DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Glucokinase Polymerization Regulates Enzyme Activity in Response to Environmental Shifts

Citation
Stoddard, Patrick. 2019. Glucokinase Polymerization Regulates Enzyme Activity in Response to Environmental Shifts. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:42029739

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

Glucokinase polymerization regulates enzyme activity in response to environmental shifts

A Dissertation Presented

by

Patrick Stoddard

to

The Department of Molecular and Cellular Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biochemistry

Harvard University

Cambridge, Massachusetts

May 2019

© 2019 – Patrick Stoddard

All rights reserved.

Ethan Garner and Andrew Murray **Patrick Stoddard**

Glucokinase polymerization regulates enzyme activity in response to environmental shifts

Abstract

Unlike cells in a multicellular organism, free living microbes are subject to abrupt changes in environment that can come with no warning. Cells must be prepared to react to these changes on timescales too short for them to change their composition and thus must have mechanisms in place to maintain homeostasis until they can reach a new steady-state. One particularly dangerous transition is the reintroduction to rich carbon sources. In the case of the budding yeast, *Saccharomyces cerevisiae*, the reintroduction to a glucose-rich environment after a period of starvation can result in death; cells can over invest energy in the first steps of glycolysis before they begin to receive returns on this investment. In our investigation of this phenomenon, we have found that one mechanism in place to prevent overinvestment is the polymerization of the glucokinase, Glk1, which catalyzes the first step in glycolysis. Glk1 is a structural homolog of actin, one of the best studied polymer forming proteins, however, that the polymerization of these two proteins has arisen independently. Glk1 polymerization appears to have arisen in the last common ancestor of a group of yeast that existed approximately 150 million years ago. I have found that Glk1 polymerization inhibits Glk1 enzyme activity. Structural studies of the Glk1 protein and Glk1 filaments have

allowed us to design a mutant that cannot polymerize and thus lacks the concentration-dependent inhibition of Glk1 enzyme activity. I have leveraged this mutant to demonstrate that without Glk1 polymerization, cells die more frequently when they are refed glucose and that these cells show additional defects growing on sugars. Thus, the polymerization of Glk1 is important for yeast cells ability to respond to unexpected increases in the concentration of glucose.

Contents

	Abstract	***iii***
	List of Figures	***vii***
	List of Tables	***viii***
	Acknowledgements	***ix***
Chapter 1	**Introduction**	**1**
1.1	Metabolic Regulation in Light of Environmental Transitions	2
1.2	The Actin Superfamily	10
1.3	Summary	15
Chapter 2	**The Actin Superfamily**	**17**
2.1	Background	20
2.2	The Structure of Actin and Actin-like Filaments	22
2.3	The Function of Actin-like Proteins	30
2.4	The Evolution of Actin-Family Proteins	35
2.5	Conclusions and Speculation	41
Chapter 3	**Glucokinase Polymerization Regulates Enzyme Activity in Response to Environmental Shifts**	**43**
3.1	Introduction	46
3.2	Glk1 Forms Polymers that Inhibit Enzyme Activity	47
3.3	The Structure of Glk1 and Glk1 Filaments	54
3.4	Glk1 Polymerization Evolved Independently of Other Actin-Related Polymers	60
3.5	Disrupting Glk1 Polymers	63
3.6	Materials and Methods	69
Chapter 4	**Conclusions and Future Directions**	**73**
4.1	Summary of Results	74
4.2	Yeast Glucokinases and Hexokinases	76

4.3 Fluorescent Proteins Effect on Polymerization 77
4.4 Enzyme Filaments 79
4.5 The Evolution of Metabolism in Yeast 81
4.6 Actin-family Protein Polymers 82
4.7 Glucokinase Becoming a Polymer 83
4.8 Fitness effects of Glk1 Polymerization 84
4.9 Redundancy of Enzymes 86

Appendix A: Additional Information 88
A-1 Glk1 Filament Contacts 89
A-2 Glk1 Filaments in Other Conditions 90
A-3 Table of Strains 91

Appendix B: An Archaeal Tubulin Homolog Forms Tubules 92
B-1 Introduction 93
B-2 Odin Tubulin Polymerizes 93
B-3 Odin Tubulin Forms Tubules 94

Works Cited 98

List of Figures

Chapter 1

1-1 Yeast Glycolysis 9
1-2 Structure of Actin-Related Filaments 12

Chapter 2

2-2 Polymer Formation is a Repeated Feature Within the Actin Superfamily 24
2-4 Evolutionary Relationships Among Members of the Actin Superfamily 38

Chapter 3

3-2-1 Glk1-GFP Polymerizes in Cells 49
3-2-2 Glk1 Polymerization is Rapid and Reversible 51
3-2-3 Glk1 Polymerizes in Response to Substrate of Product 52
3-2-4 Glk1 Polymerization Inhibits Glk1 Activity 53
3-3-1 Glk1 Polymers are Short, Linear Filaments 56
3-3-2 The Structure of Glk1 Filaments 58
3-3-3 Glk1 N and C Terminus Differ from Hxk2 59
3-4-1 Other Glucokinases Also Polymerize 61
3-4-2 Major Filament Contacts Are Conserved in Polymerizing Glucokinases 62
3-5-1 Glk1 Polymerization Can Be Disrupted and Does Not Require Catalytic Activity 64
3-5-2 NonPol Enzyme Activity Lacks Negative Cooperativity 65
3-5-3 Non-Polymerizing Glk1 Causes Growth and Viability Defects 68

Appendix A

A-1 Glk1 Filament Contacts 89
A-2 Glk1 Filaments in Other Conditions 90

Appendix B

B-2-1 Odin Tubulin Oligomerizes With GTP 95
B-2-2 Odin Tubulin Polymerization Can be Observed by Light Scattering 96
B-3-1 Odin Tubulin Forms Slinky-Like Tubules 97

List of Tables

Chapter 3

3-3-1 X-Ray Crystallography Statistics 57

Appendix

A-3 Table of Strains 63

Acknowledgements

The work described herein has not been my work alone. I had many people helping me carry this project and many others cheering me along from the sidelines.

Firstly, my thesis advisors Andrew Murray and Ethan Garner have contributed immensely to the production of this work. Both of them have created spaces where I was free to explore ideas and projects that I found interesting. They gave me room to breathe but were always available to provide guidance and advice when I asked for it. When I was considering being joint between the two of them, I was told that having two advisors either rounds up to two or down to zero. Their contribution to my development and the relationships we have developed over the past several years have certainly been the former.

When I embarked on this project, it was under the wing of Quincey Justman. Quincey was the best daily mentor that I could have hoped for and is a wonderful friend. Her intellectual contributions improved the clarity of my thought, the quality of my writing and speaking, and honed the rigor with which I now approach problems. Few people care for their wards as strongly as Quincey. It can be difficult to find a person whose deep honesty can deliver hard truths that you don't want to hear, but need to hear, while in the same conversation listening to concerns in a way that makes you feel safe, heard, and understood. Quincey is one of these rare people.

The lab that Andrew has cultivated has felt like more of a home than a place of work. Central to this feeling was Beverly. As my adoptive lab mother, she kept me in line, helped me when I screwed up, and was an endless supply of warmth and laughter.

I was sad and to lose her daily presence when she moved on this past year, but I will always treasure her friendship.

My lab siblings were also critical to the happiness of the lab. Special thanks go out to Laura for being my only full science sibling and for all the swolechats, to Phoebe for always having a smile and brightening the bay, to Sri for the late-night comradery, to Zain for taking the edge off, to Miguel for always staying positive, to Felix for all of the singing, and to Nichole, Mary, Marco, Eric, Cara, and Sara. Also, a special thank you to Linda for keeping the ship running and herding us all like a bunch of cats.

The Garner Lab has also been a joy to be a part of. Ethan has been described as ‘a shining beacon of hope for misfits’ and this is reflected in the amazing, ambitious, and delightfully unusual people that he attracts. Georgia, Sean, Matt, and Will have been an integral part of maintaining my sanity during this trek, and Alex Bisson and Ying-Jie Sun both possess drive that I find inspiring. I am also thankful to have had Peter as a lab administrator. Your hard work, cheerful chats, and perennial offer of a Friday tipple made coming to work that much more fun.

I would also like to extend my thanks to my thesis committee: Rachelle Gaudet, Vlad Denic, and Tim Mitchison. They have provided useful feedback throughout my project and have been more than generous with their time and resources. A special thank you goes to the Gaudet lab for their advice in parts of this project as well as their generosity in allowing me to use equipment and in sharing beam time.

I was lucky to have arrived in the program in a year with so many fun and interesting people. The friends that I have made in the program are too numerous to

name. I appreciate what a support they have been and feel privileged to have them in my life.

I have had the privilege of spending the last year of my PhD living at Brentwood amongst friends and cats. The community in that house is wonderful, and they are all special to me. Especially Fritz and Arthur.

I also would not have been able to make it through this last year without the support of my partner, Amy. Her cheerleading has kept me driving forward despite being so far apart. The hours of a scientist are neither short, nor normal, yet she has always understood how important my work is to me and has always encouraged me to chase my dreams. I could not have dreamed that I would find such a wonderful person as my partner.

And lastly, I would like to thank my parents, Bob and Seanna. Everything that I have, I ultimately owe to them. Their willingness to sacrifice for me and their unending, unconditional love has been the bedrock onto which I have built everything that I have ever done. Their support means the world to me, and it makes me happy to know that they are proud.

For my parents,

Seanna and Bob

Chapter 1

Introduction

1.1 Metabolic Regulation in Light of Environmental Transitions
- ***Environmental Transitions***
- **S. cerevisiae *Carbon Metabolism***
- ***Division of Labor in Carbon Metabolism***
- ***Substrate Accelerated Death in* S. cerevisiae**
- ***Glucokinases and Hexokinases***

1.2 The Actin Superfamily
- ***Overview***
- ***Actin and Actin-like Filaments***
- ***Evolution of Polymerization Within the Actin Superfamily***
- ***The Hexokinases Within the Actin Superfamily***

1.3 Summary

Introduction

1.1. Metabolic Regulation in Light of Environmental Transitions

Environmental Transitions

Existing as a unicellular organism presents several unique problems that most cells within a multicellular organism do not encounter. The cells within a large, multicellular organism experience a very narrow range of local conditions; their local environment is defined by the contributions of the other cells within the organism. In contrast, free-living microbes are subject to a wide variety of environments that differ in several of their parameters: temperature, salinity, hydration, pH, and concentration and quality of nutrients. It is critical for their success that they are able to maintain homeostasis within these diverse environments.

In addition to maintaining homeostasis within variable environments, microbes also have to be able to navigate sudden transitions from one environment to another without dying. The transition from one environment to another can often come without warning. Thus, cells must be poised to respond rapidly to any change that they have seen regularly throughout their evolutionary history.

The responses to environmental shifts take many forms, depending on what conditions have changed. In some cases, the cells respond with what is considered a

stress response, e.g. when temperature increases drastically, when an oxidizer enters the environment [1], or when the osmolarity of the solution changes suddenly [2]. Other transitions are also dangerous, and even though they do not trigger the canonical stress response pathways, the cell has mechanisms in place to bridge this passage.

My work examines the transitions between environments with different carbon sources. These transitions come in two forms: withdrawal of a carbon source or addition of new carbon sources. In the case of sudden withdrawal of a preferred carbon source, the cell must cease all activity until it can begin extracting energy from a new source. This is a particularly dangerous transition because the failure to maintain certain energetically expensive structures (e.g. replication forks and translating ribosomes) can lead to damage or collapse events that culminate in cell death. Surprisingly, it is also quite dangerous for microbes to be suddenly introduced to high concentrations of particular carbon sources. This change when not handled properly can result in cell death, which has been dubbed 'substrate accelerated death' [3]. Substrate accelerated death is observed in a variety of prokaryotic and eukaryotic microbial species.

Substrate accelerated death occurs when a population of cells has been starved for a preferred carbon source and are then fed a large bolus of the preferred carbon source. This transition results in cell death. The mechanism of death varies from case to case, and the details of most mechanisms are not understood. Proposed mechanisms for substrate accelerated death include excessive substrate import leading to osmotic rupture [4], accumulation of toxic metabolic intermediates, and depletion of necessary metabolites [5]. One of the best understood examples of substrate

accelerated death occurs when the budding yeast, *Saccharomyces cerevisiae*, is refed glucose.

S. cerevisiae *Carbon Metabolism*

S. cerevisiae and a group of closely related yeast species will ferment glucose and other simple sugars despite the presence of oxygen. This aerobic fermentation is known as the Crabtree effect [6]. This generates ethanol as a byproduct, which has the added benefit of being toxic to many other microbes. Once glucose is depleted, however, the yeast will then express their respiratory genes and respire the accumulated ethanol.

The catabolism of sugars for energy can be split into two phases: glycolysis, which does not require oxygen, and oxidative phosphorylation, which does require oxygen. Glycolysis yields two ATP per molecule of glucose, while oxidative phosphorylation yields nearly thirty ATP per molecule of glucose in yeast [6].

Despite the higher yield achieved by utilizing both glycolysis and oxidative phosphorylation, the number of ATP generated per second is higher for glycolysis, and the machinery necessary for respiration is expensive to produce. Yeast have evolved to repress their respiratory machinery in the presence of high concentrations of sugars and commit fully to fermentation.

Glycolysis can be divided into two main phases as well: upper glycolysis (energy investment) and lower glycolysis (energy extraction). In the first several reactions of glycolysis, two phosphates are added to sugar. In lower glycolysis, the phosphates added during energy investment are transferred to ADP molecules yielding the original

ATP investment and two inorganic phosphates are added to ADP to produce the net yeild of two ATP. The two processes are linked together through a bottleneck reaction: the substrate level phosphorylation of glyceraldehyde 3-phosphate to 1,3-bisphosphoglycerate, coupled to the reduction of NAD+, and catalyzed by glyceraldehyde 3 phosphate dehydrogenase. This reaction requires high concentration of free P_i to drive it forward.

Division of Labor in Carbon Metabolism

The cell uses different enzymes to catalyze the same reactions depending on whether the cell is fermenting or respiring. For example Hxk2 is expressed during fermentation while Hxk1 and Glk1 are expressed during respiration [7]. This division of labor is not completely understood. The enzymes frequently differ in their rates, their affinity for substrate, and the methods of regulation. It is thought that they represent different needs to yield a balanced, functioning system under different conditions.

***Substrate accelerated death in* S. cerevisiae**

When *S. cerevisiae* is refed glucose, a portion of the population will lose its energy charge and these cells die [8]. Certain mutations change the amount of the population that dies. For example, deletion of TPS1, a gene involved in storage carbohydrate synthesis, results in an increase in the number of cells that die. In the *tps1*Δ background, deletion of HXK2, a gene encoding a hexokinase, the number of cells dead returns to normal [9]. Studying these mutants with a combination of genetic, biochemical, and modeling experiments have revealed that cell death in this case

occurs by an imbalance of upper glycolysis (energy investment) and lower glycolysis (energy extraction) [10].

When fermenting at steady state, cells generate four molecules of ATP per two molecules of ATP invested for sugar phosphorylation. Thus, for every two molecules of ATP produced, one molecule is reinvested into glycolysis while the other is used for other cellular processes, i.e. as an energy source, hydrolyzed to ADP and inorganic phosphate (P_i). Hydrolysis of ATP is one of the principal sources or P_i within the cell, and high levels of P_i are needed to drive the downstream steps of glycolysis.

When starved cells are introduced to a glucose-rich environment, overinvestment of energy into sugar phosphorylation is possible. The sudden appearance of high substrate concentration can break the balance between phosphorylation of sugars and ATP hydrolysis in the cell. With much of the phosphate within the cell tied up in sugar-phosphates, the P_i concentration becomes too low to drive lower glycolysis. The final result is a cell with large amounts of sugar phosphates from which it cannot extract energy, and a completely depleted ATP pool (Figure 1-1 A) [10].

S. cerevisiae's primary means of preventing P_i depletion upon glucose refeeding is through futile cycling of trehalose synthesis. Trehalose (1,1-α,α-diglucose) is a storage carbohydrate that is synthesized from glucose-6-phosphate (the product of glucose phosphorylation). The synthesis of trehalose releases the sugar-bound phosphate as P_i, thereby maintaining sufficient P_i to drive the energy extraction steps of glycolysis. Elimination of the trehalose synthetic enzymes (*tps1*Δ) increases the number of cells that die during glucose refeeding. I find that in the case of wild-type cells, 0-10% of cells die, while only approximately 1/10,000 *tps1*Δ cells survive.

The trehalose cycle also serves an important secondary function. The intermediate of trehalose synthesis, trehalose-6-phosphate, inhibits the activity of the hexokinases responsible for glucose phosphorylation. The inhibition of these enzymes also slows the rate of energy investment into upper glycolysis until lower glycolysis can catch up. Thus, the transient accumulation of trehalose-6-phosphate reduces the number of cells that die during refeeding. Both overexpression of the hexokinases and replacing them with a hexokinase that is not inhibited by trehalose-6-phosphate from the fission yeast *Schizzosacharomyces pombe* results in a small increase number of dead cells (~10%) [9,10]. Interestingly, the glucokinase in *S. cerevisiae* is not inhibited by trehalose-6-phosphate (Figure 1-1 C).

Glucokinases and Hexokinases

Much of yeast's ability to handle glucose refeeding comes from decreasing the efficiency of the first steps in glycolysis: the trehalose cycle consumes glucose-6-phosphate, and inhibition of hexokinase activity by T6P reduces the rate of glucose-6-phosphate formation. In *S. cerevisiae*, three different enzymes catalyze the phosphorylation of glucose, two hexokinases (Hxk1 and Hxk2) and one glucokinase (Glk1) [7] (Figure 1-1 B). The amount of each of these three enzymes depends on the nutritional history of the cell.

Hxk2 is the primary sugar kinase, and it acts as a transcription factor that regulates the expression of Hxk1 and Glk1. Hxk2 is most strongly expressed during fermentative growth, while Glk1 and Hxk1 are both repressed. When cells are respiring or starved, however, Glk1 and Hxk1 are more strongly expressed.

In this work I found that Glk1 forms linear polymers while Hxk1 and Hxk2 do not. Hxk1, Hxk2, and Glk1 are actin-fold proteins and are thus structural homologs of several other polymer forming proteins. I found that the polymerization of Glk1 regulates Glk1 activity. Thus, understanding the polymerization of actin and related polymers is essential for interpreting these effects.

