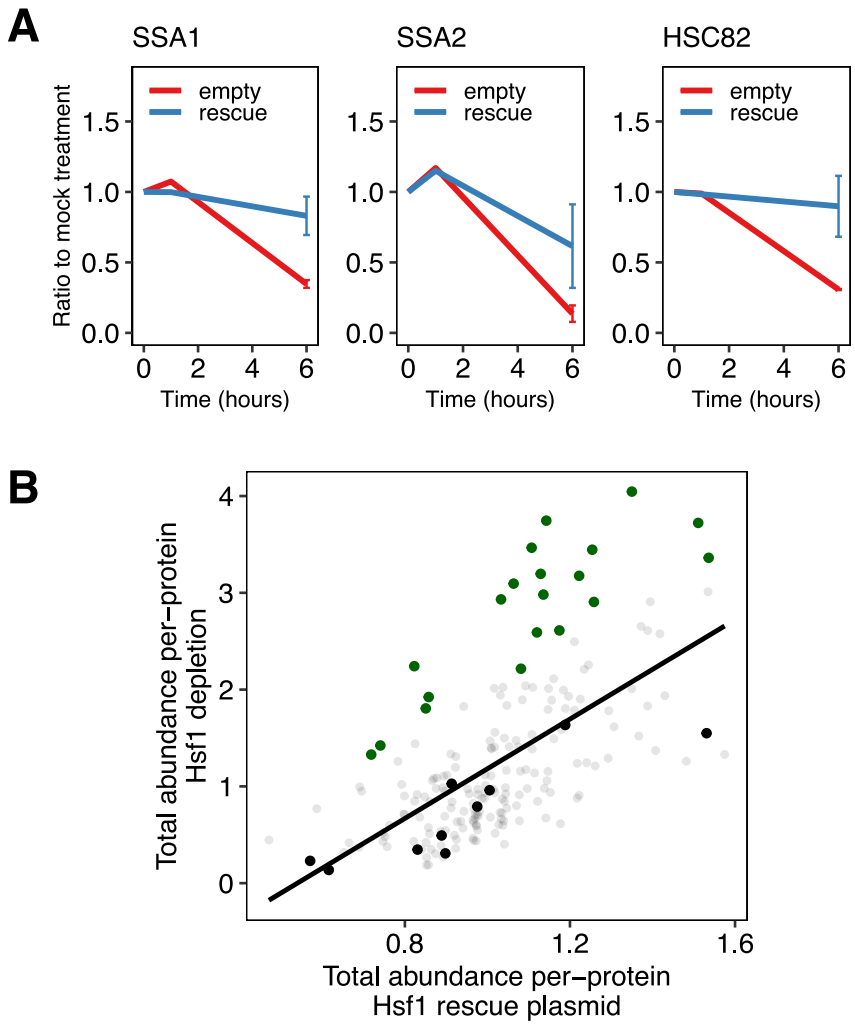

Figure 4.3: Proteins are both induced and downregulated upon nuclear depletion of Hsf1
A) Abundance of several Hsf1-dependent genes as a function of time after Hsf1 depletion ("empty", red) or Hsf1 depletion with exogenous Hsf1 to rescue ("rescue", blue). Each dot is the mean of two replicates; errorbars are the standard deviation. **B)** Abundance of all detected proteins in Hsf1 depleted and rescued cells. Black dots show Hsf1-dependent genes as identified in [116, 98]. Green dots are proteins upregulated in response to Hsf1 depletion. Black line is the best-fit linear least squares correlation between the two samples.

protein between the Hsf1 depleted and rescued samples after 6 hours of treatment (where differences are likely to be more pronounced). Of the genes the absolutely require Hsf1 for their expression detected in our dataset, nearly all were lower in abundance in Hsf1-depleted cells, as demonstrated in Figure 4.3B (black dots). Interestingly, there were many targets that *increased* in abundance in response to loss of Hsf1. These proteins, shown as green dots in Figure 4.3B, are listed in Table 4.1. The list is enriched (13/16) for mitochondrial (mt) membrane proteins, including a mt-associated chaperone (CDC48).

	Standard Name	Systematic Name
1	CDC48	YDL126C
2	COQ1	YBR003W
3	COX13	YGL191W
4	CUE5	YOR042W
5	CYT1	YOR065W
6	DEP1	YAL013W
7	ENO1	YGR254W
8	FIS1	YIL065C
9	GLN1	YPR035W
10	GPM1	YKL152C
11	MRP4	YHL004W
12	MSC1	YML128C
13	PHB1	YGR132C
14	PHB2	YGR231C
15	PIL1	YGR086C
16	PRB1	YEL060C
17	RSP5	YER125W
18	SBP1	YHL034C
19	VPH1	YOR270C
20	YDL124W	YDL124W

Table 4.1: Detected proteins that are upregulated after Hsf1 is depleted from the nucleus

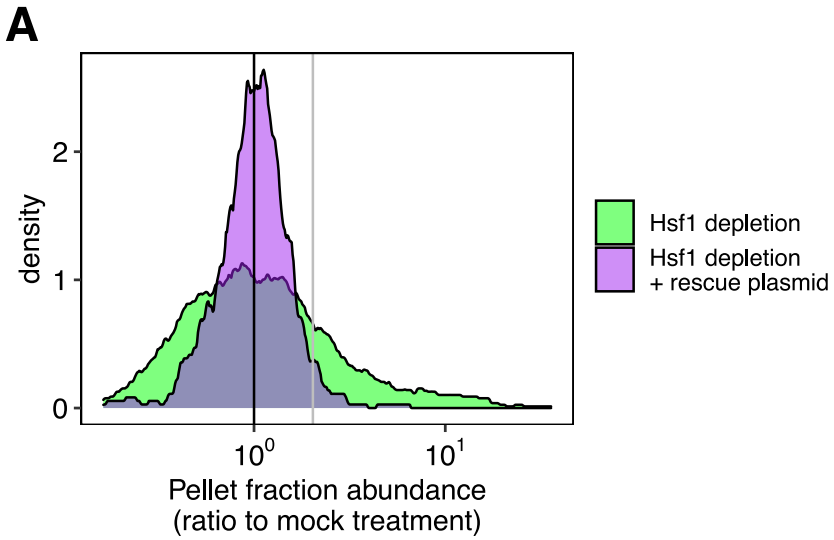

Figure 4.4: Proteome-wide distribution of abundance in the pellet fraction
A) Normalized abundance in the pellet fraction for all proteins detected. Green represents the distribution for cells where Hsf1 is depleted; purple are data from cells where Hsf1 is depleted but rescued via exogenous expression.