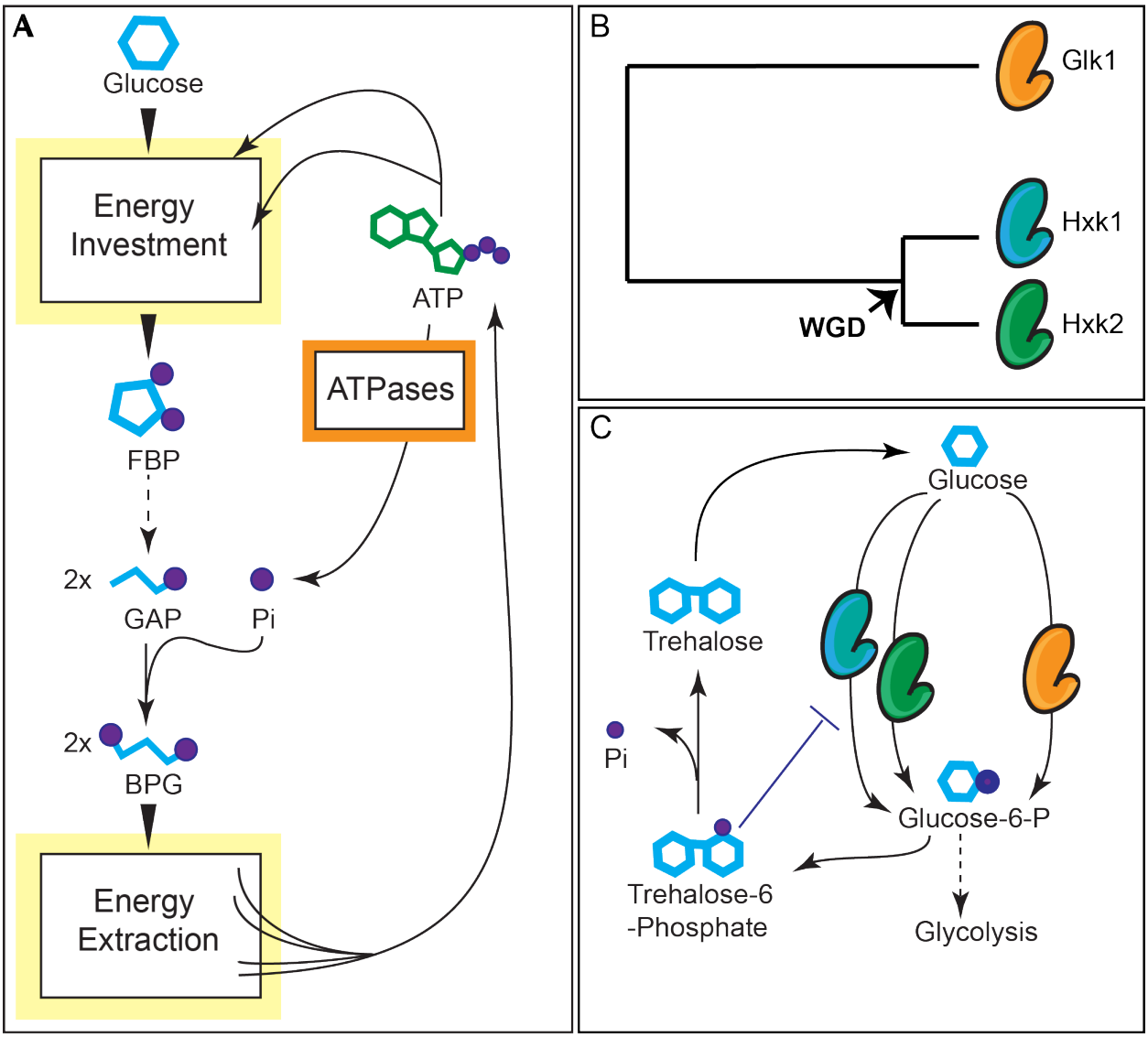


Figure 1-1. A: Simplified schematic of energy balance in glycolysis. Approximately one half of ATP generated is reinvested into the phosphrylation of sugars while the other half is hydrolyzed or used for other processes. This keeps Pi levels high, which drives the unfavorable substrate-level phosphorylation of glyceraldehyde-phosphate (GAP) to bisphosphoglycerate (BPG). B: Three enzymes in *S. cerevisiae* phosphrylate glucose; two hexokinases, Hxk1 and Hxk2 (blue and green), and one glucokinase, Glk1 (orange). The two hexokinases diverged at the yeast whole genome duplication. C: When Glucose-6-phosphate accumulates trehalose synthesis increases. Trehalose-6-phosphate increase inhibits the activity of the hexokinases, but not the activity of Glk1.

1.2. The Actin Superfamily

Overview

The actin superfamily is a functionally diverse group of structurally similar proteins that is present in all three domains of life: eukaryotes, eubacteria, and archea [11,12]. Its members are so broadly diverged that it is often difficult to identify members from sequence homology alone, however, the structural similarity is immediately evident.

All members of the actin superfamily undergo a large conformational change upon substrate binding and/or substrate hydrolysis [13]. This conformational change is used by certain members of the superfamily to do work [14,15]. Several protein families within the actin superfamily use the energy of ATP hydrolysis to form linear polymers. These polymers can be used to organize the cell interior [16], to transport material [17-20], to exert force [21,22], to convey cell shape [23], and to divide the cell [24,25].

Actin and Actin-like Filaments

The most studied member of the actin superfamily is actin itself. Actin is a nucleation-limited, helical polymer composed of two parallel strands [26]. Actin filaments can reach several microns in length and have a persistence length of nearly 10 microns [27]. Actin is present in all eukaryotes and is one of the principal cytoskeletal proteins. Very similar polymers have also been found in several archaeal species, suggesting that actin predates the endosymbiotic event between a eubacterium and an archaeal cell that gave rise to eukaryotes [12,28].

Several other families of actin homologs have also been found to form filaments. Despite the conservation in their monomer structure, these filaments have diverse quaternary structure and show varied filament kinetics [29]. The structure and kinetics of each filament are suited to its particular function. The diversity in filament structure arises mostly from changes in the inter-strand contacts. The longitudinal contacts along an individual strand are loosely conserved: the orientation of one monomer to the next is conserved, and subdomain IB contacts subdomain IA, while subdomain IIB contacts subdomain IIA [30] (Figure 1-2A).

Within the known actin-like polymers nearly every structural parameter is varied. While actin is a right-handed helix, ParM, a filament that segregates plasmids in *E. coli*, forms left-handed helices [31]. Actin filament monomers are staggered between strands, providing interaction between more monomers within the filament. MamK filaments are used by magnetotactic bacteria to organize magnetosomes (magnetic field sensing organelles). MamK filaments are not staggered, and thus, monomers only contact one other monomer on the opposite strand [32]. MreB filaments, which are responsible for orienting bacterial cell wall synthesis, are anti-parallel, non-staggered, and lack a helical twist [30]. These structural variations continue to accumulate as we investigate more actin-like filaments (Figure 1-2B).

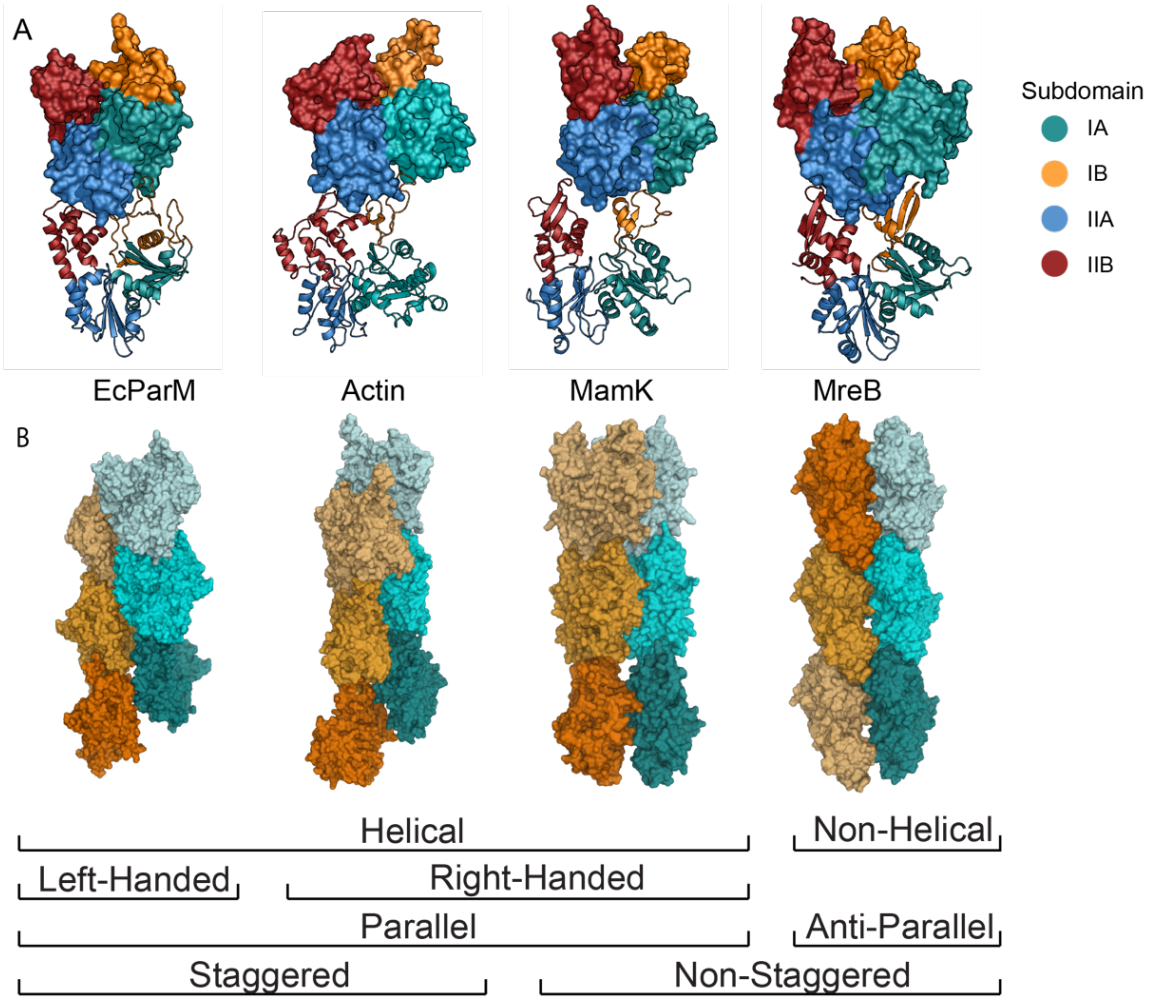


Figure 1-2: Structure of Actin-Related Filaments. A selection of filaments showing that the intra-strand geometry is conserved (**A**), while the filament morphology varies (**B**). Within one strand, subdomain IB contacts subdomain IA, and subdomain IIB contacts subdomain IIA, shown here as a surface representation above a ribbon diagram. The variation observed within the filament structures arises from variation in inter-strand contacts and the curvature of individual strands. Here filament structures are displayed as surface representations. Monomers in one strand are shades of orange or blue. From left to right: *E coli* ParM (PDB ID: 5AEY) (Bharat et al., 2015) is a parallel, left-handed, helical filament with staggered strands. Actin (PDB ID: 4A7N) (Behrman et al., 2012) is a parallel, helical filament with staggered strands. MamK (PDB ID: 5LJV) (Löwe et al., 2016) filaments are parallel, helical filaments whose strands are not staggered. MreB (PDB ID: 4CZJ) (van den Ent et al., 2014) is a nonhelical, antiparallel filament whose strands are not staggered.

Evolution of Polymer Formation Within the Actin Superfamily

The high degree of sequence divergence between actin-fold protein families makes determining their relatedness challenging. It is difficult to differentiate between a model of divergence from a single ancestor and a model where different proteins converged on the same protein fold. However, primary sequence degenerates much faster than tertiary structure and the structural similarity of all of these proteins suggests that they are all from a single origin. In fact, when hidden Markov models are built for each individual family within the actin superfamily and these models are aligned, there is better than chance similarity between all of the different families. This suggests that all of the protein families are diverged from a single actin-fold protein that was present in the last universal common ancestor [29].

It is similarly difficult to determine if polymerization has evolved within the actin superfamily once, or if the extant actin-like polymers represent several independent origins of polymerization. Like the actin-fold proteins, actin-like polymers are present in all known organisms. This allows for two possible models: either the last universal common ancestor contained a polymer-forming actin homolog, and actin and all actin-like polymers are descended from this protein, or the last universal common ancestor had an actin-fold protein that gave rise to several protein lineages, several of which independently evolved the capacity to form polymer. Phylogenetic analysis of the protein families within the actin superfamily reveal that the polymer-forming protein families form a monophyletic group along with the non-polymerizing chaperone, Hsp70 [29]. This suggests that the polymer forming actin homologs are all descended from a single polymer forming protein that was present in the last universal common ancestor.

Furthermore, because we understand that Hsp70 is a protein of bacterial origin, it is likely that Hsp70 evolved from a bacterial filament-forming protein, losing its ability to form polymer in the process. Consistent with this hypothesis, the hidden Markov model of Hsp70 is most similar to that of MreB, a bacterial actin-like polymer [33].

The Hexokinases Within the Actin Superfamily

The actin superfamily contains several families of kinases, including the hexokinases. Hxk1, Hxk2, and Glk1 are all members of the hexokinase protein family. Despite sharing their fold with actin and its related proteins, the sequence similarity between the hexokinases and most other actin-fold protein families is nearly zero. When comparing hidden Markov models, the hexokinase family shows moderate similarity to other families of carbohydrate kinases within the actin superfamily, but very poor similarity to any other polymer forming actin-fold protein families. This suggests that the hexokinases are very distantly removed from the polymer forming actin-fold proteins.

1.3. Summary

Microbes are subject to large changes in condition that can occur on second timescales, and without warning. This requires that they are prepared to rapidly respond to dangerous changes in conditions. Cells have numerous mechanisms in place to navigate the bumpy landscape from one steady state to the next. A dangerous change in condition that all microbes are subject to is the sudden abundance of preferred carbon source.

I have found that the *S. cerevisiae* glucokinase, Glk1, forms polymers upon glucose refeeding. Glk1 is a structural homolog of actin, an important cytoskeletal polymer, and we were therefore curious about how and why Glk1 formed polymer while other sugar kinases do not. Furthermore, all of the known actin-related proteins that also form polymer are more closely related to actin than they are to other actin-fold proteins. Thus, it appeared that Glk1 independently evolved polymerization, and represented the first definitive example of parallel evolution of this behavior within the actin superfamily.

In chapter 2 I give a detailed overview of the actin superfamily and actin related filaments. This chapter was initially printed in Molecular Biology of the Cell in 2017 under the title “Evolution of polymer formation in the actin superfamily.” In chapter 3.2, I describe how we found Glk1 polymers in cells and how I studied their polymerization in a purified system, which lead to our understanding of the effects of Glk1 polymerization. In chapter 3.3, I describe how I was able to solve the structure of Glk1 and how we were able to build a model of Glk1 filaments. This model revealed sequence constraints on

Glk1 polymerization that I discuss in chapter 3.4. These conserved regions allowed me to pinpoint when during yeast evolution this trait arose. In chapter 3.5 I discuss how I was able to use the model of Glk1 filaments to design a point mutant that did not polymerize but was still catalytically active. I used this mutant to understand what the physiological consequences of Glk1 not polymerizing were, which led to insights about the role of Glk1 polymerization within yeast. Finally, in chapter 4 I summarize this work again and describe its significance in the fields of microbiology, protein evolution, protein polymers, and metabolism. I also describe open questions regarding Glk1 polymers and describe extensions to the project that I wish to see performed.

Chapter 2

The Actin Superfamily

2.1 Background

2.2 The Structure of Actin and Actin-like Filaments
- ***Allostery and Monomer Contacts Within an Actin-like Filament***
- ***Variation in the Structure of Actin-related Filaments***

2.3 The Function of Actin-like Proteins
- ***Bridging Long-length Scales***
- ***Cell Mechanics***
- ***Force Generation Through Polymerization Dynamics***
- ***Cell Division***
- ***Sensing and Remodeling Cell Shape***
- ***Macromolecular Scaffolds***

2.4 The Evolution of Actin-family Proteins
- ***Structural similarity and shared ancestry among actin-like proteins***
- ***Evolution of polymerization in the actin superfamily***
- ***The Prokaryotic Origins of Eukaryotic-Type Actin***

2.5 Conclusions and Speculation

Attributions

This chapter was previously published under the title “Evolution of Polymer Formation in the Actin Superfamily” in Molecular Biology of the Cell in 2017. The authors were me (Patrick R. Stoddard), Tom A. Williams, Ethan Garner, and Buzz Baum. The idea for the review was conceived by Ethan Garner and Buzz Baum. I wrote the initial draft of the paper and the paper was edited by me, Ethan Garner, and Buzz Baum. Tom A. Williams performed the phylogenetic analysis in the paper and wrote the section describing that work.

Abstract

While many are familiar with actin as a well-conserved component of the eukaryotic cytoskeleton, it is less often appreciated that actin is a member of a large super-family of structurally-related protein families found throughout the tree of life. Actin-related proteins include chaperones, carbohydrate kinases and other enzymes, as well as a staggeringly diverse set of proteins that utilize the energy from ATP hydrolysis to form dynamic, linear polymers. Despite differing widely from one another in filament structure and dynamics, these polymers play important roles in ordering cell space in bacteria, archaea, and eukaryotes. It is not known if these polymers descended from a single ancestral polymer or arose multiple times by convergent evolution from monomeric actin-like proteins. In this work, we provide an overview of the structures, dynamics, and functions of this diverse set. Then, using a phylogenetic analysis to examine actin evolution, we show that the actin-related protein families that form polymers are more closely related to one another than they are to other non-polymerizing members of the actin superfamily. Thus, all the known actin-like polymers are likely to be the descendants of a single ancestral polymer-forming actin-like protein.

2.1 Background

Actin was one of the first biological polymers to be characterized. It was first extracted from skeletal muscle tissue by Haliburton in 1887, who referred to it as 'myosin ferment' due to its ability to coagulate in the presence of the motor protein, Myosin [34]. In 1942, Straub purified monomeric actin, opening up the possibility of studying polymer behavior *in vitro* [35]. Later, Huxley and Hanson showed that actin filaments are one of the major components of contractile muscle sarcomeres [36], and Pollard identified actin as the thin filaments present in the contractile cytoplasm of *Acanthamoeba* [37]. Since then, work by many researchers has elucidated the ultrastructure, regulation, dynamics and function of actin filaments from a wide range of eukaryotes, where it functions in concert with other cytoskeletal filaments and regulators to control dynamic, internal cellular organization. Although this work revealed the ubiquity of eukaryotic actin, the view of the actin cytoskeleton as being one of the distinguishing features of eukaryotes changed with the discovery of similar filaments within bacteria, archaea, and mobile genetic elements [38,39]. It is now clear that actin represents a single member of a much larger family of polymer-forming proteins that are present in organisms from across the tree of life.

Despite the wide divergence of these different actin-like polymers, they carry out a number of conserved functions across the domains of life: they control the shape of both eukaryotic and many prokaryotic cells; are common, conserved components of the cell division and chromosome segregation machinery [40]; and control the dynamic spatial organization of the cytoplasm in many eukaryotes and some prokaryotes.

Interestingly, the wider family of actin related proteins includes a number of proteins that have a fold that is strikingly similar to that of actin, but which do not form filaments. These include the molecular chaperone Hsp70/DnaK, and enzymes such as hexokinase and glutamate mutase [11]. The structural similarity between both evolutionarily distant filament forming proteins, and their similarity to these non-polymerising proteins raises a number of questions. How are the actin-like polymers and actin-like enzymes related? Did polymerization arise once, or multiple times, during evolution? Given the conserved functions of these evolutionarily distant actin-like filaments, what common features do they have that allows them to accomplish these tasks, and what features vary across evolution to change their activity and role? In this review, we aim to explore these questions. In doing so, we will consider two reciprocal hypotheses:

1. The self-assembly of actin filaments is ancient and evolved only once, i.e. all extant actin-like filaments are descended from a single ancestral polymer. The diversity of functions observed for modern actin-like polymers then represent descent with modification of this ancestral polymerization activity.
2. Filament formation is a convergent phenomenon within the actin superfamily, with polymerization arising independently, multiple times from an ancestral monomeric ATPase.

In exploring the evolution of actin-like proteins, we begin by examining their structure and functions across different systems. We then look at the phylogenetic tree of actins and related proteins to explore whether the sequences of the extant actins can

shed light on the evolution of this diverse family of proteins, and to test these alternative hypotheses about the origin of polymerization within the family.

2.2. The structure of actin and actin-like filaments

Actin is one of the most conserved and highly expressed proteins in eukaryotes, being 91% identical between *Caenorhabditis elegans* and *Homo sapiens*. Nevertheless, in some eukaryotes, such as plants, actin has diverged more widely [41], implying that the conservation in animals reflects, in part, constraints imposed by the need to preserve interactions of the monomer with a wide set of actin-binding proteins and/or as the result of its incredibly high level of expression [42], rather than some fundamental limit to the ability of the protein itself to tolerate changes in its sequence. This divergence becomes much more evident when bacterial, archaeal and eukaryotic actins are compared where, in many cases, it is challenging to identify their homology using sequence alone. Despite their divergence, these proteins show highly similar fold, and it was through this structural homology that Bork and Sander were able to build a sequence motif predicting the presence of filament forming actins in prokaryotes [39].

The first low-resolution structure of the actin monomer was obtained by Kabsch, Mannherz, and Suck in 1985, and the same team produced the first atomic model in 1990 with the help of Pai and Holmes [43,44]. It was immediately evident from their analysis that, despite their low sequence similarity, actin shared the same fold as hexokinases [45]. At the heart of actin, flanked by two globular domains, is a nucleotide-binding pocket that binds and hydrolyzes ATP. Importantly, the structure of the monomer is profoundly altered by presence and state of the nucleotide, i.e. ATP or

ADP. As a result, ATP hydrolysis and phosphate release leads to large-scale movement of the two domains of the protein relative to one another [13,46,47].

Allostery and monomer contacts within an actin-like filament

When embedded within a filament, actin-like monomers make extensive contacts with one another. These include head to tail contacts between monomers within the two protofilaments and contacts occurring across protofilament strands (Figure 2-2). In most actin-like polymers, the location of the contacts between monomers within protofilaments is conserved: subdomain Ia in one monomer interacts with residues in subdomain Ib in the next, while subdomain IIa interacts with subdomain IIb (Figure 2-2C). Even in FtsA, which is missing subdomain Ib, the longitudinal contacts appear to be conserved; an insertion into subdomain Ia takes the place of subdomain Ib within a presumed polymer [48,49]. Thus, these contacts may be essential across actin-like proteins to enable the formation of dynamic filaments, since nucleotide binding shortens the distance separating subdomains Ib and IIb, matching the distance between subdomain Ia and IIa [30] (Figure 2-2).

As a result of these contacts, changes in the conformation of individual monomers, e.g. as the result of ATP hydrolysis and phosphate release, are felt by neighboring monomers in the filament. As well as making actin an excellent chassis for use in the generation of cytomotive filaments, this monomer-monomer communication also enables actin-like proteins to function as an allosteric ATPase with high cooperativity: nucleotide hydrolysis (either random, or vectorial) occurs at a greater rate

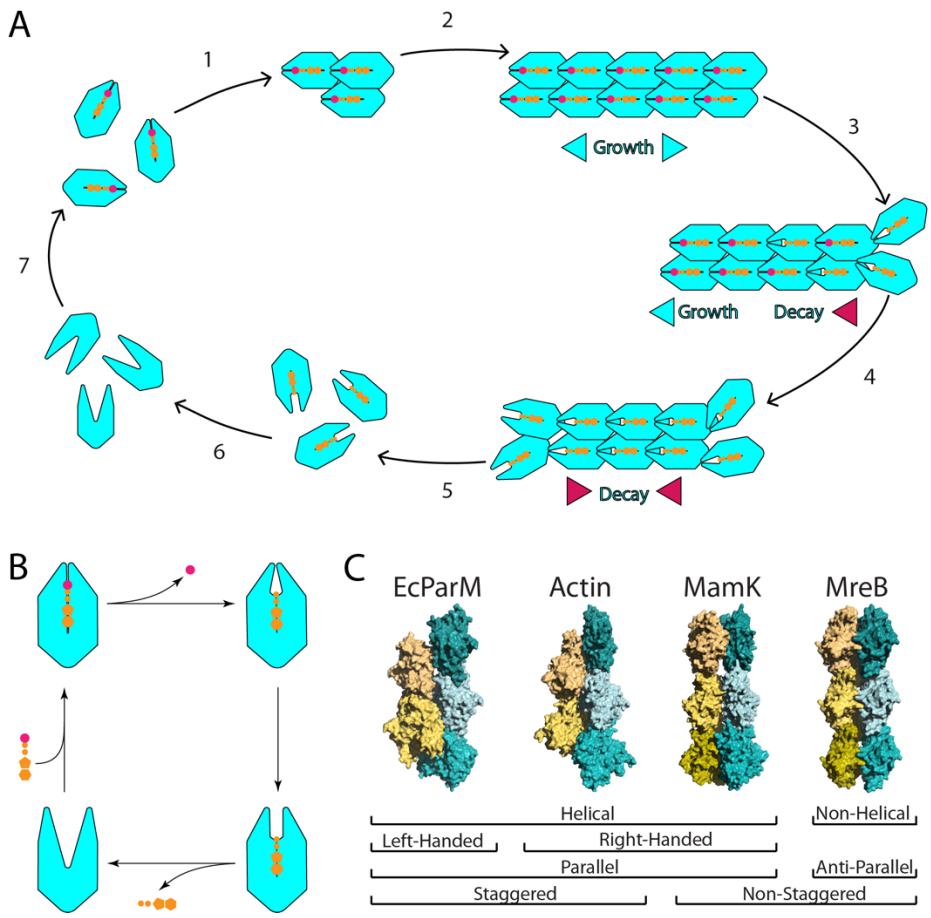


Figure 2-2: Polymer Formation is a Repeated Feature Within the Actin Superfamily.
A: Polymerization of actin homologues requires the formation of a filament nucleus (1). Once a nucleus has formed, filament elongation is rapid (2). Once monomers have been incorporated into the filament, they begin to hydrolyze ATP (3) until the filament end(s) is/are ADP bound monomers (4). Filaments initially grow in both directions. Under certain conditions, some filaments grow from one end and shrink from the other. Eventually, lower affinity between ADP bound monomers allows for filament disassembly from the end (5). ADP then dissociates from the ADP bound monomers (6), which then rapidly rebind ATP (7). **B:** When ATP is hydrolyzed by actin monomers within filaments, inorganic phosphate is released, inducing a conformational change within the monomer. This conformational change reduces the affinity between monomers. Once the ADP monomer has dissociated from the filament, it has low affinity for the polymer. The monomer then releases ADP. The now open monomer rebinds ATP and enters a conformation with a high affinity for the filament. **C:** Amongst actin-like filaments, the contacts within individual protofilaments vary little. By varying lateral contacts between protofilaments, different filament structures with different properties and behaviors can be generated. A subset of these structures is displayed here. Left: *E. coli* ParM is a parallel, left-handed, helical filament whose protofilaments are staggered (PDB ID: 5AEY) (Bharat *et al,* 2015). Left-Center: actin is a parallel, helical filament whose protofilaments are staggered (PDB ID: 4A7N) (Behrman *et al,* 2012). Right-Center: MamK filaments are parallel, helical filaments whose protofilaments are not staggered (PDB ID: 5LJV) (Löwe *et al,* 2016). Right: MreB is a non-helical, anti-parallel filament whose protofilaments are not staggered and which has an intrinsic curvature (PDB ID: 4CZJ) (van den Ent *et al,* 2014).

in the filament, and thus polymerization may function as a way to regulate enzyme (e.g. ATP hydrolysis) activity. In this way, it is formally possible that some members of the enzyme/chaperone actin subfamilies form polymers. This would not be unprecedented, as other enzymes have been shown to assemble into filaments, which serve to regulate their activity. One of the best examples of this is the non-actin-related protein, CTP synthetase [50]. Depending on the organism, polymerization can activate or inhibit CTP synthetase activity [51,52]. Similarly, several metabolic enzymes in yeast have been shown to form pH dependent polymers in response to starvation [53]. This raises the possibility of there being metabolic enzymes within the clade of non-polymerizing actin-like proteins (Figure 2-4) that employ polymerization as a mechanism for allostery.

Variation in the structure of actin-related filaments

A great deal of variation of filament structure comes from changes in the lateral contacts between protofilaments [30]. This can alter helical pitch, filament handedness, and can determine whether proto-filaments lie parallel or antiparallel to one another (Figure 2-2C). Because of this, filaments built from monomers with a similar fold can have very different properties.

In most actin-like filaments, two protofilaments twist around one another like rope, giving the structure physical strength under tension and a high degree of flexibility. In eukaryotic actin, these two protofilaments align in parallel to form a right-handed helical filament that is both polar and chiral. The polarity, in particular, is an essential property when actin is used as a substrate for directed molecular motors, like myosin [54].

However, there are exceptions to these structural characteristics. Many bacterial actins differ from eukaryotic actin in their pitch, in their handedness, in the polarity of protofilaments, and in the degree to which their protofilaments are staggered. Moreover, these structural differences have significant consequences for the function of these polymers. These consequences are revealed by an examination of the most notable of outliers in terms of filament architecture: MreB, MamK and certain members of the ParM family of proteins.

MreB polymers are involved in bacterial cell wall synthesis and maintenance of prokaryotic rod shape. MreB forms two-stranded, antiparallel, non-helical filaments, with no stagger between the subunits in each protofilament [30]. The lack of a twist is essential to MreB function in two ways. First, since MreB associates with the membrane through a single contact face, in the context of a helical polymer, fewer monomers would be able to form tight contacts with the membrane, weakening the membrane-filament interaction. Furthermore, because MreB lacks a helical twist, MreB filaments can sustain an intrinsic curvature - something that appears to be crucial for its function in the regulation of bacterial cell shape [30,55].

MamK spatially organizes magnetic organelles (magnetosomes) within magnetotactic bacteria. MamK filaments form parallel helical filaments like actin but with no stagger between the subunits in each protofilament. The buried surface area between protofilaments is approximately two-thirds of that of other actin-like filaments, leaving an open cavity between each of the protofilaments [32,48,56]. This raises some interesting questions about how these structural differences influence MamK polymer

dynamics, the interaction of the polymer with magnetosomes, and how the lack of stagger affects the mechanical properties of MamK filaments.

ParM from the *Escherichia coli* R1 plasmid (*Ec*ParM) forms filaments that differ from actin filaments most strikingly in their handedness, with a left handed twist. ParM from the *Bacillus thuringiensis* pBMB67 plasmid (*Bt*ParM), however, forms two-stranded, anti-parallel, supercoiled, helical filaments. In the presence of its DNA adaptor protein (ParR), these filaments associate to form a four stranded nanotubule with an open core; a structure somewhat reminiscent of microtubules [57]. While this structure's mechanical properties have not been studied in detail, it is likely that this filament is stiffer than a two-stranded filament, again, reminiscent of the mechanical properties of microtubules or of stiff fascin-crosslinked actin bundles seen in many eukaryotic cells [58,59]. While speculative, the fact that *Bt*ParM has converged on microtubule-like structure, suggests that such properties may be desirable for filament systems which function in the segregation of DNA.

Variation in Filament Kinetics

Actin-like polymers also differ from one another widely in their dynamics. Changes in kinetic parameters (e.g. nucleation rate, monomer on/off rate, ATPase activity, phosphate release kinetics) can yield polymers with different length distributions, stability, and polymerization and depolymerization rates. Importantly, however, since polymer dynamics are sensitive to changes in amino acid sequence that

cause little or no change in secondary and tertiary structure, this may not be evident from a simple analysis of the structure of the polymer.

In most eukaryotes, a small number of actins are used to perform a wide variety of cellular tasks. This flexibility of function is made possible by the presence of a large number of actin-binding proteins that modify different kinetic or structural parameters of the monomer and/or polymer. This tunability enables cells to exert dynamic local control over the actin cytoskeleton to determine where actin filaments form, how fast they elongate, and for how long they persist [60]. Crosslinks between actin filaments further expand this repertoire to enable networks with different filament organization [61,62].

In eukaryotes, much of the control is exerted at the level of nucleation, as actin is a nucleation-condensation polymer: the affinity of a monomer for another monomer is much lower than its affinity for a growing filament or a filament nucleus [61,63]. This is both a natural consequence of monomers having multiple contacts in the context of a filament, and a useful property since it disfavors the spontaneous nucleation of filaments even in the presence of high concentrations of the monomer. As a result, nearly all actin-like polymers whose assembly dynamics have been elucidated have been found to be nucleation-limited. However, relative to eukaryotic actin, the nucleation rate tends to be much faster for bacterial actins. For example, *Ec*ParM forms polymers 200 times more easily than eukaryotic actin, even though filaments elongate at a similar rate [64]. Similarly, the plasmid-segregating actin AlfA from *Bacillus subtilis* also assembles via a rapid, yet nucleation-limited pathway [65]. In these cases, the ability to undergo spontaneous polymerization at a suitable site may do away with the need for a

dedicated nucleator. Interestingly, there are eukaryotic actins within parasites such as *Toxoplasma gondii* that assemble isodesmically [66] (the affinity of a monomer for another monomer is equal to the affinity between a monomer and a filament end), thus obviating the need for a nucleator. These actins have been reported to exhibit atypical lateral filament contacts [67] and form short microfilaments just beneath the cell membrane.

Following polymerization, the behavior of a filament will depend on the rates of ATP hydrolysis and phosphate release. As ATP-bound monomers are incorporated into actin filaments, bound ATP is hydrolyzed such that the ends of the filament tend to be composed of ATP-actin as long as the filament continues to grow [68] (Figure 2-2). This in turn affects polymer dynamics, since, for eukaryotic actin, the ADP bound form has an off rate at the minus end that is much higher than that of ATP actin [69]. This, together with a slight asymmetry in the on/off rates at each end of the filament, can lead to treadmilling, whereby monomers flux through a filament of constant length as the filament grows via monomer addition at its (+) end at the same rate as it shrinks via monomers loss from its (-) end [70]. Other actin-like polymers display this behavior. These include the plasmid segregators Alp7a and AlfA, which show evidence of treadmilling [65,71], and MamK, which treadmills to evenly distribute magnetosomes between daughter cells [14].

In polymers where the off rate of the ADP bound monomers is much higher than the ATP bound form, filaments can even exhibit dynamic instability, suddenly transitioning between polymerization and depolymerization - a key feature of microtubule dynamics [72]. While this is not the case for eukaryotic actin under most conditions (for a possible exception see [73]), the plasmid segregating polymers *Ec*ParM

and Alp7a (from *B. subtilis*) exhibit dynamic instability [64,74]. Although there are exceptions to this rule, e.g. AlfA [75], it is interesting to ask whether this property aids the segregation of DNA, as is the case with microtubules. If so, it shows how, in some instances, evolution may have settled on a common solution in terms of polymer dynamics, as well as polymer structure (see *Bt*ParM above), to achieve a common purpose.

At the same time, there are many clades of structurally similar enzymes that do not appear to form polymers. These include the many "actin related proteins" (ARPs) in eukaryotes. Although, in the case of the Arp2/3 complex, the actin-related proteins in the complex have lost their ability to form filaments themselves, the large conformational change associated with their ATPase activity is still used to cause allosteric changes required for the protein to function. It determines whether the dimer is in an open or closed state, which regulates its ability to seed new actin filaments [76,77] and modulates Arp2/3 affinity for other nucleation and branching factors [78]. It remains to be seen whether there are similar non-polymerizing proteins within the bacterial family of "actin like proteins" (Alp) [71].

2.3. The functions of actin-like proteins

Next we explore the functional roles of the different actin proteins. Importantly, many of these proteins function in the regulation of cell shape, organization, and division across kingdoms.

Bridging long-length scales

One of the primary cellular roles of the actin family is to bridge molecular and cellular length scales. Both actin and ParM filaments have a persistence length of ~10 microns [27,79]. These rigid, long range tracks allow micron-sized cells a means to establish or read out long-range order. When filaments are crosslinked together or are coupled to other cellular systems, they can be used to exert control over much greater distances.

Importantly, the mechanism used by actin-related filaments to bridge length-scales also differs depending on cell size [80]. In large eukaryotic cells, components are usually trafficked along internal highways of parallel bundles of actin filaments by processive myosin motors [19,54,81,82]. In prokaryotes, where molecular motors have not been found, cargo (such as plasmids) is usually trafficked by the association of the object with the growing ends of one or more actin-like filaments, which are stiffer at short length-scales [14,74,75,83].

Filaments can be used to organize cell space in several ways. Long and stiff filaments (or bundles of filaments), when elongating within a confined space with a simple geometry (like a cylinder) will orient along its long axis, growing until the filament reaches its tips. Actin-related filaments have been shown to function in this way [15,22]. In Ashbya gossypii, for example, cell wall material is trafficked to the growing tips of the hypha along bundles of actin filaments [17], and in budding yeast, actin-mediated transport plays a critical role in moving material from the mother to the emerging bud [18-20,84]. In a similar way, actin-related proteins drive the segregation of large cellular objects during bacterial cell division [85]. Thus, MamK is responsible for the partitioning of

magnetosomes into daughter cells [14] and in distributing them along the bacterial long axis [16], while bidirectionally growing ParM filament bundles segregate the R1 plasmid to ensure that each daughter cell inherits a single copy. Similar “end finding” strategies have been observed for other plasmid segregators, but with subtle kinetic twists [65,71].