Finally, I assessed the abundance of proteins in the pellet fraction to determine which species required the activity of Hsf1 in order to remain soluble. I plotted the distribution of abundances in the pellet fraction for proteins extracted from Hsf1-depleted cells (Figure 4.4A, green), and Hsf1-rescued cells (Figure 4.4A, purple), Using the data from the rescue sample to establish a cutoff, I identified proteins that had normalized abundance in the pellet fraction greater that two standard deviations above the mean (shown as a gray line in Figure 4.4). These proteins, and their associated pellet enrichment (i.e. the ratio of abundance in Hsf1-depleted vs. rescued samples) are listed in Table 4.2.

This list of proteins is enriched (10/37) for proteasome and peptidase components, and all are cytoplasmic proteins. There is no significant enrichment for GO terms for biological process or for function. However, a large portion of them are essential (13/37), perhaps providing an explanation for why the depletion of chaperones eventually leads to cell death; if these proteins are inactive in their aggregated state, this could phenocopy a null mutant, leading to loss of viability.

Interestingly, although anecdotal, the group includes several glycolytic enzymes (GLN1, HSP48, PRE7) [86, 97, 90] which have been shown to aggregate during periods of glucose starvation. In particular Gln1 forms filaments during glucose starvation in a mechanism that is mediated through a drop intracellular pH. [97] Although filamentation has not been shown to require chaperone activity directly, it may be possible that interaction with chaperones helps to regulate when filamentation occurs and the integrity of the filaments.

Conclusions and Future Directions

In this work I identified a group of proteins that require Hsf1 activity to retain solubility; when Hsf1 is depleted from the nucleus and the protein products of Hsf1 genes, mainly molecular chaperones, are downregulated, these proteins become insoluble. There are many interesting questions that remain to be answered about these proteins – whether they are functional in their aggregated state, whether the loss of one or more of the group is responsible for the cell death that occurs with loss of Hsf1, and which genes in the Hsf1 regulon are responsible for maintaining the solubility of these proteins, among others.

A way to begin answering these questions lies in deeper analysis of the targets – identification of a common fold or structural motif unique to this group, or even compositional features, could help explain why these proteins aggregate without chaperone activity. It

would also be useful to analyze the sequences to look for predicting binding sites of molecular chaperones, particularly Hsp70 and Hsp90. It is also important to validate these targets through a means orthogonal to sedimentation and mass spec – for example, tagging the genomic copies of the proteins with a fluorophore and looking for the formation of foci or aggregates after Hsf1 depletion.

This work extends our understanding of the essential function of Hsf1, an ancient transcription factor with a unique role in environmental response, protection of proteostasis, and human disease.

Methods

Sample preparation

Strain VDY2578 [116] (W303 MATα TPK1/2/3/as TOR1-1(S1972I) fpr1Δ::NATMX PrTEF2-mKATE2URA3 4xHSE-Pr$_{CYC1}$-EmGFP RPL13A-2xFKBP12TRP1 HSF1-FRBHIS3MX6RPB3-FLAGKANMX), containing either plasmid pRS412 (empty) or pVD565 (pRS412 + HSF1) were grown at 30°C; the OD600 was below 0.1 for at least 12 hours, and cells were treated when they reached OD600 0.1. Treatment consisted of addition of 1000X rapamycin in DMSO to a final concentration of 1μM, or addition of an equal volume of carrier only. The addition of rapamycin induces dimerization of FRB and FKBP, depleting Hsf1 from the nucleus. Cells were allowed to continue to grow at 30°C with 250 rpm shaking, and samples were collected and flash-frozen 1 and 6 hours post treatment.

Flash-frozen samples were lysed and centrifuged as previously described. [131] Chloroform-methanol extraction was performed on all samples, and precipitated proteins were submitted to the mass spectrometry facility at the FAS Center for Systems Biology at Harvard University for proteomic mass spectrometry using TMT labeling. Protocols for both procedures can be found at http://drummondlab.org.

Data Analysis

The mass spectrometry data was quantified and assigned to the yeast proteome using MaxQuant. Proteins for which more than one peptide was detected in all samples were considered for downstream analysis.

The cutoff for pellet enrichment was determined by plotting the distribution of protein abundance in the insoluble fraction in cells where endogenous Hsf1 was depleted, but cells

were rescued by exogenous Hsf1 expression. Values greater than two standard deviations were considered as being enriched in the pellet, and these values for cells where Hsf1 had been depleted without rescue, and which were also not found enriched in the pellet fraction of rescued cells were considered to be pellet enriched, and are found in Table 4.2.

	Gene name	Systematic Name	Pellet Enrichment with Hsf1 depletion
1	YDL124W	YDL124W	16.39
2	ENO1	YGR254W	12.27
3	PGM2	YMR105C	11.09
4	GLO1	YML004C	10.91
5	MSC1	YML128C	7.81
6	PRE6	YOL038W	6.76
7	DEP1	YAL013W	5.84
8	PRE7	YBL041W	5.74
9	MCR1	YKL150W	5.48
10	ACO1	YLR304C	5.14
11	PRB1	YEL060C	4.92
12	GLN1	YPR035W	4.35
13	MRP4	YHL004W	4.28
14	RPT5	YOR117W	3.95
15	HSP104	YLL026W	3.51
16	ARO8	YGL202W	3.45
17	VMA2	YBR127C	3.32
18	RPT3	YDR394W	3.32
19	MLC1	YGL106W	3.32
20	RPT6	YGL048C	3.11
21	SEC53	YFL045C	2.92
22	ABP1	YCR088W	2.82
23	MGE1	YOR232W	2.75
24	QCR7	YDR529C	2.74
25	CUE5	YOR042W	2.68
26	ILV5	YLR355C	2.59
27	FMP52	YER004W	2.55
28	FIS1	YIL065C	2.53
29	PHB1	YGR132C	2.44
30	AYR1	YIL124W	2.32
31	RPN8	YOR261C	2.26
32	MRPS9	YBR146W	2.17
33	EIS1	YMR031C	2.15
34	ENO2	YHR174W	2.11
35	FUR1	YHR128W	2.10
36	ARC1	YGL105W	2.04
37	YML6	YML025C	2.04

Table 4.2: Detected proteins that require Hsf1 activity for constitutive solubility

CHAPTER 5
CONCLUSIONS AND FUTURE DIRECTIONS

My work here has explored the role of physical signals in the cellular response to changing environments. Specifically, I defined the role that long-observed intracellular acidification plays in triggering the response and connections to specific known stress-responsive molecular mechanisms; I have also explored the role that charge patterning plays in determining the molecular properties and evolutionary trajectory of intrinsically-disordered regions as a first step towards understanding the effect that intracellular pH change has on the molecules within cells.

I am particularly excited about the fascinating questions that this work has raised. From a cell-biological point of view: how are cells sensitive to pH changes? What cellular species activates Hsf1 in a pH-specific manner, and how is this accomplished at the molecular level? From a molecular point of view, we might ask how does pH change affect the structure, localization, and higher-order associations of proteins, and is pH-sensitivity harnessed on the molecular level to create pH-sensors or integrate pH-sensing into cellular pathways?

It is often easier to think about systems at rest or at equilibrium, but cells live in dynamic environments, constantly facing change in their surroundings. This situation is only exacerbated when the cells in question are sessile – that is they cannot escape the changes, they must ride them out. In this sense cells can be incredibly resilient, and it is our job to understand this resiliency – not only is it an excellent tool for understanding regulation, but it also teaches us about the fundamental limits of the stuff we're made of. When the temperature increases, a dizzying array of other changes come along for the ride – protons increase in concentration by an order of magnitude, the cytoplasm becomes far more viscous and diffusion slows, entire organelles that were once dynamic freeze in place, membranes of all sorts grow vast holes and leakiness between compartments increases, making once-sharp gradients fuzzy or non-existent, to name a few. These changes, being physical in their nature, have existed as long as the cells have existed, and are thus a 'natural' part of cellular life.

The fact and extent of these changes is often overlooked, as they are happening at a scale at which we cannot watch them unfold with our own eyes, yet we have ample physical evidence for all of them. Imagining all these changes happening simultaneously paints a dramatic picture of the cellular interior; to understand how the cell responds to these changes,

not just as stressors but indeed as cues and sources of information about its environment, was the goal of my work. I considered it both at the macro- and regulatory scale, and at the scale of individual proteins. By continuing to keep such changes in mind, I hope to broaden the view of what happens when cells get too hot – by understanding these processes in 'simpler' organisms, we can begin to understand both fundamental aspects of biology, and the origin of behaviors we observe in our own cells and tissues – the pathways considered here came from the same raw evolutionary material from which our own cells evolved regulation.

REFERENCES

[1] G. , U. Pilatus, and L. Rensing. Similar Dose Response of Heat Shock Protein Synthesis and Intracellular pH Change in Yeast. *Exp Cell Res*, pages 252–256, 1985.

[2] Benjamin Albert, Isabelle C Kos-Braun, Anthony K Henras, Christophe Dez, Maria Paula Rueda, Xu Zhang, Olivier Gadal, Martin Kos, and David Shore. A ribosome assembly stress response regulates transcription to maintain proteome homeostasis. *Elife*, 8:e45002, May 2019.

[3] Mark P Ashe, Susan K De Long, and Alan B Sachs. Glucose depletion rapidly inhibits translation initiation in yeast. *Mol. Biol. Cell*, 11:833–848, 2000.

[4] Tanja Bagar, Kirsten Altenbach, Nick D. Read, and Mojca Benčina. Live-cell imaging and measurement of intracellular pH in filamentous fungi using a genetically encoded ratiometric probe. *Eukaryot. Cell*, 8(5):703–712, 2009.

[5] R Baler. Heat shock gene regulation by nascent polypeptides and denatured proteins: hsp70 as a potential autoregulatory factor. *J. Cell Biol.*, 117(6):1151–1159, 1992.

[6] C. A. Barnes, R. A. Singer, and G. C. Johnston. Production of heat shock protein is independent of cell cycle blockage in the yeast Saccharomyces cerevisiae. *J. Bacteriol.*, 169(12):5622–5625, 1987.

[7] R P Beckmann, M Lovett, and W J Welch. Examining the function and regulation of hsp 70 in cells subjected to metabolic stress. *J. Cell Biol.*, 117(6):1137–1150, June 1992.

[8] Tatiana Benaglia, Didier Chauveau, David R Hunter, and Derek S Young. mixtools: An R package for analyzing mixture models. *J. Stat. Softw.*, 32(6):1–29, 2009.

[9] Walter F. Boron. Regulation of intracellular pH. *Adv. Physiol. Educ.*, 28:160–179, 2004.

[10] Clifford P. Brangwynne, Peter Tompa, and Rohit V. Pappu. Polymer physics of intracellular phase transitions. *Nat. Phys.*, 11(11):899–904, 2015.

[11] Nicolas L Bray, Harold Pimentel, Páll Melsted, and Lior Pachter. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.*, 34(5):525–527, 2016.

[12] Catherine M. Bright and David Ellis. Intracellular pH changes induced by hypoxia and anoxia in isolated sheep heart Purkinje fibres. *Exp. Physiol.*, 77:165–175, 1992.

[13] Alistair J P Brown, Michelle D Leach, and Susan Nicholls. The relevance of heat shock regulation in fungal pathogens of humans. *Virulence*, 1(4):330–332, 2010.

[14] Christine E. Brown and Alan B. Sachs. Poly(A) Tail Length Control in Saccharomyces cerevisiae Occurs by Message-Specific Deadenylation. *Mol. Cell. Biol.*, 18(11):6548–6559, 1998.

[15] J. Ross Buchan and Roy Parker. Eukaryotic Stress Granules: The Ins and Outs of Translation. *Mol. Cell*, 36(6):932–941, 2009.

[16] Laia Castells-Roca, José García-Martínez, Joaquín Moreno, Enrique Herrero, Gemma Bellí, and José E. Pérez-Ortín. Heat Shock Response in Yeast Involves Changes in Both Transcription Rates and mRNA Stabilities. *PLoS One*, 6(2):e17272, 2011.

[17] H C Causton, B Ren, S S Koh, C T Harbison, E Kanin, E G Jennings, T I Lee, H L True, E S Lander, and R A Young. Remodeling of yeast genome expression in response to environmental changes. *Mol Biol Cell*, 12(2):323–337, 2001.