Cell mechanics

When crosslinked, actin-related filaments can also generate higher order networks, whose material properties can be precisely tuned. In this state the actin cytoskeleton plays a critical role in setting the physical properties of cells. Thus, in animal cells, which have a flexible form, actin has been shown to define the dynamics of the nucleoplasm [86] and the cytoplasm [87]. The actin meshwork that underlies the plasma membrane also plays an essential in providing the plasma membrane of animal cells with structural support and the capacity to undergo dynamic changes in shape, e.g. to form filopodia / phagosomes [88,89].

Force generation through polymerization dynamics

The polymerization of filaments from monomers can also exert force. In the context of actin networks, this has been proposed to be achieved in several ways [90]. The most easily understood is a Brownian ratchet mechanism [91] where polymerizing filaments iteratively rectify fluctuations in the position of a load, using ATP hydrolysis as a way to keep the polymer system far from equilibrium. In *Ec*ParM, insertional polymerization of monomers at ParM filament ends causes directional plasmid

movement [22]. Actin depolymerization has also been shown to generate mechanical pulling forces *in vitro* [21].

Cell division

Filamentous actins are an essential component of the cell division machinery across kingdoms [92]. Actin is an essential component of the cytokinetic-ring in eukaryotes [25]. In the presence of multivalent motors, like Myosin II, which can bind multiple actin filaments at the same time, the forces exerted by the motors can cause an actin network to contract or expand [93]. However, since individual filaments tend to buckle under compressive forces, this typically leads to a break in the symmetry, so that crosslinked actin-networks tend to contract under the influence of Myosin motors [94]. When organized into a ring, these contractile forces cause constriction as aligned filaments slide past one another [25] as seen in the cytokinesis of eukaryotes. At a larger scale, similar rings, encompassing many cells, provide the contractile forces required for wound-healing (flies) and tissue morphogenesis. Importantly, in these cases, the tension generated around the ring can function to drive changes in cell/tissue shape and/or to provide a template with a simple uniform geometry that can then be used to guide and organize external processes such as cell wall synthesis during dorsal closure in fly embryos [95].

Interestingly, a similar process occurs during prokaryotic cell division, where the bacterial actin FtsA polymerizes at the septum of the dividing cell [96] along with the tubulin homolog, FtsZ. FtsA treadmills around the ring along with the bacterial tubulin, FtsZ [97,98], although FtsA's ATPase activity has no effect on this behavior [98]. It is unclear

if the treadmilling of FtsA/FtsZ is responsible for generating forces that contribute to division, as has been suggested for actin and Myosin II in yeast [24], or if the filaments themselves bend the membrane. Alternatively, it may be the creation of new cell wall guided by the FtsA/FtsZ associated enzymes that generates the force required to pinch the cell into two [99]; a process that also contributes to division forces in eukaryotes with a wall [100].

Sensing and remodeling cell shape

Although actin-like filaments have a persistence length of around 10 microns, at smaller length scales short polymers can be used to sense local curvature. In many systems their role in generating long-range order is due to their associations with membranes. The best example of this is MreB. MreB filaments are highly curved [30,55]. This enables filaments to read out the differences in the local curvature of the bacterial plasma membrane, causing them to align along the circumference of the cell, perpendicular to the straight, long cell axis (Hussain et al, in review). This orientation guides the associated cell wall synthetic machinery so that new cell wall material is laid down in hoops, ensuring that bacteria retain their rod shape against the internal pressure [23]. Interestingly, an archaeal actin, crenactin, is present in rod-shaped crenarchaea, but is absent from many more pleomorphic archaea, suggesting that it could perform a similar function in guiding archaeal cell shape [12], though it is likely that this would be via a different mechanism.

While eukaryotic actin also controls the shape of many eukaryotic cells, its function differs depending on whether or not they have a cell wall (fungi, plants and

animals). Thus, the cortical actin network generates the forces required to shape animals cells (see above), whereas in fungi and plant cells actin filaments define sites of cell wall insertion. So, for example, fission yeast cells that have defects in actin dynamics lose their normal elongated shape to become round [101], just as bacteria do that lack MreB.

Macromolecular Scaffolds

Finally, actin and actin-related filaments serve as scaffolds for larger complexes. An example of this is the dynactin complex. This complex links vesicles to the motor dynein, and contains short filaments of actin and ARP1 [102]. This filament helps to orient and adapt various organelles for microtubule mediated transport [103,104]. The polarity of this actin-like filament ensures that dynein engages with its cargo in the proper orientation. This type of role for short filaments may be more widespread, as actin and ARPs have recently been suggested to play a similar function in the regulation of eukaryotic transcription [105].

2.4. The evolution of actin-family proteins

Structural similarity and shared ancestry among actin-like proteins

To what extent do the related structure and function of prokaryotic and eukaryotic actins reflect the action of selection, and to what extent does it reflect the constraints imposed by common ancestry? To answer this, we will take a look at the evolution of the actin-related proteins.

Since primary sequence evolves faster than tertiary structure [106], structural similarities can preserve evidence of common evolutionary ancestry even after sequences have diverged so far that they are not longer detectably similar. At the same time, because the function of a protein is directly specified by its three-dimensional structure, sequences with different evolutionary origins might converge on (that is, evolve towards) similar structures in response to selection for a particular function, although this possibility is thought to occur only rarely. Given the diversity of functions performed by actin-like proteins across the tree of life – from filament formation to chaperoning other proteins and catalyzing the rate-limiting step of glycolysis – it seems reasonable to ask whether their shared fold is the result of convergence or common ancestry.

Here, to provide a formal test of the common ancestry of all protein families sharing the actin fold, we built hidden Markov models from sequence alignments of each family, and aligned these profiles using the hhalign program in the HH-suite package [107]. These comparisons are more sensitive than a standard BLAST search because they incorporate information about the evolutionary conservation, and therefore functional importance, of the residues in each protein family. However, because they are still based on comparisons of primary sequences, they provide a straightforward way to distinguish common ancestry from convergence: given the rate at which sequences evolve, and the biochemical degeneracy of the 20 amino acid alphabet, better-than-chance levels of sequence identity between two families provides good evidence of descent from a common ancestor or of a common structural constraint. The results provide strong evidence of significant sequence similarity, and

therefore common ancestry, among many of the actin-like protein families, consistent with their structural similarities, and demonstrate that filament-forming and monomeric actin-like families share a common ancestor. Note that a lack of significant similarity does not rule out an evolutionary relationship between two families, particularly in light of the structural similarities among all these families.

Evolution of polymerization in the actin superfamily

Having established the common ancestry of actin-like proteins, we next inferred a phylogenetic tree to determine the relationships among the subfamilies, and to resolve the number of times that polymerization has evolved. Since homology (evolutionary relatedness) among sequences is a prerequisite of phylogenetic analysis, we limited our tree inference to the set of actin-like families for which all pairs show significant pairwise sequence similarity (with an E-value cutoff of $< 10^{-5}$); this set included the polymers Actin, MreB, FtsA, ParM, and PilM; and the enzymes DnaK, Type II secretion protein L, diol dehydratase reactivase, peptidase M22, BcrAD-BadFG, glutamate mutase, 2-hydroxyglutaryl-CoA-reductase, EutA, FGGY carbohydrate kinase, AnmK, GDA1-CD39, and PPX GPPA phosphatase. HMMs built from representative sequences for each of these subfamilies were aligned against each other using the E-INS-i mode in mafft [108], and poorly-aligning positions were identified and removed using trimAl [109]. Bootstrapped maximum likelihood phylogenies were inferred under the LG+C20+G+F model in IQ-Tree [110], providing a model that accommodates variation in exchange rates among amino acids (LG) [111,112], variation in equilibrium composition frequencies across sites

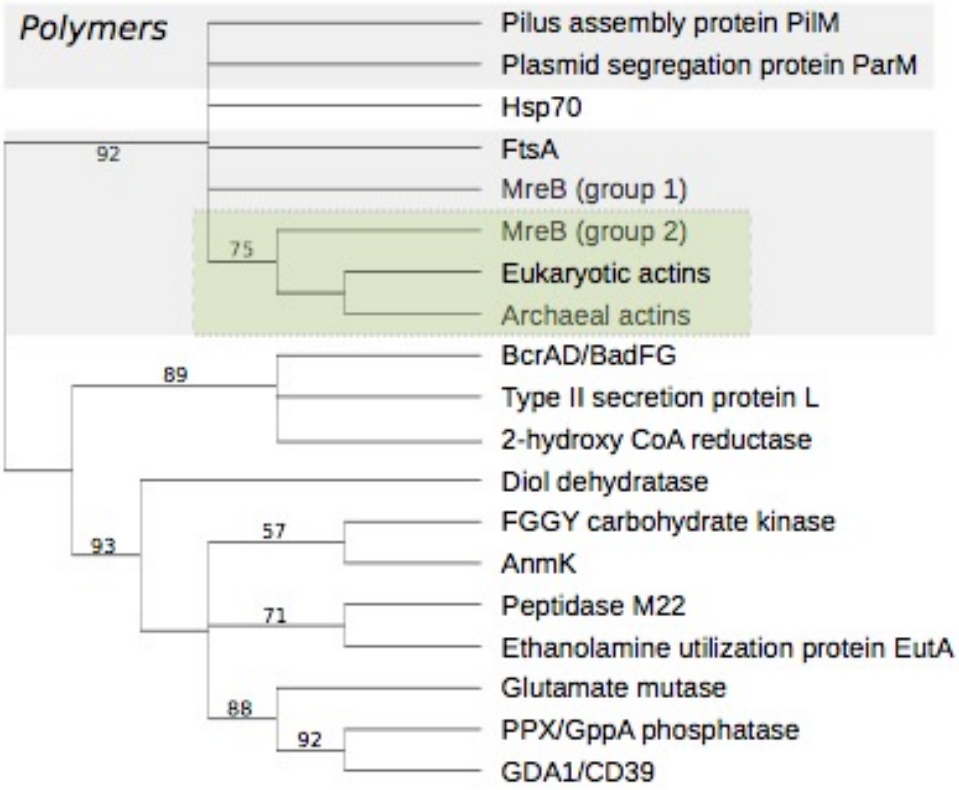


Figure 2-4: Evolutionary relationships among members of the actin superfamily.
The phylogeny was inferred under the LG+C20+G+F model in IQ-Tree (Nguyen *et al.*, 2015), and branch supports are maximum likelihood bootstrap values. All polymer-forming actins cluster together in the tree, suggesting that the capacity to form filaments arose once during the evolution of the actin fold. Interestingly, the molecular chaperone DnaK/Hsp70 falls within the polymer lineage, suggesting that it may have evolved from an ancestral polymer-forming actin by loss of polymerization. The actin-like proteins of Crenarchaeota and the Asgard archaea (indicated as 'Archeal actins') are the closest prokaryotic relatives of bona fide eukaryotic actins, consistent with a close relationship between the Asgard superphylum and the archaeal host cell for the mitochondrial endosymbiont (Braun *et al.*, 2015; Spang *et al.*, 2015; Zaremba-Niedzwiedzka *et al.*, 2017). The Archeal actins and the Eukaryotic actins together form a lineage that is most closely related to the cell shape determining protein MreB, found in rod-shaped bacteria. Our phylogenetic analysis suggest that actins are nested within the diversity of MreB proteins (green box), although statistical support for the specific relationship is low. We have depicted the actin tree as unrooted: the divergences between superfamily members are ancient, with some likely occuring before the time of the last universal common ancestor. It nonetheless seems reasonable to suppose that the polymer-forming actins evolved from an ancestral monomer, suggesting that the root may lie somewhere among the monomeric actins. Among modern actin-like protiens, only two proteins, both monomeric --- Benzoyl-CoA Reductase (BcrAD/BadFG) and hydantoinase (not depicted in this tree, due to high levels of sequence divergence) --- are broadly distributed in Bacteria, Archaea, and amongst some Eukaryotic lineages, and may represent good candidates for the oldest extant members of the superfamily; both also perform functions that may have been important during the evolution of early life.

(C20) [111], and variation in evolutionary rates across sites (G, discrete gamma distribution) [113].

The maximum likelihood phylogeny (Figure 2-4) robustly resolves the filament-forming actin families as a clan, or single lineage, on the tree, with the split between polymers and monomers receiving good (93) bootstrap support. This result strongly suggests that polymerization evolved once during the evolution of actins, and that all filament-forming actins are descended from an ancestral polymer.

Given the lack of an obvious outgroup, it is difficult to unambiguously root the actin tree. It nonetheless seems reasonable to assume that polymers would evolve from monomers. However, monomeric actins do not form a single lineage in our tree: a striking, and perhaps unexpected, result of our analysis is that the DnaK family of molecular chaperones is nested within the polymer clade, separated from the other monomeric actins with good bootstrap support. Although our unrooted tree does not enable us to exclude the possibility that the root of the entire actin family lies within one of these two groups, the high level of sequence similarity between DnaK and the filament-forming actins is, in our view, more readily explained by an origin of this molecular chaperone via a secondary loss of polymerization. Consistent with this view, it is interesting to note that while none of the actin family polymers can be confidently mapped to the last universal common ancestor, some of the monomeric forms - including Benzoyl-CoA Reductase and hydantoinase - are broadly distributed in modern Bacteria, Archaea, and some Eukaryotes (Table S1), suggesting that the radiation of the monomeric actins may have pre-dated the deepest divergence among extant cellular life.

The Prokaryotic Origins of Eukaryotic-Type Actin

Given that eukaryotes appear to have arisen from the symbiosis of cells from both bacterial and archeal lineages, both of which contain actin-like proteins, the origins of eukaryotic actin have long remained an important question in the field. Fortunately, this mystery now appears to have been resolved. Recent discoveries have identified close homologues of the eukaryotic-type actins in the genomes of members of the TACK and "Asgard" superphyla [28,114,115]; lineages of Archaea that appear to be most closely related to eukaryotes. Furthermore, functional studies of "crenactin" filaments from *Pyrobaculum* indicate that they are nearly identical to the filaments formed by eukaryotic actin [114,116], and at least one gelsolin-like protein (arcadin-2) seems to affect crenactin analogously to how gelsolin affects actin [117]. While functional studies have not yet been performed on Asgard actins, both crenactins and Asgard actins cluster with the actins of canonical eukaryotes in our phylogenetic analysis (Figure 2-4). Thus, direct evolutionary precursors of eukaryotic actin appear to have first evolved in the archaeal ancestors of eukaryotes. Furthermore, it is interesting to note that the metagenomic assembly of *Lokiarchaeum*, the first member of the Asgard group to be discovered, also encodes potential homologues of the actin modulators gelsolin and profilin. If these proteins prove to modulate the polymer kinetics of a lokiarchaeal actin *in vitro*, it would suggest that at least some of the functions and regulatory mechanisms of eukaryotic actins evolved prior to the split between eukaryotes and Archaea. While the cellular function of these archaeal actin-homologues remains a matter of speculation, this is an exciting time for research into the origin of the eukaryotic cytoskeleton. It is clear that

much more functional information is needed to answer the following questions: Do Asgard actins form actin-like filaments? If they do, what are the kinetic properties of archaeal actin polymers? Do the predicted actin modulators present in the lokiarchea modulate lokiarcheal actin? Does crenactin perform similar cellular functions to that of eukaryotic actin?

2.5. Conclusions and Speculation

In this review, we have examined the similarities in the structure and function of actin-like proteins across the tree of life. Importantly, an evolutionary analysis suggests that the known polymer forming actin-like proteins from bacteria, archaea and eukaryotes have all arisen from a single ancestral polymer forming protein. This makes actin and its relatives an ancient protein [118]. It also suggests that the actin fold provides an important and thus evolutionary conserved scaffold for the formation of dynamic filaments: the actin fold binds ATP and associates with other monomers in a head to tail fashion, generating protofilaments, which associate through lateral contacts to form multi-stranded polymers. Importantly, much of the structural variation that appears to underpin changes in the biological function of actin across domains is based on changes in lateral contacts between protofilaments [30], yielding changes in twist, packing, curvature, and polarity. In addition, it is clear that there are wide differences in the dynamics of polymer formation and disassembly across the clade. More work needs to be done to survey the variations in kinetic properties of different actin-like sub-families (MreB, actin, MamK, ParM etc) to understand how subtle differences in these similar folds leads to such dramatically different behaviour. For the one well studied

case of ParM, a weakening of the cross protofilament contacts occurs when filaments are ADP bound [31]. In addition, our analysis motivates a systematic investigation of whether DnaK can form polymers, or whether it has lost that capability during evolution.

It seems remarkable that so many of the cytomotive filaments used to generate intracellular forces, to order cellular organization and drive cell division across living systems are based on this single scaffold. This reflects the role of evolution as tinkerer. It may also point to the special properties of the actin-fold within the polymer forming enzymes. Actin couples ATP hydrolysis and phosphate release to large-scale changes in conformation, altering monomer-monomer contacts in the context of a filament. Of course, there are instances in which biological polymers appear to have evolved from non-polymerizing proteins. These include CTP synthetase, sickle cell hemoglobin and MSP from *C. elegans* [119-121]. However these polymers appear to lack the functional flexibility of actin. Sickle cell haemoglobin forms long static polar filaments that distort cells, and CTP synthase and MSP form non-polar filaments. CTP synthase polymerization appears to inhibit the enzyme activity in most cases, and MSP polymerization cannot do work since it is not coupled to a chemical reaction. As a result, its directed polymerization must be driven by the action of an external pH gradient.

In sum, actin is a remarkable protein family. Polymer forming actin appears to be an ancient protein whose ability to generate long range order and force has been used again and again in evolution to perform many of the tasks that help give cells their dynamic shape and internal organization.

Chapter 3

Glucokinase polymerization regulates enzyme activity in response to environmental shifts

3.1 Introduction

3.2 Glk1 Forms Polymers That Inhibit Enzyme Activity
Substrate binding induces polymerization of the glucokinase Glk1
Glk1 polymerization inhibits Glk1 enzyme activity

3.3 The Structure of Glk1 and Glk1 Filaments
Glk1 polymers are two-stranded, anti-parallel helices

3.4 Glk1 Polymerization Evolved Independently of Other Actin-Related Polymers
Glk1 polymerization arose within the ascomycetes

3.5 Disrupting Glk1 Polymers
Glk1 activity and polymerization are separable
Non-polymerizing Glk1 reduces the fitness of yeast growing on sugars

Attributions

This chapter will be submitted as a paper. The authors on the paper will be me (Patrick R. Stoddard), Quincey Justman, Eric Lynch, Justin Kollman, Ethan Garner, and Andrew Murray. I have performed the majority of the experiments in this chapter. The initial observation of Glk1-GFP forming polymers in cells was done by Quincey Justman. The cryo-EM reconstruction of Glk1 filaments was performed by Eric Lynch and Justin Kollman. I wrote the draft of this work and it has been edited by me, Ethan Garner, and Andrew Murray.