[18] Valeria Cherkasov, Sarah Hofmann, Silke Druffel-Augustin, Axel Mogk, Jens Tyedmers, Georg Stoecklin, and Bernd Bukau. Coordination of translational control and protein homeostasis during severe heat stress. *Curr. Biol.*, 23(24):2452–2462, 2013.

[19] Jeong-Mo Choi, Furquan Dar, and Rohit V. Pappu. LASSI: A lattice model for simulating phase transitions of multivalent proteins. *bioRxiv*, 2019.

[20] Surabhi Chowdhary, Amoldeep S Kainth, and David S Gross. Heat Shock Protein Genes Undergo Dynamic Alteration in Their Three-Dimensional Structure and Genome Organization in Response to Thermal Stress. *Mol. Cell. Biol.*, 37(24):1–22, 2017.

[21] Surabhi Chowdhary, Amoldeep S Kainth, David Pincus, and David S Gross. Heat Shock Factor 1 Drives Intergenic Association of Its Target Gene Loci Upon Heat Shock. *bioRxiv*, 2018.

[22] Richard; Christen, Robert W; Schackmann, and Bennett M; Shapiro. Elevation of the intracellular pH activates respiration and motility of sperm of the sea urchin, Strongylocentrotus purpuratus. *J Biol Chem*, 257(24):14881–14890, 1982.

[23] Joaquin F Christiaens, Luis M Franco, Tanne L Cools, Luc De Meester, Jan Michiels, Tom Wenseleers, Bassem A Hassan, Emre Yaksi, and Kevin J Verstrepen. The fungal aroma gene ATF1 promotes dispersal of yeast cells through insect vectors. *Cell Rep.*, 9(2):425–432, 2014.

[24] P. J. Coote, M. B. Cole, and M. V. Jones. Induction of increased thermotolerance in Saccharomyces cerevisiae may be triggered by a mechanism involving intracellular pH. *J. Gen. Microbiol.*, 137(7):1701–1708, 1991.

[25] P J Coote, M V Jones, I J Seymour, D L Rowe, D P Ferdinando, A J McArthur, and M B Cole. Activity of the plasma membrane H(+)-ATPase is a key physiological determinant of thermotolerance in Saccharomyces cerevisiae. *Microbiology*, 140(1):1881–1890, 1994.

[26] Chengkai Dai, Luke Whitesell, Arlin B. Rogers, and Susan Lindquist. Heat Shock Factor 1 Is a Powerful Multifaceted Modifier of Carcinogenesis. *Cell*, 130(6):1005–1018, 2007.

[27] Leonardo Dapporto, Irene Stefanini, Damariz Rivero, Mario Polsinelli, Paolo Capretti, Paolo De Marchi, Roberto Viola, Stefano Turillazzi, and Duccio Cavalieri. Social wasp intestines host the local phenotypic variability of Saccharomyces cerevisiae strains. *Yeast*, 33(7):277–287, 2016.

[28] Rahul K. Das and Rohit V. Pappu. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U. S. A.*, 110(33):13392–13397, 2013.

[29] Carl G. De Boer and Timothy R. Hughes. YeTFaSCo: A database of evaluated yeast transcription factor sequence specificities. *Nucleic Acids Res.*, 40:D169–D179, 2012.

[30] Reinhard Dechant, Matteo Binda, Sung Sik Lee, Serge Pelet, Joris Winderickx, and Matthias Peter. Cytosolic pH is a second messenger for glucose and regulates the PKA pathway through V-ATPase. *EMBO J.*, 29(15):2515–2526, 2010.

[31] R W DeHaven and R L Hothem. Estimating bird damage from damage incidence in wine grape vineyards. *Am. J. Enol. Vitic.*, 32(1):1–4, 1981.

[32] Fernando Erra Díaz, Ezequiel Dantas, Maia Cabrera, Constanza A Benítez, María V Delpino, Gabriel Duette, Julia Rubione, Norberto Sanjuan, Analía S Trevani, and Jorge Geffner. Fever-range hyperthermia improves the anti-apoptotic effect induced by low pH on human neutrophils promoting a proangiogenic profile. *Cell Death Dis.*, 7(10):e2437, 2016.

[33] Gregory L. Dignon, Wenwei Zheng, Robert B. Best, Young C. Kim, and Jeetain Mittal. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proc. Natl. Acad. Sci.*, 115(40):9929–9934, 2018.

[34] Gregory L. Dignon, Wenwei Zheng, Young C. Kim, Robert B. Best, and Jeetain Mittal. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput. Biol.*, 14(1):1–23, 2018.

[35] Ken A. Dill. Dominant Forces in Protein Folding. *Biochemistry*, 29(31):7133–7155, 1990.

[36] I A Drummond, S A McClure, M Poenie, R Y Tsien, and R A Steinhardt. Large changes in intracellular ph and calcium observed during heat shock are not responsible for the induction of heat shock proteins in drosophila melanogaster. *Mol. Cell. Biol.*, 6(5):1767–1775, 1986.

[37] Peter Eastman, Jason Swails, John D. Chodera, Robert T. McGibbon, Yutong Zhao, Kyle A. Beauchamp, Lee Ping Wang, Andrew C. Simmonett, Matthew P. Harrigan, Chaya D. Stern, Rafal P. Wiewiora, Bernard R. Brooks, and Vijay S. Pande. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.*, 13(7):1–17, 2017.

[38] Robert C. Edgar. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, 5:1–19, 2004.

[39] Francisco Estruch. Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast. *FEMS Microbiol. Rev.*, 24(4):469–486, 2000.

[40] Martin Falk, Yana Feodorova, Natalia Naumova, Maxim Imakaev, Bryan R. Lajoie, Heinrich Leonhardt, Boris Joffe, Job Dekker, Geoffrey Fudenberg, Irina Solovei, and Leonid A. Mirny. Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature*, 570(7761):395–399, 2019.

[41] Nicola Francesca, Domenica E Canale, Luca Settanni, and Giancarlo Moschetti. Dissemination of wine-related yeasts by migratory birds. *Environ. Microbiol. Rep.*, 4(1):105–112, 2012.

[42] Nicola Francesca, Mario Chiurazzi, Raffaele Romano, Maria Aponte, Luca Settanni, and Giancarlo Moschetti. Indigenous yeast communities in the environment of "rovello bianco" grape variety and their use in commercial white wine fermentation. *World J. Microbiol. Biotechnol.*, 26(2):337–351, 2010.