Abstract

The actin fold is utilized across the tree of life for various functions. While a subset polymerize into filaments, most members of the actin superfamily, including carbohydrate kinases, do not. Here we find that the *S. cerevisiae* glucokinase, Glk1, forms two-stranded filaments, with novel interactions between its monomers. Glk1 polymers can be observed *in vivo*, forming and depolymerizing when sugars are added or withdrawn. Glk1 polymers form when the substrate/product bound form exceeds a threshold concentration, and polymerization inhibits enzymatic activity, indicating that polymerization acts as a 'clamp' on total Glk1 activity. A point mutation eliminating Glk1 polymerization alleviates enzymatic inhibition. Glk1 polymerization serves an important *in vivo* function: yeast containing non-polymerizing Glk1 have reduced competitive fitness growing on sugars and have reduced viability upon reintroduction to glucose. Glucokinase polymerization arose within the ascomycete fungi is conserved across a group of highly divergent (> 150 mya) yeast. Our data demonstrates that Glk1 polymerization evolved independently from other actin-related filaments and allows cells to rapidly modulate glucokinase activity in the face of rapid shifts in nutrient concentration.

3.1. Introduction

The actin superfamily is a diverse group of structurally related proteins that are ubiquitous across the domains of life [12,55,122]. While several of the actin fold sub-families form linear polymers, the proteins that act as catalytic enzymes, such as hexose kinases and the Hsp70/DnaK family of chaperones, do not [39]. It is unclear if the ubiquity and diversity of actin-like filaments means that the superfamily's fold is especially suited to polymerization. Phylogenetic analysis suggests that the known actin-related polymers are diverged from an ancestral, polymer-forming protein [29].

Free living cells are subject to rapid environmental shifts, many occurring too quickly for them to respond by changing their proteome. Failure to respond appropriately reduces competitive fitness and can kill cells. Cells use a variety of mechanisms, including allosteric and post-translational regulation, to maintain homeostasis until they can alter their composition to match their environment.

Some enzymes form conglomerates, self-associating into filaments or phase-separated gels. This behavior is thought to regulate enzyme activity or to act as a 'storage depot' for temporarily unnecessary proteins [123]. Self-assembly may serve as a sensitive and tunable way to control enzyme activity when polymerization depends not only on an enzyme's concentration, but also the concentration of its substrates and products. Polymerization of enzymes has been proposed as a dynamic strategy to

regulate flux through certain reactions [124], as a way to store enzymes during starvation [125], or to measure and signal key information about the cell state.

Here we find that in the budding yeast *S. cerevisiae*, the actin homolog Glk1, a glucokinase, forms filaments in response to its substrates. Polymerization of Glk1 has evolved independently of polymerization of actin and other closely related filaments. Filament formation acts as an activity clamp, setting an upper limit to the rate of the entire Glk1 pool. This activity limit protects cells from toxic overactivity during a glucose pulse.

3.2. Glk1 Forms Polymers That Inhibit Enzyme Activity

Substrate binding induces polymerization of the glucokinase Glk1.

In the budding yeast, *S. cerevisiae*, the first step in glycolysis is catalyzed by three enzymes, two hexokinases (Hxk1 and Hxk2) and one glucokinase (Glk1). The relative amount of these enzymes changes depending on cellular state and environmental conditions. Glk1 expression is induced in the absence of glucose and suppressed in glucose-rich conditions [7] (Figure 3-2-1 A). To examine Glk1 localization, we created a Glk1-GFP fusion expressed from the native locus. In the absence of glucose, Glk1-GFP is cytoplasmic and diffuse. However, when glucose-starved cells are reintroduced to glucose, Glk1-GFP assembles into large filament bundles. These filaments rapidly disassemble upon washing out glucose Hxk1-GFP and Hxk2-GFP do not exhibit this behavior (Figure 3-2-1 B and 3-2-2 A-B). We also observe Glk1-GFP polymerization in response to mannose, glucosamine, fructose, and galactose (Figure

3-2-1 C-D). The polymers persist while sugar is present and cell growth and division dilute Glk1 below a critical concentration.

To understand what triggers Glk1 filament formation we purified the enzyme and studied its polymerization *in vitro*. Many of the previously observed enzyme condensates form due to decreased cytosolic pH induced by stress or starvation [125]. We found that Glk1 does not polymerize in response to pH changes. Rather, Glk1 polymerization requires its enzymatic substrates, ATP and glucose, mannose, or glucosamine, or its products, ADP and sugar-6- phosphate. Modest polymerization also occurs with N-acetylglucosamine and N-acetylglucosamine-6-phosphate neither of which are substrate or product, but which have been shown to inhibit Glk1 activity [126] (2-2-3 A-C). While we see Glk1 polymers form in vivo upon addition of fructose or galactose, these sugars do not cause Glk1 to polymerize *in vitro*, suggesting the *in vivo* polymers assemble from the accumulation of glucose-6-phosphate, which is an intermediate in galactose metabolism and must be formed from fructose to supply UDP-glucose for cell wall synthesis [127].

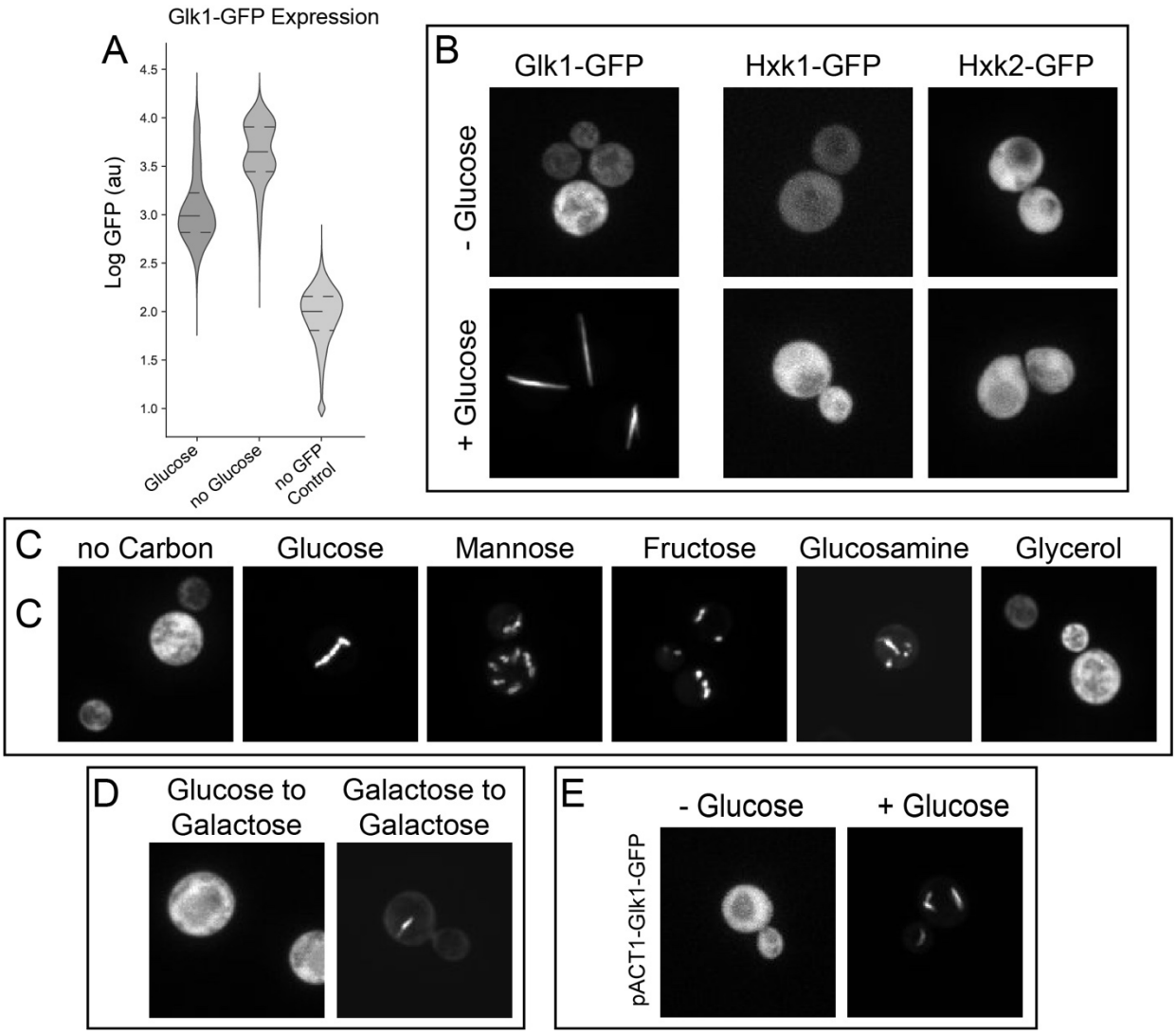


Figure 3-2-1. Glk1-GFP Polymerizes in Cells.
A: Flow cytometry of Glk1-GFP cells in mid-log in glucose (Glucose) and several hours after glucose depletion (no Glucose). Glk1 expression is regulated by carbon source, and is glucose repressed. Expression increases when growing on non-fermentable carbon sources. **B**: Fluorescence images of saturated phase cells expressing Glk1-GFP (left), Hxk1-GFP (middle), and Hxk2-GFP (right) both before the addition of glucose (top) or after (bottom). Neither hexokinase forms higher-order structures in response to glucose, but Glk1-GFP forms large filaments. **C**: Glk1-GFP forms filaments when saturated phase cells are refed sugars but not glycerol. **D**: Glk1-GFP forms filaments in response to refeeding galactose only in galactose conditioned cells. The galactose transporters are not expressed unless galactose genes are induced. **E**: Cells expressing Glk1-GFP under the actin promoter (strong, constitutive promotor) show filaments in the presence of glucose regardless of cell history.

Like the polymerization of actin, Glk1 polymerization shows a critical concentration. No Glk1 polymerization is seen below 2μM enzyme. Above 2uM, the amount of polymer increases with increasing Glk1, while the amount of unpolymerized Glk1 remains constant (Figure 3-2-4). This is consistent with the lack of polymers in exponentially growing cells because Glk1 expression is suppressed in the presence of glucose. We only observe Glk1-GFP polymers in exponentially growing cells if the protein is expressed under a strong, constitutive promoter (Figure 3-2-1 E). Consistent with the rapid polymer formation observed within cells, Glk1 polymerization in vitro is also rapid, occurring in a matter of seconds upon addition of ligands (Figure 3-2-2 C).

Glk1 polymerization inhibits Glk1 enzyme activity

Filament formation of enzymes can either activate or inhibit enzyme activity [50,128]. To determine the effects of Glk1 polymerization on its enzymatic activity, we measured the rate of glucose-6-phosphate (G6P) production as we varied Glk1 concentration. We find that beyond 2μM Glk1, there is no significant increase in the total rate of product formation (Figure 3-2-4). This transition corresponds to the critical concentration for Glk1 polymerization. Thus, polymerization inhibits Glk1 activity, and Glk1 activity is held constant by the monomer-polymer equilibrium at any enzyme concentration above the critical concentration.

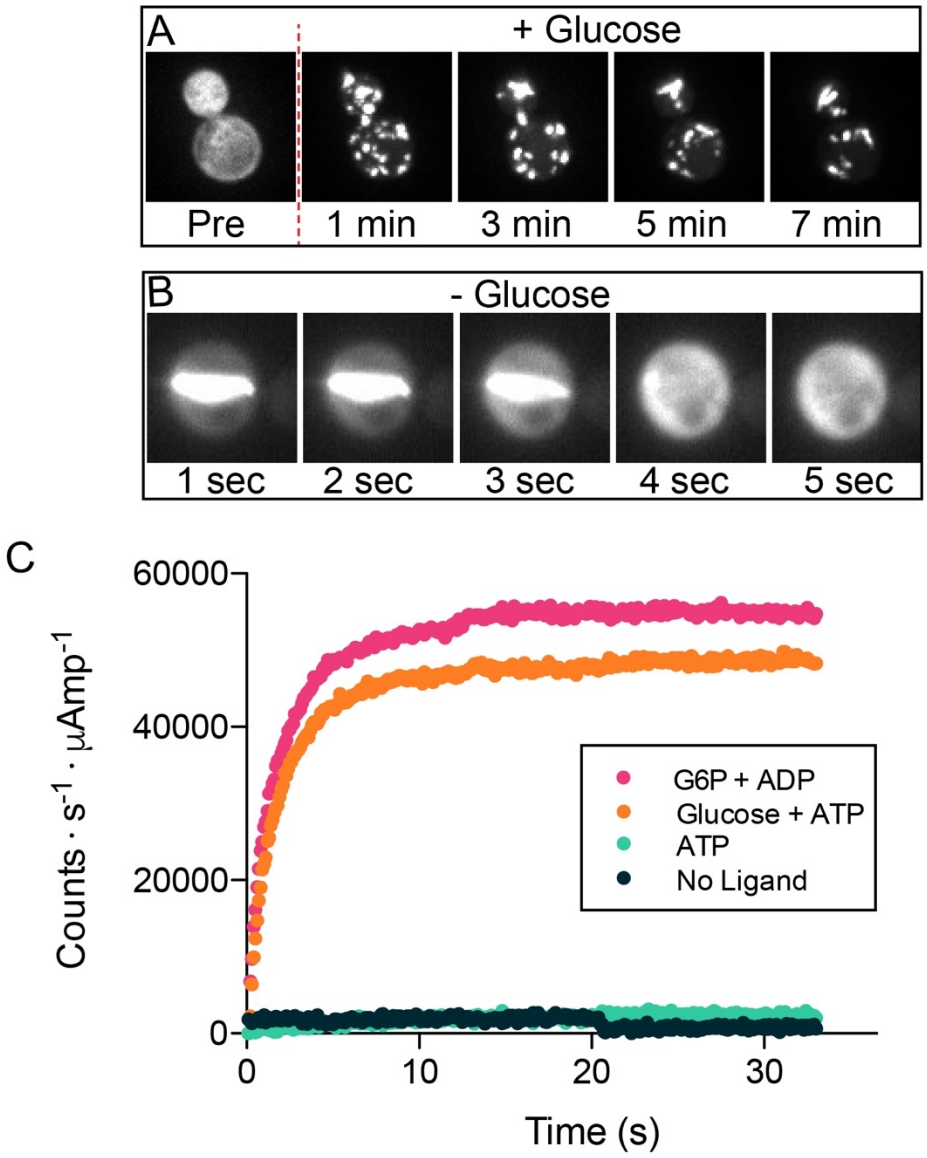


Figure 3-2-2. Glk1 Polymerization is Rapid and Reversible
A, B: Fluorescence images of cells expressing Glk1-GFP from a timelapse where glucose is added to the medium in a flow cell (top) or washed out (bottom). Glk1-GFP filaments form and dissolve rapidly in response to changes in the concentration of glucose in the medium. **C**: Glk1 polymerization assayed by light scattering. Glk1 polymerization was measured by following 90 degree light scattering at 315nm. Polymerization is rapid and reaches equilibrium in under 10 seconds.

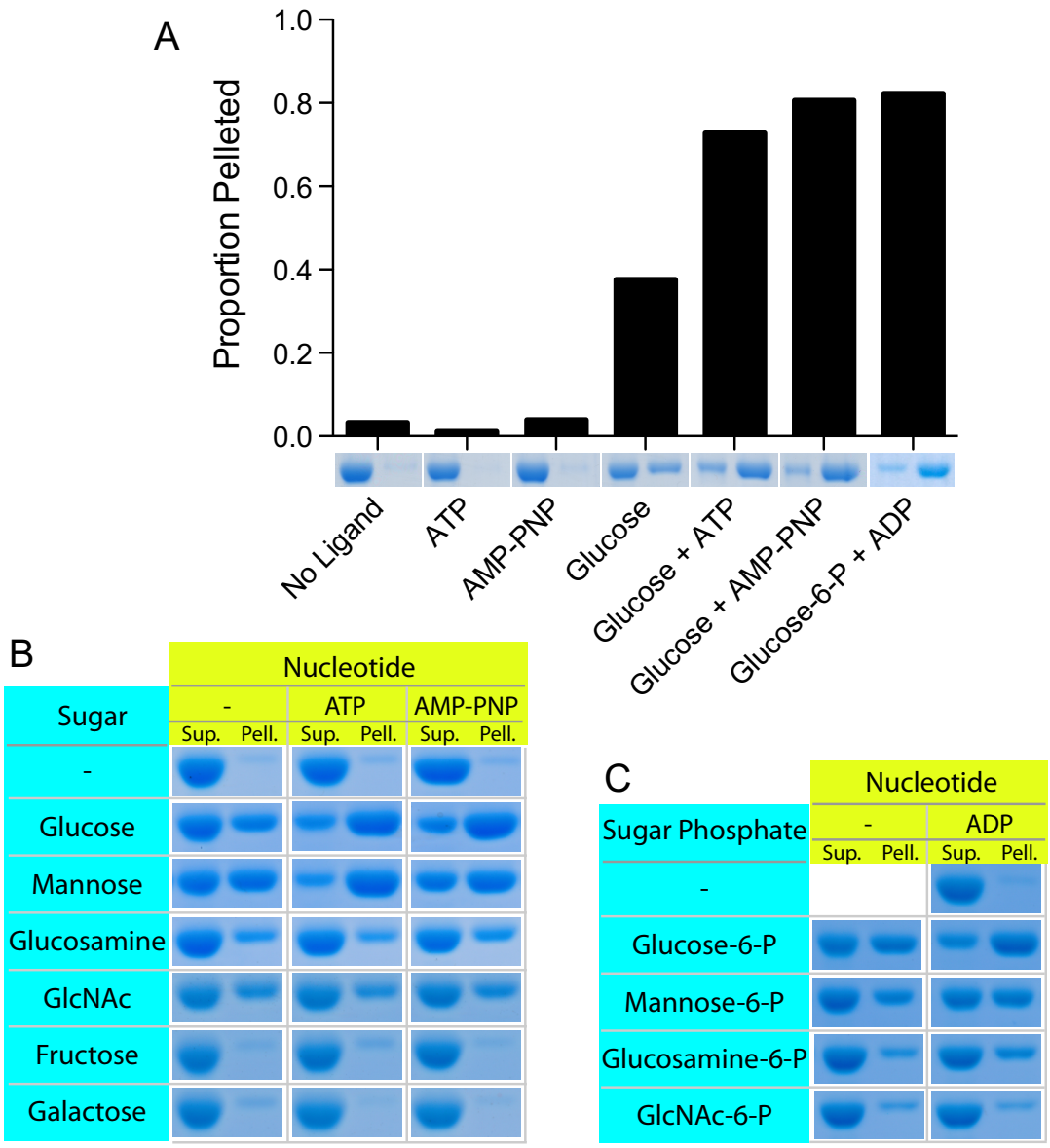


Figure 3-2-3. Glk1 Polymerizes in Response to Substrate or Product
A: Purified Glk1 was ultracentrifuged in the presence of different ligand combinations. The supernatant (left) and pellet (right) of each condition were run on SDS-PAGE. The relative band intensities are reported in the bar graph. Glk1 polymerizes in response to glucose, and polymerization is enhanced by the presence of ATP. Both the substrate- and product-bound state form polymer. **B, C**: Glk1 was untracentrifuged with substrates (left) and products (right). Only its substrate sugars and their products cause polymerization while other sugars do not.

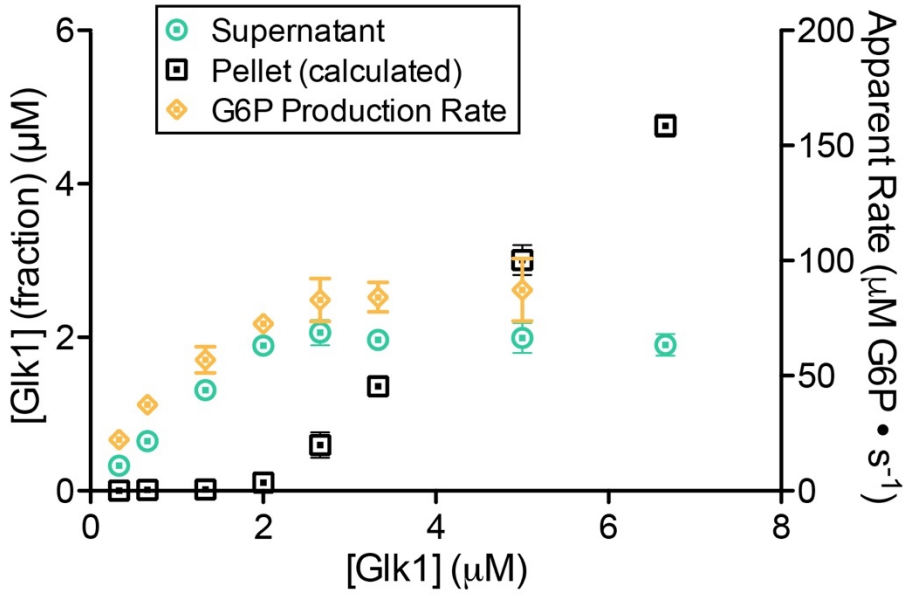


Figure 3-2-4. Glk1 Polymerization Inhibits Glk1 Activity
The concentration of purified Glk1 was varied in the presence of saturating glucose and ATP, and the mixtures were assayed for enzyme activity (glucose-6-phosphate production rate) or ultracentrifuged. The amount of Glk1 remaining in the supernatant was measured, and the amount in the pellet was inferred from this value.

3.3. The Structure of Glk1 and Glk1 Filaments

Glk1 polymers are two stranded, anti-parallel helices.

To determine if the Glk1 condensates were filaments, we imaged them with negative stain electron microscopy revealing short, helical filaments (Figure 3-3-1). This suggests the thick filaments observed in cells could be bundles that arise from depletion interactions from crowding [129], stabilization by filament binding proteins [130] or perhaps higher order structures arising from self-interactions of the fluorescent protein [131].

To understand the mechanism of Glk1 polymerization, we solved the Glk1 crystal structure (Figure 3-3-2 C and Table 3-3-1), and compared it to that of the yeast hexokinases. This comparison revealed interesting differences. The Glk1 N-terminal alpha helix extends beyond the end of the large subdomain and contains a solvent-exposed phenylalanine, while the N-terminal helix of Hxk2 ends before the end of the subdomain, and the remaining residues are predicted to be disordered by several protein disorder predictors (PredictProtein[132], IUPred2[133], DisEMBL[134]) (Figure 3-3-3 A). Also, in the hexokinases the C-terminal alpha helix extends beyond the end of the small subdomain, but in Glk1 the C-terminal helix ends within the small subdomain, creating a hydrophobic patch (Figure 3-3-3 B). These features were later found to be involved in important filament contacts.

We next used cryo-electron microscopy (cryo-EM) to build a model of Glk1 filaments with a resolution of 3.8Å (Figure 3-3-2 A) revealing that Glk1 forms two stranded, antiparallel filaments. The longitudinal interactions in the Glk1 polymer differ from the conserved interactions seen in polymers formed by the polymer-forming clade in the actin superfamily suggesting that the polymerization of Glk1 has evolved

independently of that of the polymer-forming clade; the N-terminal, solvent-exposed phenylalanine of one monomer apposes the hydrophobic patch at the C terminus of the next monomer [29,30] (Figure 3-3-2 B). The lateral associations between strands are mediated by loop 380-387 and loop 438-444. Notably, the loop containing residues 438-444 is smaller in the hexokinases.

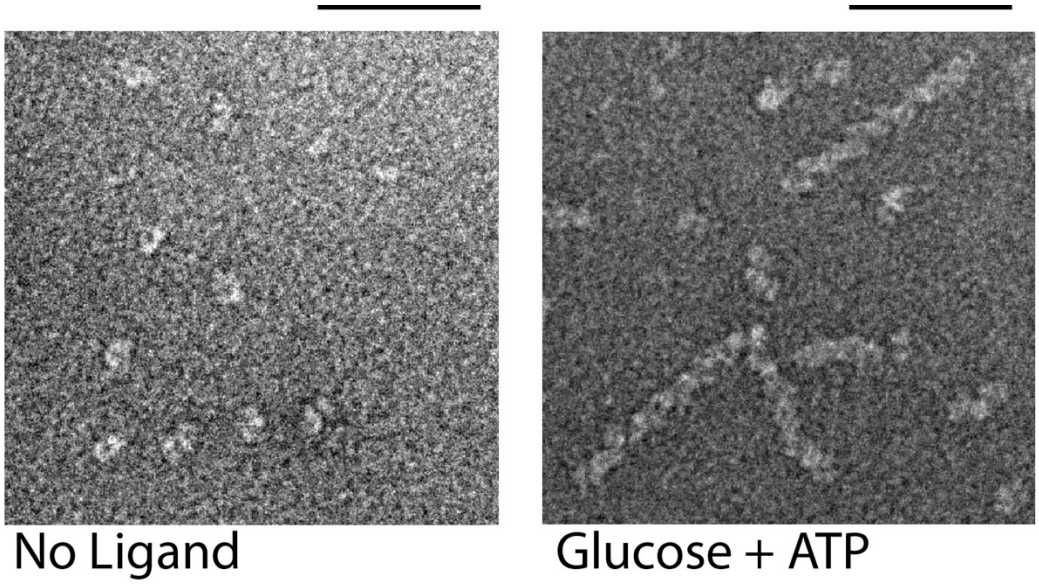


Figure 3-3-1. Glk1 Polymers are Short, Linear Filaments
Negative-stain electron micrographs of Glk1. 7.5μM Glk1 was applied of grids in the absence of ligands (left) or in saturating glucose and ATP (right) and stained with uranyl formate. Scale Bar = 50nm.

	Glk1
Data Collection	
Wavelength (Å)	0.9793
Resolution Range (Å)	155.8 - 3.59 (3.718 - 3.59)
Space Group	P 43 21 2
Unit Cell (a,b,c)	174.5, 174.5, 323.947
Unit Cell (α,β,γ)	90, 90, 90
Number of Crystals	1
Total Reflections	122662 (12052)
Unique Reflections	61332 (6026)
Redundancy	2.0 (2.0)
Completeness (%)	99.97 (99.95)
Mean I/σ (I)	8.28 (1.56)
R-merge	0.1003 (0.4831)
R-meas	0.1418 (0.6832)
R-pim	0.1003 (0.4831)
CC 1/2	0.988 (0.638)
Refinement	
Resolution Range (Å)	155.8 - 3.59 (3.718 - 3.59)
R-work	0.2652 (0.3358)
R-free	0.3001 (0.3565)
Number of Atoms	22953
Protein	22888
Ligands	65
Protein Residues	2944
Ramachandran Plot	
Favored (%)	98.73
Allowed (%)	1.27
Outliers (%)	0.0
RMS (bonds)	0.004
RMS (angles)	1.47
Average B-factor	89.65
Protein	89.69
Ligands	76.21
Number of TLS Groups	20

Table 3-3-1. X-Ray Crystallography Statistics
Statistics for data collection and refinement of the Glk1 crystal structure. Values in parentheses indicate the value for the highest resolution shell.

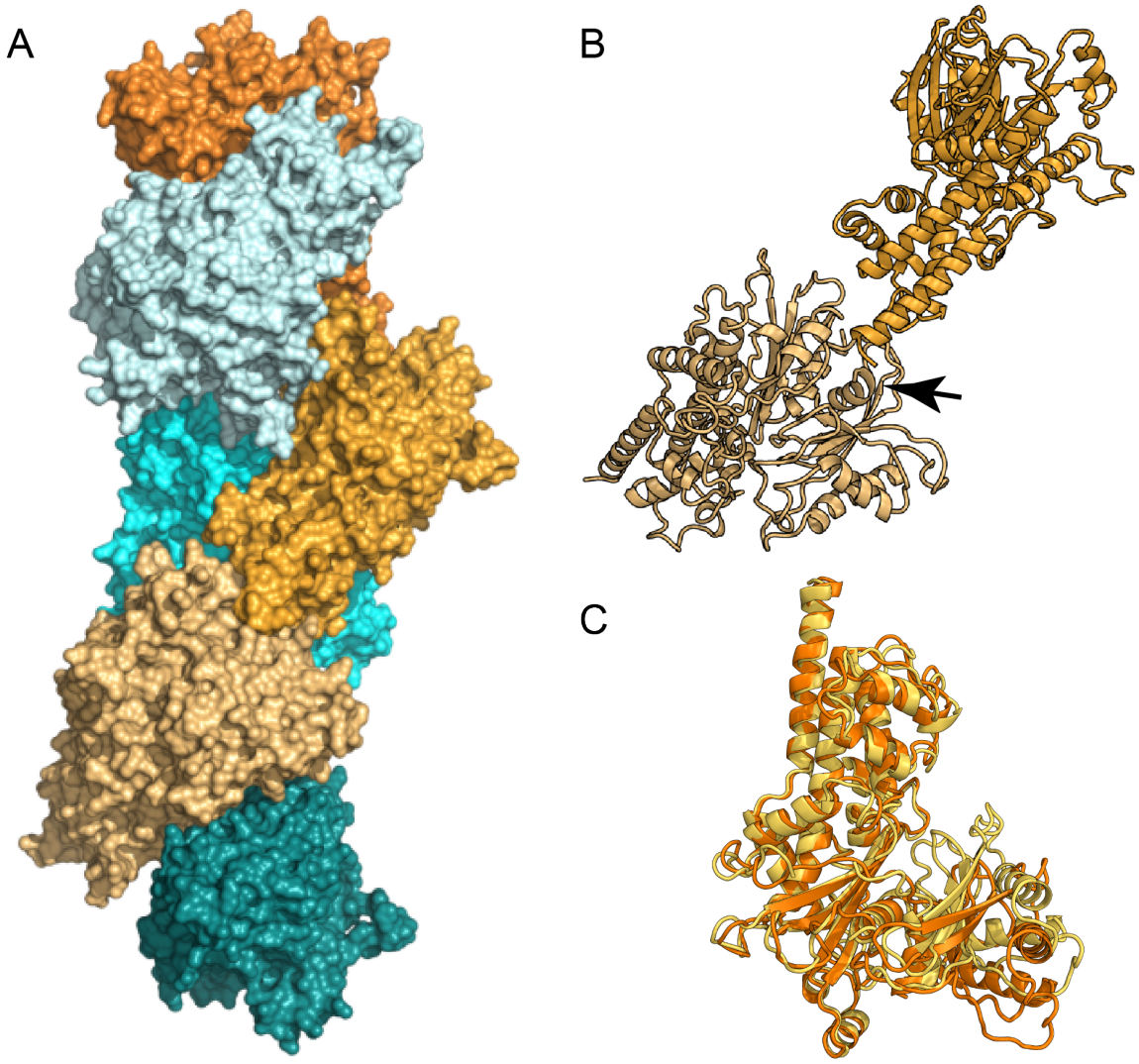


Figure 3-3-2. The Structure of Glk1 Filaments
A: Surface representation of Glk1 filaments. Filaments are antiparallel helices. **B**: Ribbon diagram of two monomers along one strand. The contacts are different than in other actin-fold protein polymers. The N terminus of one monomer makes close contact with the C terminus of the next monomer (arrow). **C**: Superimposition of Glk1 crystal structure (orange) with the Glk1 filament structure (yellow). The crystal structure is in the apo state and the filament structure is closed. This is consistent with the conformational change seen in other glucokinases and hexokinases.

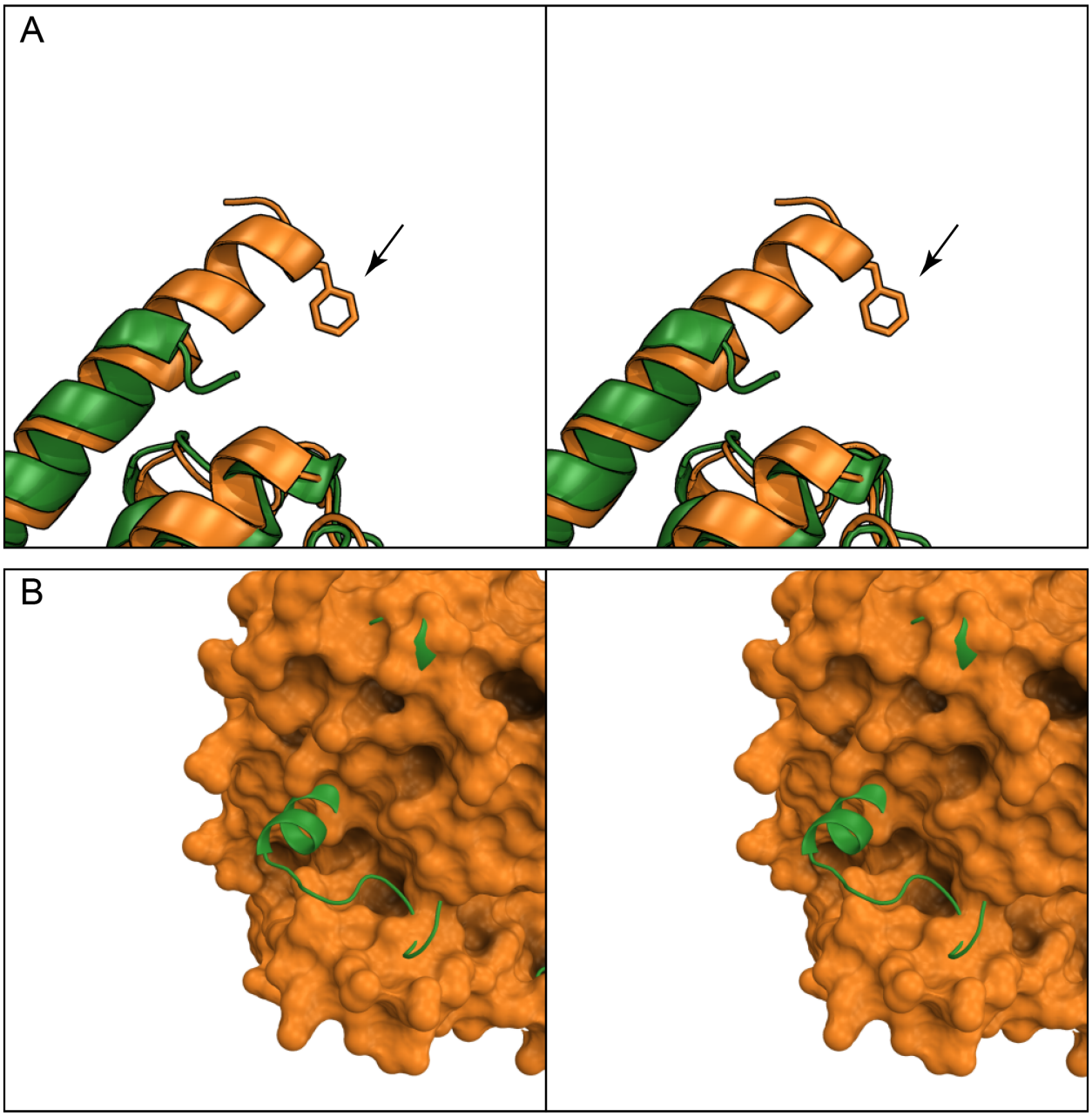


Figure 3-3-3. Glk1 N and C Terminus Differ from Hxk2
Glk1 (orange) and Hxk2 (green) (PDB ID: 1IG8, Kuser et al, 2000). **A**: Comparison of the N terminus of Glk1 and Hxk2. The N-terminal helix of Glk1 extends beyond the domain further than that of Hxk2. The residues not modeled at the Hxk2 N-terminus are predicted to be disordered by several protein disorder predictors. The helix of Glk1 ends in a solvent-exposed phenylalanine (arrow). **B**: Surface representation of Glk1 with cartoon diagram of Hxk2 showing the C terminus. The C-terminal helix of Hxk2 extends beyond the end of the domain while the C-terminal helix of Glk1 ends within the domain.

3.4. Glk1 Polymerization Evolved Independently of Other Actin-Related Polymers

Glk1 polymerization arose within the ascomycetes.

The two major extant fungal clades are the Ascomycota, which includes *S.cerevisae*, and Basidiomycota. Within the Ascomycota, the sequences of hexokinases and glucokinases fall into two groups: one contains the yeast hexokinases, Hxk1 and Hxk2, and the other contains Glk1. Within the ascomycetes, all members of the Saccharomycetaceae, excepting some members of the Kluyveromycetes, contained three conserved motifs in their Glk1 homologs that differentiate them from other ascomycete glucokinases and all ascomycete hexokinases: their N terminus (MSF(e/d)(e/d)LHK), a conserved loop (Glk1 residues 438-444, KRYHGEV) and their C terminus (LCALVA). To test if these motifs were predictive of polymerization in glucokinases and hexokinases, we purified several of the enzymes and tested their ability to polymerize. Of the enzymes tested, only glucokinases with conserved N and C-termini and conserved loop 438-444 polymerized in response to glucose and ATP (Figure 3-4-1 and Figure 3-4-2); the Glk1 ortholog from *Ashbya gossypii* contains all three motifs and polymerizes, whereas the protein from *Kluveromyces lactis*, which differs at all three motifs, does not.

These results argue that the ability of members of the Glk1 family to polymerize arose 150-200 million years ago, before the lineages leading to *S.cerevisiae* and *A. gossypii* diverged from each other [135].