[43] Titus M. Franzmann, Marcus Jahnel, Andrei Pozniakovsky, Julia Mahamid, Alex S. Holehouse, Elisabeth Nüske, Doris Richter, Wolfgang Baumeister, Stephan W. Grill, Rohit V. Pappu, Anthony A. Hyman, and Simon Alberti. Phase separation of a yeast prion protein promotes cellular fitness. *Science*, 359(6371), 2018.

[44] Jeffrey C. Freedman. Ionophores in planar lipid bilayers. In *Cell Physiol. Source B.*, chapter 4, pages 61–66. Elsevier Inc., 2012.

[45] Audrey P. Gasch, Paul T. Spellman, Camilla M. Kao, Orna Carmel-Harel, Michael B Eisen, Gisela Storz, David Botstein, and Patrick O Brown. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell*, 11(12):4241–4257, 2000.

[46] Kerry A. Geiler-Samerotte, Michael F. Dion, Bogdan A. Budnik, Stephanie M. Wang, Daniel L. Hartl, and D. Allan Drummond. Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. *Proc. Natl. Acad. Sci. U. S. A.*, 108(2):680–685, 2011.

[47] Maxim V Gerashchenko and Vadim N Gladyshev. Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic Acids Res.*, 42(17):1–7, 2014.

[48] Tali Gidalevitz, Veena Prahlad, and Richard I Morimoto. The stress of protein misfolding: from single cells to multicellular organisms. *Cold Spring Harb. Perspect. Biol.*, 3(6), 2011.

[49] Italo Giglioli. Insects and yeasts. *Nature*, 56:575–577, 1897.

[50] Michael J. Gray, Wei Yun Wholey, Nico O. Wagner, Claudia M. Cremers, Antje Mueller-Schickert, Nathaniel T. Hock, Adam G. Krieger, Erica M. Smith, Robert A. Bender, James C.A. Bardwell, and Ursula Jakob. Polyphosphate Is a Primordial Chaperone. *Mol. Cell*, 53(5):689–699, 2014.

[51] J. Ignacio Gutierrez, Greg Brittingham, Xuya Want, David Fenyo, and Liam J. Holt. The largest SWI/SNF polyglutamine domain is a pH sensor. *bioRxiv*, 2017.

[52] Tyler S. Harmon, Alex S. Holehouse, Michael K. Rosen, and Rohit V. Pappu. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife*, 6:1–31, 2017.

[53] L H Hartwell. Saccharomyces cerevisiae cell cycle. *Bacteriol. Rev.*, 38(2):164–198, 1974.

[54] S. Blair Hedges, Julie Marin, Michael Suleski, Madeline Paymer, and Sudhir Kumar. Tree of life reveals clock-like speciation and diversification. *Mol. Biol. Evol.*, 32(4):835–845, 2015.

[55] Nikolai Hentze, Laura Le Breton, Jan Wiesner, Georg Kempf, and Matthias P Mayer. Molecular mechanism of thermosensory function of human heat shock transcription factor hsf1. *Elife*, 5, January 2016.

[56] Christopher J. Hewitt and Gerhard Nebe-Von-Caron. The Application of Multi-Parameter Flow Cytometry to Monitor Individual Microbial Cell Physiological State. *Adv Biochem Engin/Biotechnol*, 89:197–223, 2004.

[57] Robert A. Hoffman. Pulse width for particle sizing. *Curr. Protoc. Cytom.*, pages 1–17, 2009.

[58] Thomas Hottiger, Claudio De Virgilio, Walter Bell, Thomas Boller, and Andres Wiemken. The 70-kilodalton heat-shock proteins of the SSA subfamily negatively modulate heat-shock-induced accumulation of trehalose and promote recovery from heat stress in the yeast, Saccharomyces cerevisiae. *Eur. J. Biochem.*, 210(1):125–132, 1992.

[59] Nicholas T Ingolia, Sina Ghaemmaghami, John R S Newman, and Jonathan S Weissman. Genome-wide analysis in vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science*, 324(5924):218–223, 2009.

[60] Kimiharu Ishizawa. *Intracellular pH Regulation of Plant Cells Under Anaerobic Conditions*, pages 59–74. Springer Vienna, Vienna, 2014.

[61] Ryan P Joyner, Jeffrey H Tang, Jonne Helenius, Elisa Dultz, Christiane Brune, Liam J Holt, Sebastien Huet, Daniel J Muller, and Karsten Weis. A glucose-starvation response regulates the diffusion of macromolecules. *Elife*, 5:1–26, 2016.

[62] Bruce S Klein and Brad Tebbets. Dimorphism and virulence in fungi. *Curr. Opin. Microbiol.*, 10(4):314–319, 2007.

[63] Christos M Kougentakis, Ananya Majumdar, and Bertrand García-Moreno, E. Electrostatic effects in proteins are governed by pH-redistribution of the conformational ensemble. *bioRxiv*, 2020.

[64] Joanna Krakowiak, Xu Zheng, Nikit Patel, Jayamani Anandhakumar, Kendra Valerius, David S Gross, Ahmad S Khalil, and David Pincus. Hsf1 and Hsp70 constitute a two-component feedback loop that regulates the yeast heat shock response. *Elife*, 7:1–17, 2018.

[65] Sonja Kroschwald, Shovamayee Maharana, Daniel Mateju, Liliana Malinovska, Elisabeth Nüske, Ina Poser, Doris Richter, and Simon Alberti. Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *Elife*, 4:1–32, 2015.

[66] Sonja Kroschwald, Matthias C Munder, Shovamayee Maharana, Titus M Franzmann, Doris Richter, Martine Ruer, Anthony A Hyman, and Simon Alberti. Different material states of Pub1 condensates define distinct modes of stress adaptation and recovery. *Cell Rep.*, 23(11):3327–3339, 2018.

[67] Robin Van Der Lee, Marija Buljan, and Benjamin Lang. Classification of intrinsically disordered regions and proteins. *Chem. Rev.*, 114:6589–6631, 2014.

[68] Jian Li, Johnathan Labbadia, and Richard I Morimoto. Rethinking HSF1 in stress, development, and organismal health. *Trends Cell Biol.*, 27(12):895–905, December 2017.

[69] Yi Hsuan Lin and Hue Sun Chan. Phase Separation and Single-Chain Compactness of Charged Disordered Proteins Are Strongly Correlated. *Biophys. J.*, 112(10):2043–2046, 2017.

[70] S Lindquist and G Kim. Heat-shock protein 104 expression is sufficient for thermotolerance in yeast. *Proc. Natl. Acad. Sci. U. S. A.*, 93:5301–5306, 1996.

[71] Susan Lindquist. The Heat-Shock Response. *Ann. Rev. Biochem*, 55:1151–91, 1986.

[72] Anne A Madden, Mary Jane Epps, Tadashi Fukami, Rebecca E Irwin, John Sheppard, D Magdalena Sorger, and Robert R Dunn. The ecology of insect-yeast relationships and its relevance to human industry. *Proc. Biol. Sci.*, 285(1875), 2018.

[73] David A. Mangus, Matthew C Evans, and Allan Jacobson. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.*, 4(7):223, 2003.

[74] Erik W. Martin and Tanja Mittag. Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry*, 57(17):2478–2487, 2018.

[75] Anna E Masser, Wenjing Kang, Joydeep Roy, Jayasankar M Kaimal, Jany Quintana-Cordero, Marc R Friedländer, and Claes Andréasson. Cytoplasmic protein misfolding titrates Hsp70 to activate nuclear Hsf1. *Elife*, 8:1–27, 2019.

[76] Marc L. Mendillo, Sandro Santagata, Martina Koeva, George W. Bell, Rong Hu, Rulla M. Tamimi, Ernest Fraenkel, Tan A. Ince, Luke Whitesell, and Susan Lindquist. HSF1 drives a transcriptional program distinct from heat shock to support highly malignant human cancers. *Cell*, 150(3):549–562, 2012.

[77] Bálint Mészáros, Gábor Erdös, and Zsuzsanna Dosztányi. IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.*, 46(W1):W329–W337, 2018.

[78] Gero Miesenböck, Dino A De Angelis, and James E Rothman. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature*, 394(6689):192–5, 1998.

[79] Souvik Modi, Swetha M G, Debanjan Goswami, Gagan D Gupta, Satyajit Mayor, and Yamuna Krishnan. A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat. Nanotechnol.*, 4(5):325–330, 2009.

[80] Fuad Mohammad, Christopher J. Woolstenhulme, Rachel Green, and Allen R. Buskirk. Clarifying the Translational Pausing Landscape in Bacteria by Ribosome Profiling. *Cell Rep.*, 14(4):686–694, 2016.

[81] Amandine Molliex, Jamshid Temirov, Jihun Lee, Maura Coughlin, Anderson P. Kanagaraj, Hong Joo Kim, Tanja Mittag, and J Paul Taylor. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*, 163:123–133, 2015.

[82] Kevin A. Morano, Chris M. Grant, and W. Scott Moye-Rowley. The response to heat shock and oxidative stress in Saccharomyces cerevisiae. *Genetics*, 190:1157–1195, 2012.

[83] R Mortimer and M Polsinelli. On the origins of wine yeast. *Res. Microbiol.*, 150(3):199–204, 1999.

[84] Matthias Christoph Munder, Daniel Midtvedt, Titus Franzmann, Elisabeth Nuske, Oliver Otto, Maik Herbig, Elke Ulbricht, Paul Müller, Anna Taubenberger, Shovamayee Maharana, Liliana Malinovska, Doris Richter, Jochen Guck, Vasily Zaburdaev, and Simon Alberti. A pH-driven transition of the cytoplasm from a fluid- to a solid-like state promotes entry into dormancy. *Elife*, 5:1–30, 2016.

[85] Eiji Nakata, Yoshihiro Yukimachi, Yoshijiro Nazumi, Yoshihiro Uto, Hiroshi Maezawa, Toshihiro Hashimoto, Yasuko Okamoto, and Hitoshi Hori. A newly designed cell-permeable SNARF derivative as an effective intracellular pH indicator. *Chem. Commun.*, 46(20):3526, 2010.

[86] Rammohan Narayanaswamy, Matthew Levy, Mark Tsechansky, Gwendolyn M Stovall, Jeremy D O'Connell, Jennifer Mirrielees, Andrew D Ellington, and Edward M Marcotte. Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation. *Proc. Natl. Acad. Sci. USA*, 106(25):10147–10152, 2009.

[87] Joseph R. Nevins. Induction of the synthesis of a 70,000 dalton mammalian heat shock protein by the adenovirus E1A gene product. *Cell*, 29(3):913–919, 1982.

[88] Susan Nicholls, Michelle D Leach, Claire L Priest, and Alistair J P Brown. Role of the heat shock transcription factor, Hsf1, in a major fungal pathogen that is obligately associated with warm-blooded animals. *Mol. Microbiol.*, 74(4):844–861, 2009.

[89] Susan Nicholls, Melissa Straffon, Brice Enjalbert, André Nantel, Susan Macaskill, Malcolm Whiteway, and Alistair J P Brown. Msn2- and Msn4-like transcription factors play no obvious roles in the stress responses of the fungal pathogen Candida albicans. *Eukaryot. Cell*, 3(5):1111–1123, 2004.

[90] C. Noree, B. K. Sato, R. M. Broyer, and J. E. Wilhelm. Identification of novel filament-forming proteins in Saccharomyces cerevisiae and Drosophila melanogaster. *J. Cell Biol.*, 190(4):541–551, 2010.

[91] Rick Orij, Stanley Brul, and Gertien J. Smits. Intracellular pH is a tightly controlled signal in yeast. *Biochim. Biophys. Acta - Gen. Subj.*, 1810(10):933–944, 2011.

[92] Rick Orij, Jarne Postmus, Alex Ter Beek, Stanley Brul, and Gertien J. Smits. In vivo measurement of cytosolic and mitochondrial pH using a pH-sensitive GFP derivative in Saccharomyces cerevisiae reveals a relation between intracellular pH and growth. *Microbiology*, 155(1):268–278, 2009.

[93] Rick Orij, Malene L Urbanus, Franco J Vizeacoumar, Guri Giaever, Charles Boone, Corey Nislow, Stanley Brul, and Gertien J Smits. Genome-wide analysis of intracellular pH reveals quantitative control of cell division rate by pHc in Saccharomyces cerevisiae. *Genome Biol.*, 13(9):R80, 2012.