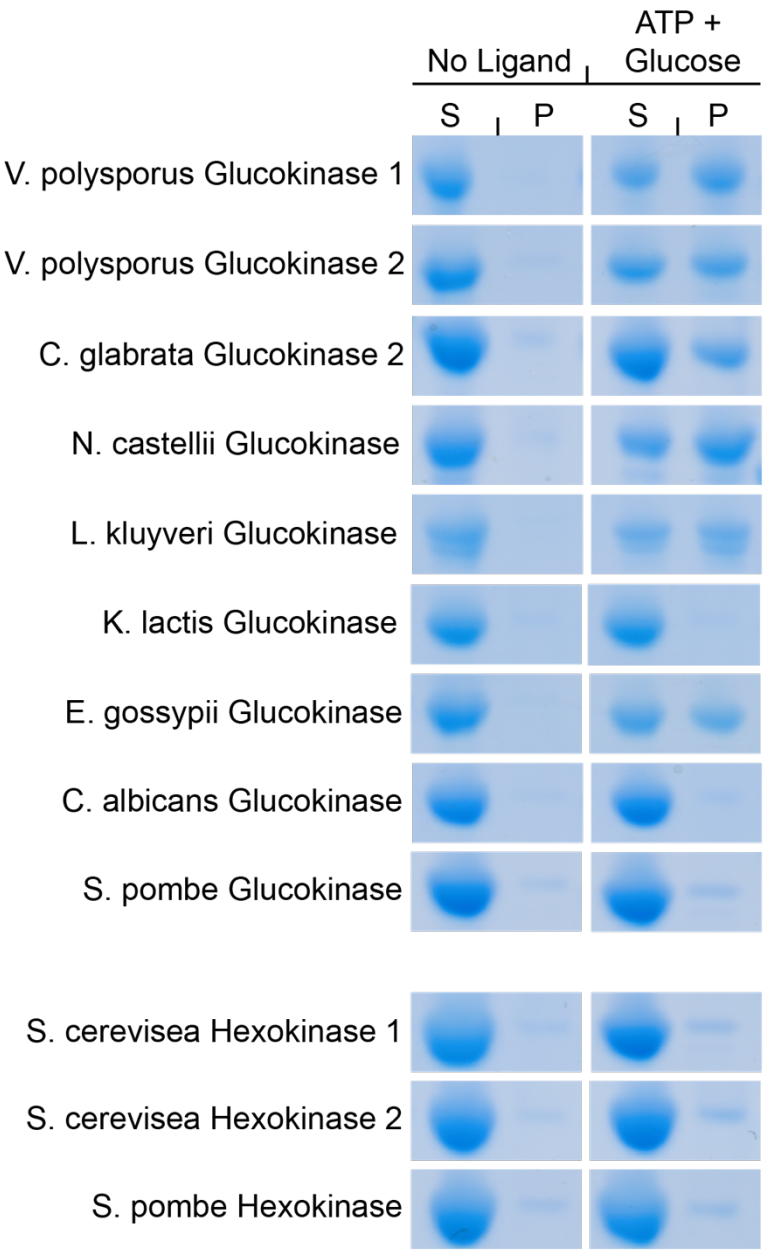
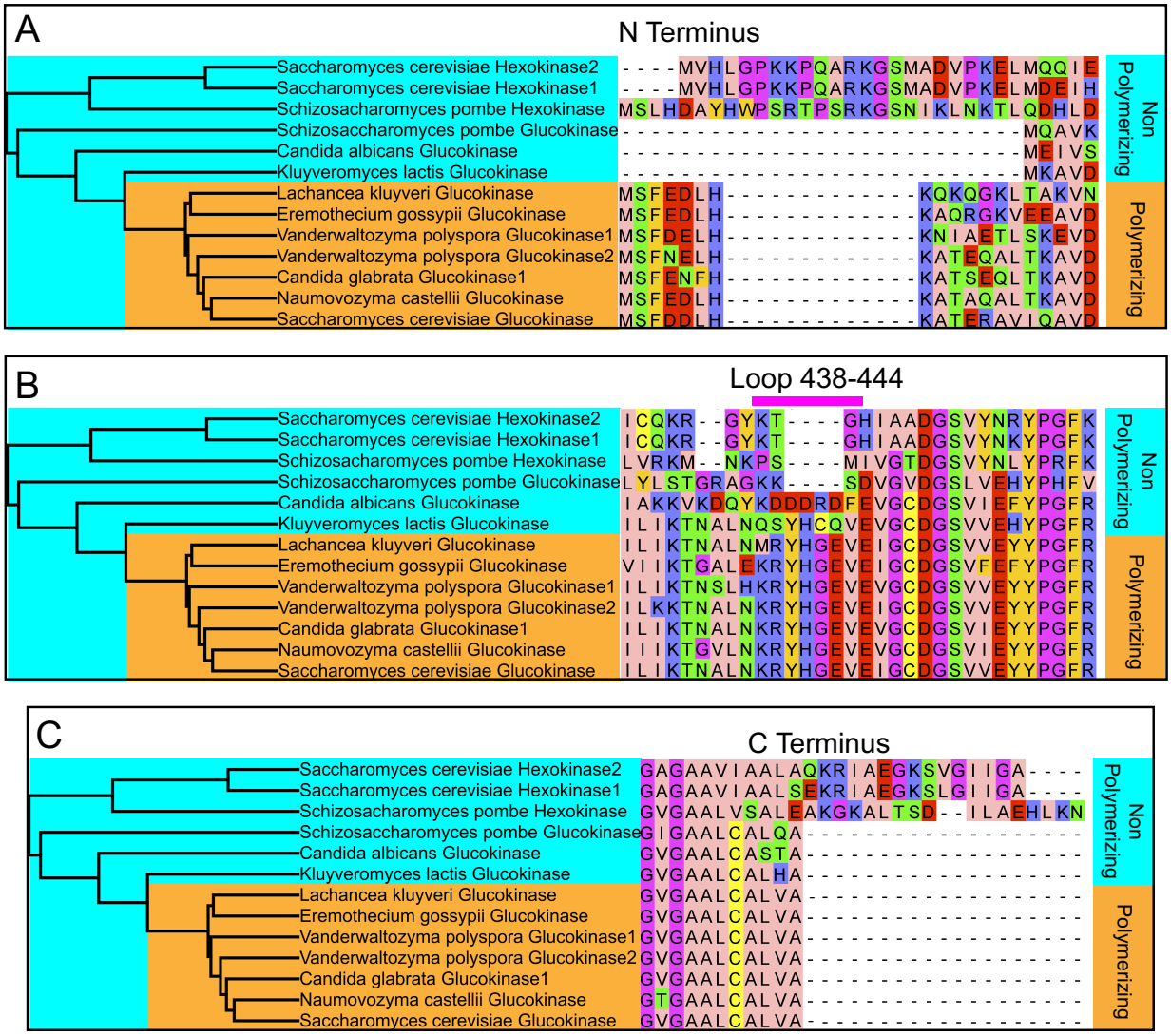


Figure 3-4-1. Other Glucokinases Also Polymerize
Glucokinases from other yeast were purified and tested for polymerization in response to glucose and ATP. Hexokinases from *S. cerevisiae* and *S. pombe* were also tested. Reactions were ultracentrifuged and the supernatant (left) and pellet (right) were run on SDS-PAGE.



3-4-2. Major Filament Contacts Are Conserved in Polymerizing Glucokinases
Alignment of all tested glucokinases and hexokinases with trees showing the relationship between proteins (Blosum62 neighbor joining). The N terminus (**A**), loop 438-444 (**B**), and C terminus (**C**) are conserved amongst polymerizing glucokinases but divergent in the non-polymerizing glucokinases and the hexokinases.

3.5. Disrupting Glk1 Polymers

Glk1 activity and polymerization are separable.

The conservation of polymerization-regulated glucokinase activity over millions of years, prompted us to attempt to disrupt Glk1 polymerization and examine its consequences on enzyme activity and cellular physiology. We used the contacts observed above to design a non-polymerizing Glk1 (NonPol) by mutating the N-terminal phenylalanine to serine: Glk1(F3S). Pelleting assays revealed that NonPol does not form polymers in response to glucose and ATP (Figure 3-5-1 A). Likewise, NonPol-GFP did not form visible polymer in cells upon glucose addition (Figure 3-5-1 B). Furthermore, assaying the rate of G6P production revealed NonPol is not only enzymatically active, but also lacks the concentration-dependent inhibition of enzyme activity exhibited by wild-type Glk1 (Figure 3-5-2).

We were also interested in distinguishing between effects from NonPol's lack of enzyme inhibition and the absence of Glk1 polymer. To separate these effects, we made catalytically dead Glk1 (CatDead) by mutating the catalytic lysine, creating Glk1(K182A) [136], and we combined these two mutations to create non-polymerizing, catalytically dead Glk1 (NonPolCatDead) as Glk1(F3S/K182A). CatDead forms polymers in response to glucose and ATP but does not generate G6P and, as expected, NonPolCatDead neither forms polymer nor produces G6P (Figures 3-5-1 and 3-5-2).

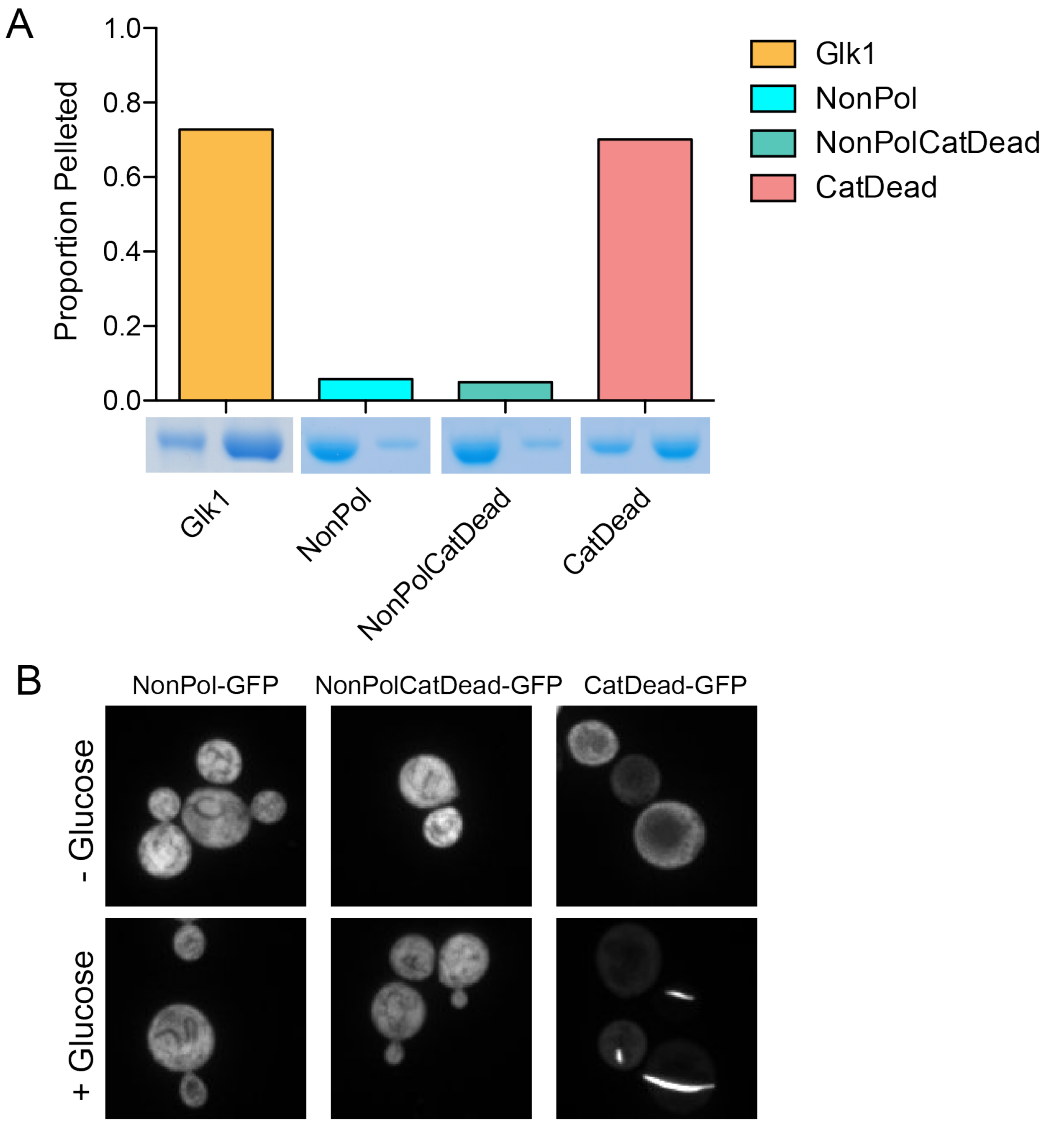


Figure 3-5-1. Glk1 Polymerization Can Be Disrupted and Does Not Require Catalytic Activity.

A: 5μM purified non-polymerizing Glk1 (NonPol, Glk1[F3S]), non-polymerizing catalytic-dead Glk1 (NonPolCatDead, Glk1[F3S,K182A]), and catalytic-dead Glk1 (CatDead, Glk1[K182A]) were ultracentrifuged in the presence of saturating glucose and ATP. The supernatant (left) and pellet (right) were run on SDS-PAGE. The relative band intensities are reported in the bar graph. Glk1 is presented here for comparison, taken from Figure 3-2-3A. **B**: Fluorescence images of cells from a saturated culture expressing NonPol-GFP (left), NonPolCatDead-GFP (middle), or CatDead-GFP (right) either before (top) or after (right) the introduction of glucose to the medium. NonPol and NonPolCatDead no longer form polymer in cells while CatDead does.

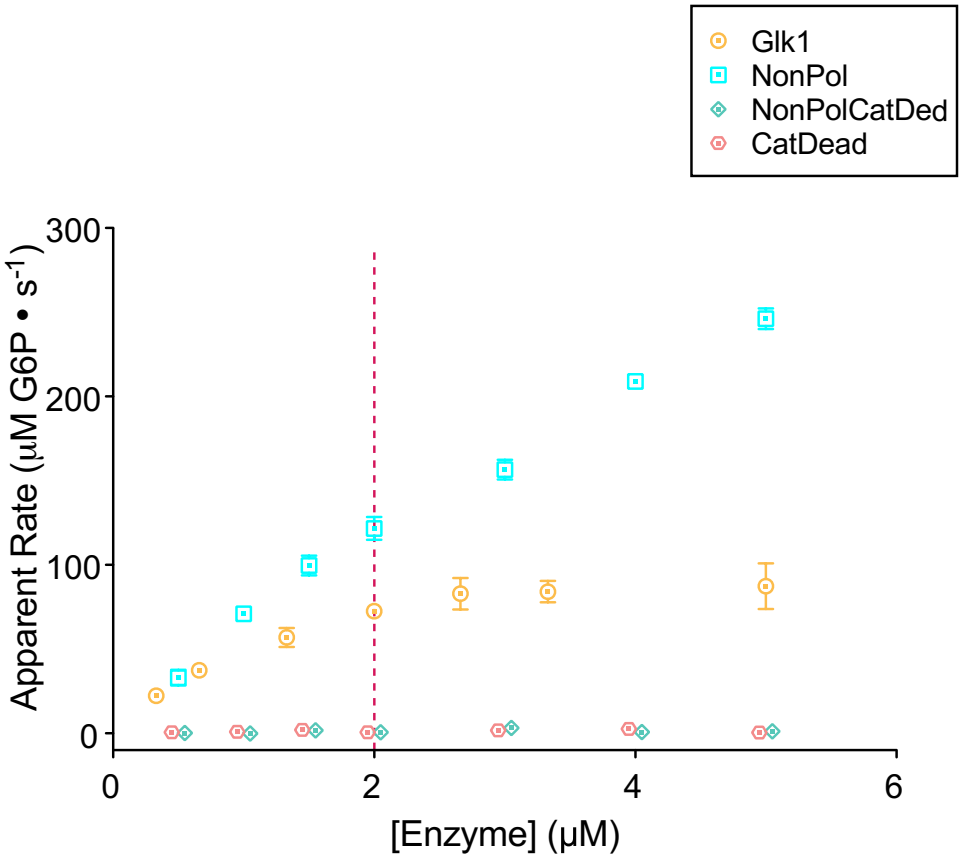


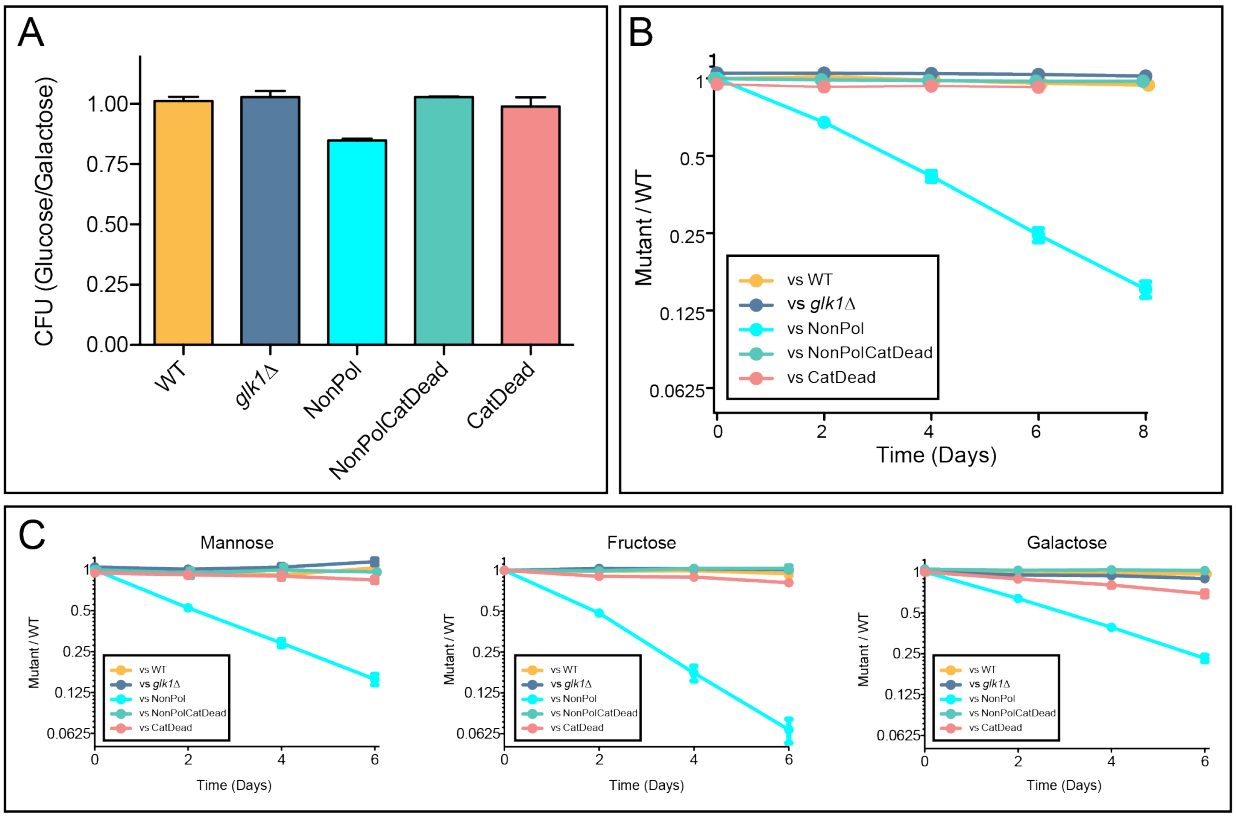
Figure 3-5-2. NonPol Enzyme Activity Lacks Negative Cooperativity
A: Rate of glucose-6-phosphate production was measured with respect to enzyme concentration for purified Glk1, NonPol, NonPolCatDead, and CatDead. Glk1 shows no significant increase in apparent rate beyond the concentration at which polymer is first observed (see Figure 3-2-4). NonPol lacks concentration-dependent inhibition while NonPolCatDead and CatDead show no activity.