[94] B. Panaretou and P. W. Piper. Plasma-membrane ATPase action affects several stress tolerances of Saccharomyces cerevisiae and Schizosaccharomyces pombe as well as the extent and duration of the heat shock response. *J. Gen. Microbiol.*, 136:1763–1770, 1990.

[95] Bradley R. Parry, Ivan V. Surovtsev, Matthew T. Cabeen, Corey S. O'Hern, Eric R. Dufresne, and Christine Jacobs-Wagner. The bacterial cytoplasm has glass-like properties and is fluidized by metabolic activity. *Cell*, 156(1-2):183–194, 2014.

[96] Dawn A. Parsell, John Taulien, and Susan Lindquist. The role of heat-shock proteins in thermotolerance. *R. Soc.*, 339:279–286, 1993.

[97] Ivana Petrovska, Elisabeth Nüske, Matthias C Munder, Gayathrie Kulasegaran, Liliana Malinovska, Sonja Kroschwald, Doris Richter, Karim Fahmy, Kimberley Gibson, Jean-Marc Verbavatz, and Simon Alberti. Filament formation by metabolic enzymes is a specific adaptation to an advanced state of cellular starvation. *Elife*, 3:1–19, 2014.

[98] David Pincus, Jayamani Anandhakumar, Prathapan Thiru, Michael J Guertin, Alexander M. Erkine, and David S Gross. Genetic and epigenetic determinants establish a continuum of Hsf1 occupancy and activity across the yeast genome. *bioRxiv*, 2018.

[99] R Prinzinger, A Preßmar, and E Schleucher. Body temperature in birds. *Comp. Biochem. Physiol. A Physiol.*, 99(4):499–506, 1991.

[100] David S.W. Protter and Roy Parker. Principles and Properties of Stress Granules. *Trends Cell Biol.*, 26(9):668–679, 2016.

[101] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017.

[102] Esmerelda Z. Reyes-Fernandez and Shimon Schuldiner. Acidification of Cytoplasmic pH in Escherichia coli Provides a Strategy to Cope with Stress and Facilitates Development of Antibiotic Resistance. *bioRxiv*, 2020.

[103] Joshua A Riback, Christopher D Katanski, Jamie L Kear-Scott, Evgeny V Pilipenko, Alexandra E Rojek, Tobin R Sosnick, and D Allan Drummond. Stress-triggered phase separation is an adaptive, evolutionarily tuned response. *Cell*, 168:1028–1040, 2017.

[104] Klaus Richter, Martin Haslbeck, and Johannes Buchner. The Heat Shock Response: Life on the Verge of Death. *Mol. Cell*, 40(2):253–266, 2010.

[105] F Ritossa. A new puffing pattern induced by temperature shock and DNP in drosophila. *Experientia*, 18(12):571–573, 1962.

[106] Adam P. Rosebrock. Analysis of the budding yeast cell cycle by flow cytometry. *Cold Spring Harbor Protocols*, 2017(1):pdb.prot088740, 2017.

[107] Adele Rowley, Gerald C Johnston, Braeden Butler, Margaret Werner-Washburne, and Richard A Singer. Heat shock-mediated cell cycle blockage and G1 cyclin expression in the yeast Saccharomyces cerevisiae. *Mol. Cell. Biol.*, 13(2):1034–1041, 1993.

[108] Stefan Rüdiger, Lothar Germeroth, Jens Schneider-Mergener, and Bernd Bukau. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.*, 16(7):1501–1507, 1997.

[109] Z Salvadó, F N Arroyo-López, J M Guillamón, G Salazar, A Querol, and E Barrio. Temperature adaptation markedly determines evolution within the genus Saccharomyces. *Appl. Environ. Microbiol.*, 77(7):2292–2302, 2011.

[110] Benjamin S Schuster, Gregory L Dignon, Wai Shing Tang, Fleurie M Kelley, Aishwarya Kanchi Ranganath, Craig N Jahnke, Alison G Simpkins, Roshan Mammen Regy, Daniel A Hammer, Matthew C Good, and Jeetain Mittal. Identifying Sequence Perturbations to an Intrinsically Disordered Protein that Determine Its Phase Separation Behavior. *bioRxiv*, pages 1–52, 2020.

[111] Guy Sella and Aaron E. Hirsh. The application of statistical physics to evolutionary biology. *Proc. Natl. Acad. Sci.*, 102(27):9541–9546, 2005.

[112] Lei Shi, Benjamin M. Sutter, Xinyue Ye, and Benjamin P. Tu. Trehalose Is a Key Determinant of the Quiescent Metabolic State That Fuels Cell Cycle Progression upon Return to Growth. *Mol. Biol. Cell*, 21:1982–1990, 2010.

[113] Y Shi, D D Mosser, and R I Morimoto. Molecular chaperones as HSF1-specific transcriptional repressors. *Genes Dev.*, 12(5):654–666, 1998.

[114] Alexander A. Shishkin, Georgia Giannoukos, Alper Kucukural, Dawn Ciulla, Michele Busby, Christine Surka, Jenny Chen, Roby P. Bhattacharyya, Robert F. Rudy, Milesh M. Patel, Nathaniel Novod, Deborah T. Hung, Andreas Gnirke, Manuel Garber, Mitchell Guttman, and Jonathan Livny. Simultaneous generation of many RNA-seq libraries in a single reaction. *Nat. Methods*, 12(4):323–325, 2015.

[115] Ishwar S Singh and Jeffrey D Hasday. Fever, hyperthermia and the heat shock response. *Int. J. Hyperthermia*, 29(5):423–435, 2013.

[116] Eric J. Solis, Jai P Pandey, Xu Zheng, Dexter X Jin, Piyush B Gupta, Edoardo M Airoldi, David Pincus, and Vladimir Denic. Defining the essential function of yeast Hsf1 reveals a compact transcriptional program for Maintaining eukaryotic proteostasis. *Mol. Cell*, pages 1–12, 2016.

[117] Christopher M Somers and Ralph D Morris. Birds and wine grapes: foraging activity causes small-scale damage patterns in single vineyards. *J. Appl. Ecol.*, 39(3):511–523, 2002.

[118] P. K. Sorger and H. R. Pelham. Purification and characterization of a heat-shock element binding protein from yeast. *EMBO J.*, 6(10):3035–3041, 1987.

[119] Peter K. Sorger and Hugh R.B. Pelham. Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation. *Cell*, 54(6):855–864, 1988.

[120] Antonia Statt, Helena Casademunt, Clifford P. Brangwynne, and Athanassios Z. Panagiotopoulos. Model for disordered proteins with strongly sequence-dependent liquid phase behavior. *bioRxiv*, 2019.

[121] M Tanabe, A Nakai, Y Kawazoe, and K Nagata. Different thresholds in the responses of two heat shock transcription factors, HSF1 and HSF3. *J. Biol. Chem.*, 272(24):15389–15395, June 1997.

[122] John A. Thomas, Robert N. Buchsbaum, Andrzej Zimniak, and Efraim Racker. Intracellular pH measurements in Ehrlich Ascites tumor cells utilizing spectroscopic probes generated in situ. *Biochemistry*, 18(11):2210–2218, 1979.

[123] Geoffrey C. Tombaugh and Robert M. Sapolsky. Evolving concepts about the role of acidosis in ischemic neuropathology. *J. Neurochem.*, 61(3):793–803, 1993.

[124] Eleanor W. Trotter, Lumilla Berenfeld, Sue Ann Krause, Gregory A. Petsko, and Joseph V. Gray. Protein misfolding and temperature up-shift cause G1 arrest via a common mechanism dependent on heat shock factor in Saccharomycescerevisiae. *Proc. Natl. Acad. Sci. U. S. A.*, 98(13):7313–7318, 2001.

[125] Eleanor W Trotter, Camilla M Kao, Ludmilla Berenfeld, David Botstein, Gregory A Petsko, and Joseph V Gray. Misfolded proteins are competent to mediate a subset of the responses to heat shock in Saccharomyces cerevisiae. *J. Biol. Chem.*, 277(47):44817–44825, 2002.

[126] Amit Tzur, Jodene K. Moore, Paul Jorgensen, Howard M. Shapiro, and Marc W. Kirschner. Optimizing optical flow cytometry for cell volume-based sorting and analysis. *PLoS One*, 6(1):1–9, 2011.

[127] R Martin Vabulas, Swasti Raychaudhuri, Manajit Hayer-Hartl, and F Ulrich Hartl. Protein folding in the cytoplasm and the heat shock response. *Cold Spring Harb. Perspect. Biol.*, pages 1–19, 2010.

[128] Mari Valkonen, Dominik Mojzita, Merja Penttilä, and Mojca Bencina. Noninvasive high-throughput single-cell analysis of the intracellular pH of Saccharomyces cerevisiae by ratiometric flow cytometry. *Appl. Environ. Microbiol.*, 79(23):7179–87, 2013.

[129] J L van de Vossenberg, A J Driessen, M S da Costa, and W N Konings. Homeostasis of the membrane proton permeability in bacillus subtilis grown at different temperatures. *Biochim. Biophys. Acta*, 1419(1):97–104, June 1999.

[130] J. Verghese, J. Abrams, Y. Wang, and K. a. Morano. Biology of the Heat Shock Response and Protein Chaperones: Budding Yeast (Saccharomyces cerevisiae) as a Model System. *Microbiol. Mol. Biol. Rev.*, 76(2):115–158, 2012.

[131] Edward W.J. Wallace, Jamie L. Kear-Scott, Evgeny V. Pilipenko, Michael H. Schwartz, Pawel R. Laskowski, Alexandra E. Rojek, Christopher D. Katanski, Joshua A. Riback, Michael F. Dion, Alexander M. Franks, Edoardo M. Airoldi, Tao Pan, Bogdan A. Budnik, and D. Allan Drummond. Reversible, Specific, Active Aggregates of Endogenous Proteins Assemble upon Heat Stress. *Cell*, 162(6):1286–1298, 2015.

[132] Robert W Walters, Denise Muhlrad, Jennifer Garcia, and Roy Parker. Differential effects of ydj1 and sis1 on hsp70-mediated clearance of stress granules in saccharomyces cerevisiae. *RNA*, 21(9):1660–1671, 2015.

[133] Claudia Weigert, Fabian Steffler, Tomas Kurz, Thomas H. Shellhammer, and Frank Jürgen Methner. Application of a short intracellular pH method to flow cytometry for determining Saccharomyces cerevisiae vitality. *Appl. Environ. Microbiol.*, 75(17):5615–5620, 2009.

[134] David E Weinberg, Premal Shah, Stephen W Eichhorn, Jeffrey A Hussmann, Joshua B Plotkin, and David P Bartel. Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. *Cell Rep.*, 14(7):1787–1799, 2016.

[135] Gabriele Weitzel, Ulrich Pilatus, and Ludger Rensing. The cytoplasmic ph, atp content and total protein synthesis rate during heat-shock protein inducing treatments in yeast. *Exp. Cell Res.*, 170(1):64 – 79, 1987.

[136] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009.

[137] Hadley Wickham. *tidyverse: Easily Install and Load the 'Tidyverse'*, 2017. R package version 1.2.1.

[138] B. J. Wu and R. I. Morimoto. Transcription of the human hsp70 gene is induced by serum stimulation. *Proc. Natl. Acad. Sci. U. S. A.*, 82(18):6070–6074, 1985.

[139] Hang Yao and Gabriel G Haddad. Calcium and pH homeostasis in neurons during hypoxia and ischemia. *Cell Calcium*, 36(3-4):247–255, 2004.

[140] Haneul Yoo, Catherine Triandafillou, and D Allan Drummond. Cellular sensing by phase separation: Using the process, not just the products. *J. Biol. Chem.*, 294(18):7151–7159, 2019.

[141] Taraneh Zarin, Bob Strome, Alex N. Nguyen Ba, Simon Alberti, Julie D. Forman-Kay, and Alan M. Moses. Proteome-wide signatures of function in highly diverged intrinsically disordered regions. *Elife*, 8:1–27, 2019.

[142] Xu Zheng, Joanna Krakowiak, Nikit Patel, Ali Beyzavi, Jideofor Ezike, Ahmad S Khalil, and David Pincus. Dynamic control of Hsf1 during heat shock by a chaperone switch and phosphorylation. *Elife*, 5:e18638, 2016.

[143] Min Zhong, Soon-jong Kim, and Carl Wu. Sensitivity of Drosophila Heat Shock Transcription Factor to Low pH. *J. Biol. Chem.*, 274(5):3135–3140, 1999.

[144] Brian M. Zid and Erin K. O'Shea. Promoter sequences direct cytoplasmic localization and translation of mRNAs during starvation in yeast. *Nature*, 514(7520):117–121, 2014.