Non-polymerizing Glk1 reduces the fitness of yeast growing on sugars.

Excess sugar kinase activity is toxic to yeast when starved cells are refed glucose because of an imbalance between the early steps of glycolysis, which consume ATP, and the late steps which generate it [10]. Yeast avoid this toxicity from excessive hexokinase activity of Hxk1 and Hxk2 because these enzymes are inhibited by trehalose-6-phosphate, an intermediate that transiently accumulates as a result of elevated levels of glucose-6-phosphate, the product of hexokinase and glucokinase activity [9,137]. Glk1 is not inhibited by trehalose-6-phosphate [138], so we hypothesized that Glk1 polymerization might prevent excessive Glk1 activity as glycolysis resumes in refed cells. Consistent with our hypothesis, feeding glucose to cells that were approaching stationary phase after growth on galactose kills 15% of the cells expressing NonPol relative to wild-type cells. Cells lacking Glk1 or expressing CatDead, or NonPolCatDead cells show no significant viability difference (Figure 3-5-3 A). These results suggest that Glk1 polymerization plays a role in limiting glucose phosphorylation during the start-up of glycolysis, but that the presence of the polymer itself is not needed for viability in these conditions.

We also tested the effect of unregulated Glk1 activity on fitness over the entire growth cycle. We competed wild-type cells with each of the mutants, diluted the cultures into fresh medium every 48 hours, and measured the proportion of each cell type. We found that NonPol cells have a severe fitness defect when growing on glucose but none of the other mutants or *glk1*Δ showed any major impairment (Figure 3-5-3 B). Similar effects were also observed in competitions on other sugars (Figure 3-5-

3 C). These experiments demonstrate that the growth defects in cells that express non-polymerizing Glk1 are caused by excess Glk1 activity, and not by a lack of Glk1 polymer.



3-5-3. Non-Polymerizing Glk1 Causes Growth and Viability Defects

A: The ratio of colonies formed on CBS-Glucose and CBS-Galactose after cells were grown in CBS-Galactose for 48 hours. Non-polymerizing Glk1 cells show increased cell death when introduced into glucose. **B,C**: Wild-type W303 labeled with mCherry were competed against green labeled W303 (wild-type, glk1Δ, NonPol, NonPolCatDead, CatDead) through repeated growth and dilution cycles in CSM-Glucose (B), CSM-Mannose, CSM-Fructose, and CSM-Galactose. The relative proportion of the strains were measured at each dilution by flow cytometry. Cells expressing NonPol show lower fitness on each carbon source, some of which are not Glk1 substrates. This suggests that there are other roles for Glk1 polymerization.

3.6. Materials and Methods

Strains and Culture Conditions

Unless otherwise noted, *S. cerevisiae* was grown at 30 ˚C in CSM supplemented with the appropriate carbon sources. For viability/plating assays, yeast were grown in CBS with appropriate carbon source [139].

Plasmid Construction and Cloning

The table in Appendix A-3 contains strains used. Constructs for the transformation of *S. cerevisiae* were generated by fusion PCR [140], and the cells were transformed with purified PCR product [141]. Clones were verified by amplification of the genomic locus followed by Sanger sequencing of the purified product.

For protein overexpression and purification from *E. coli,* glucokinases/hexokinases were cloned as his_6SUMO fusion constructs into T7 plasmids by Gibson assembly [142]. The plasmids were purified, and their sequences confirmed by sanger sequencing using universal primers.

Media Transitions

Cells grown to saturation were loaded into a microfluidic cell (CellAsic), and their spent medium flowed over them. Cells were exchanged into fresh media containing 2% glucose. Media was pumped at 2PSI.

Microscopy

Images were taken on a Nikon Ti inverted microscope with a Yokogowa spinning disc confocal unit, 447, 515, and 595 nm lasers, a Hammamatsu Orca camera operated with MetaMorph software. 8μm z-stacks were taken in 0.2μm increments. Z-stacks were converted to 8-bit, max intensity projected, and contrast adjusted in ImageJ [143].

Protein Purification

BL21(DE3)Rosetta containing the his_6SUMO fusion plasmid were grown to OD_{600} ~ 0.6 and induced overnight with 0.4mM IPTG at 16 ˚C. His_6SUMO fusion product was purified and cleaved as in [144]. Proteins were further purified and exchanged into HKM-buffer (20mM HEPES-KOH pH 7.5, 100mM KCl, 1mM $MgCl_2$, 10% Glycerol (v/v), 10mM β-mercaptoethanol) using an S200 16/600 column on an AKTA FPLC.

Ultracentrifugation

Unless otherwise noted, polymerization reactions were done in HKM-buffer and contained 5μM enzyme, 10mM Glucose, and 10mM MgATP. Polymerization reactions were spun in a TLA-100 at 100K rpm for 30 minutes at 30 ˚C. Supernatant was removed and added to equal volume of 2xSDS-Buffer. Pellet was resuspended by heating in 2 volumes of 1xSDS-Buffer. Fractions were run out on SDS-PAGE, stained with Coomassie Blue G250, and band intensities quantified in imageJ [143].

Light Scattering

Polymerization reactions were initiated by mixing 15μM Glk1 with an equal volume of 10mM Glucose and 10mM MgATP or 10mM Glucose-6-phosphate and 10mM MgADP using a SFA-20 rapid mixer (Hi-Tech Scientific), driven by a pneumatic drive unit (Hi-Tech Scientific). The 90 degree scattering of the solution at 315nm was measured using a Fluorolog-3 (Horiba).

Crystallization, Diffraction, and Refinement

Glk1 was exchanged into 5mM HEPES-KOH pH 7.5, 20mM KCl, 1mM $MgCl_2$, 0.5mM TCEP and concentrated to 11 mg/mL by centrifugal ultrafiltration (Amicon Ultra; Millipore). Conditions were screened around the conditions used in [145] as hanging

drops. Final crystallization conditions were 2.4M $(NH_4)_2HPO_4$, 0.1M CHES, pH 9.4 (2:1 protein:well solution) with crystals forming after 5 days at 20˚C. Crystals were cryoprotected by briefly soaking in well solution supplemented with 18% glycerol before freezing.

X-ray data were collected at the Advance Photon Source beamline 24-ID-C at the Argonne National Laboratory. 180˚ of data was collected with a 10μm beam at 20% transmission. Reflection data were indexed, integrated, and scaled with HKL2000. Initial phases were obtained using Phaser for molecular replacement, using the *S. cerevisiae* hexokinase-2 structure (PBD ID: 1IG8, Kuser et al. 2000) as a search model. The model was built in Coot [146] and refined using Phenix.refine [147]. Positional and B-factor refinement with TLS, torsion angle, and NCS restraints were used. Phosphates from the crystallization condition were modeled into densities of appropriate size that were coordinated by basic and polar residues. Residues 51-59 were omitted from all chains because this loop had poor density. Besides this omission, chain A contains residues 1-500, chain B contains residues 3-500, chain C contains residues 1-500, chain D contains residues 4-500, chain E contains residues 3-500, and chain F contains residues 1-500. Chain A was used for analysis and generation of figures.

Bioinformatics

Glucokinase and hexokinase homologs were found by reciprocal BLAST [148] searches and were informed by synteny when possible. Protein alignments were made using Clustal Omega [149] and trees were generated using a neighbor joining approach in Jalview [150].

Activity Assays

Glucose-6-phosphate production was measured in HKM-buffer at 30 ˚C. Timepoints were quenched in an equal volume of 100mM EDTA at 95 ˚C. Glucose-6-phosphate concentration was measured in each timepoint by a colorimetric enzyme-coupled reaction as in [151].

Competition Experiments

Strains were grown overnight in CSM with the appropriate carbon source, back diluted and grown to mid exponential phase. mCherry labeled wild-type cells were mixed with GFP labeled strains in equal numbers. Cells were back diluted 200-fold every 48 hours, and samples taken. The ratio of the two strains was measured by flow cytometry (Fortessa) and the data analyzed in FlowJo.

Viability Experiments

Strains were conditioned in CBS/Galactose for 48 hours being backdiluted to $1x10^5$ cells/mL every 12 hours. Cells were harvested, washed twice in CBS with no carbon, and plated onto either CBS/Glucose or CBS/Galactose. Colonies were counted after three days.

Chapter 4

Conclusions and Future Directions

4.1 Summary of Results

4.2 Yeast Glucokinases and Hexokinases

4.3 Fluorescent Proteins Effect on Polymerization

4.4 Enzyme Filaments
- ***Role in Activity Regulation***
- ***Protein-Protein Interactions***

4.5 The Evolution of Metabolism in Yeast

4.6 Actin-family Protein Polymers

4.7 Glucokinase Becoming a Polymer

4.8 Fitness effects of Glk1 Polymerization
- ***Polymerization During Approach to Diauxy***
- ***Polymerization During the Cell Cycle***

4.9 Redundancy of Enzymes

Conclusions and Future Directions

4.1. Summary of Results

In this thesis I investigated the polymerization of the yeast glucokinase, Glk1. This investigation shed light on enzyme regulation via oligomerization, the evolution of polymerization amongst actin-related proteins, and the regulation of central carbon metabolism in the face of a fluctuating environment. These findings, in detail, were:

- Glk1 forms polymers in cells in response to refeeding of sugars, while Hxk1 and Hxk2 do not. (Chapter 3.2)
- Glk1 polymerizes in response to its substrates or products when Glk1 concentration is sufficiently high. (Chapter 3.2)
- Glk1 polymerization inhibits Glk1 activity. This sets an upper limit to the rate of glucose-6-phosphate synthesis for the entire pool of Glk1. (Chapter 3.2)
- The Glk1 monomer structure is highly similar to the structure of glucokinases/hexokinases. The N and C termini are different than gluokinases outside the Saccharomycetaceae. (Chapter 3.3)
- Glk1 filaments are two-stranded, anti-parallel helices. The contacts within the filaments are unique amongst polymers formed by actin-related proteins. (Chapter 3.3)
- Glucokinase polymerization arose within yeast 150-200 million years ago and evolved independently of other actin-related polymers. (Chapter 3.4).
- A single point mutation (F3S) eliminates Glk1 polymerization. Glk1(F3S) activity lacks Glk1 activity's concentration dependence. (Chapter 3.5)

- Glk1 activity is not required for polymerization. A catalytic null, Glk1(K182A), forms polymer in response to ligand. (Chapter 3.5)
- Glk1 polymerization limits cell death upon refeeding of glucose; more cells die during refeeding when bearing Glk1(F3S). (Chapter 3.5)
- Yeast expressing Glk1(F3S) have additional growth defects when growing on sugars as their carbon source. (Chapter 3.5)

Additionally, many of these findings have also opened up new lines of questioning that can be pursued. Our understanding of how oligomerization of enzymes effects the metabolic networks in which they exist is still in its infancy. The questions opened up by this study include the following. In the following sections I discuss how I would begin solving several of these puzzles.

- What are all of the sequence and structural features of Glk1 and related glucokinases that allow them to form polymer?
- How does function diversify within a particular protein fold?
- How does enzyme polymerization effect metabolic networks? Does it make them more resilient to rapid fluctuations in environment? Does it buffer systems against imbalance?
- Does enzyme polymerization act as a signal regarding cell state?
- Are there Glk1 filament binding proteins? Do Glk1 filaments act as a scaffold for certain other protein-protein interactions?
- Is the role of glucokinase polymerization the same in other yeast species?

- What evolutionary pressures led to the evolution of glucokinase polymerization, and why has it been maintained for so long? Do these pressures still exist, or has the function changed over time?
- What causes the fitness defects when growing on other sugars besides glucose?
- Does Glk1 polymerization serve a role in responding to fluctuations derived from cellular activity and not from the environment?

4.2. Yeast Glucokinases and the Hexokinases

Across all ascomycete fungi, excepting the yeast descended from the whole genome duplication, each species contains one hexokinase and one glucokinase. The divergence between these two proteins is ancient and predates the ascomycetes (500 million years ago) [152].

After the whole genome duplication on the branch leading to *S. cerevisiae* occurred, most duplicated genes were eventually lost. Hexokinase and glucokinase were two of the most common genes to avoid the winnowing process and to remain duplicated in most species [153]. Duplication of genes temporarily relaxes selection against change in function, and thus new functions can arise [154,155]. In *S. cerevisiae*, one of the two glucokinase paralogues has lost its glucokinase activity and now acts as a transcription factor involved in meiotic initiation (Emi2) [156]. I have found that despite being descended from a polymerizing glucokinase, Emi2 does not polymerize in response to glucose or other sugars despite having the appropriate N and C terminal sequences. I have also found that Emi2 has a point mutation in its ATP binding site relative to Glk1 (S455Y). The mutation from serine to a bulky tyrosine likely eliminated

ATP binding, and thus, eliminated glucokinase activity. Furthermore, several of the most closely related yeast species to *S. cerevisiae* also have an Emi2-like protein (by synteny and BLAST) that also has Y455.

It would be interesting to see if mutation of Y455 to serine restores glucokinase activity in Emi2. Furthermore, I would like to test if this same point mutant also regained polymerization. It would also be interesting to know the meiotic phenotype of emi2Δ/emi2Δ cells. No direct studies have been performed on the protein, though its deletion results in early meiotic initiation in screens. A possible model to explain these observations is that Emi2 has lost its ability bind ATP, which eliminated its glucokinase activity. The conformational change that Emi2 undergoes upon glucose binding is now used as a glucose sensor. In this model, Emi2 represses meiosis when glucose/glucose-6-phosphate bound, and this repression is alleviated when glucose is absent. Our vast knowledge of point mutants that change binding affinity and specificity amongst glucokinases and hexokinases could be leveraged to investigate this model. Performing these studies would contribute to our understanding of yeast meiosis in the specific, but would also contribute to our understanding of how novel functions evolve in proteins.

4.3. Fluorescent Proteins Effect on Polymerization

In Chapter 3.2 I showed that Glk1-GFP forms polymers in cells upon refeeding with various sugars. The size of these polymers is not consistent with the size of polymers observed *in vitro*. It is possible that the polymers are longer and/ or form large bundles in cells. This difference could arise from crowding-mediated depletion

interactions driving bundling of filaments, stabilization of filaments by filament binding proteins [130], bundling driven by bivalent filament binding proteins [157], or some combination thereof. It is also possible that the fluorescent protein tag is enhancing Glk1 polymerization and driving bundling.

Fluorescent protein tags have self-affinity. This affinity can enhance oligomerization processes and can even drive the formation of unnatural aggregates or phase separated droplets by avidity [131]. It is likely that bundling is enhanced by the presence of the fluorescent-protein tag in Glk1-GFP strains. Overexpression of Glk1 tagged with fluorescent proteins for purification was highly toxic in *E. coli*, while all untagged versions of Glk1 express very strongly and are well tolerated. Thus, I was never able to test the effects of these protein tags *in vitro*.

To address these questions two lines of questioning should be followed. 1) What is the size distribution of Glk1 filaments in cells without traditional, beta-barrel fluorescent-protein tags? 2) What are the effects of fluorescent-protein tags on Glk1 polymerization? To address the first question, several other labeling techniques can be employed. The globular Halo-tag system [158] could be used to label Glk1 fluorescently without the risks associated with beta-barrel fluorescent-protein tags. This would eliminate concerns about the tag enhancing polymerization or bundling by self- affinity. The second question could be addressed by solving the problems with purification of Glk1-GFP. To obtain sufficient quantities of Glk1-GFP for *in vitro* assays it can be purified from *S. cerevisiae* either overexpressing the enzyme or grown in conditions to maximize Glk1 expression. Having an understanding of what the real Glk1 polymer

length distribution and how much Glk1 is in polymer in cells will be critical for future experiments modeling yeast metabolism.

4.4. Enzyme Filaments

Role in Activity Regulation

In Chapter 3.2 I demonstrate that Glk1 polymerization inhibits Glk1 enzyme activity. A handful of other enzyme filaments are known, however, biological function has not been established for most [124]. CTP synthetase is the best characterized and most widespread enzyme filament [119]. Interestingly, in some organisms the filament form is inhibited while in other species it is the active form [50,128]. Several other enzyme filaments have been observed, but polymerization of most of these filaments is mediated by pH [125]. They assemble when cells are starved and cytosolic pH drops. These types of filaments are not dynamic, and they are thought to act as storage depots for temporarily unnecessary enzyme so that cells are ready to begin growth rapidly upon refeeding.

Enzyme filaments that are regulated by the presence or absence of ligand, like Glk1 and CTP synthetase, have been proposed to be key regulators of flux through metabolic networks. They are thought to act as buffers against sudden changes and to maintain balance between different metabolites.

It would be valuable to study the effects of Glk1 polymerization on the yeast metabolome. Understanding how concentrations of metabolites change in response to glucose pulse in Glk1 vs Glk1(F3S) cells would shed light onto the molecular role of Glk1 polymerization. Measuring metabolite concentrations during steady state growth

on sugars in these different strains would also be important. This may help us understand the cause of Glk1(F3S) cells decreased fitness growing on sugars. This research would also help us understand the role of dynamic enzyme filaments more generally and would contribute to our understanding of when this type of adaptation evolves.

Protein-Protein Interactions

Polymerization can be an effective means of sensing the environment and cell state. While we detect no fitness defects due to the absence of Glk1 polymers specifically, it is impossible to test all conditions exhaustively. Glk1 is known to interact with other proteins, including several glycolytic proteins. Notably, Glk1 may interact with certain proteins in the trehalose pathway [159]. It seems likely that the polymeric state could either enhance or inhibit specific interactions. The scaffolding of protein-protein interactions is observed in the context of other polymers [102,105] and Glk1 polymers could be driving the formation of signaling complexes or metabolons. Alternatively, the monomeric form may bind certain proteins, and polymerization may release these proteins.

Several affinity techniques can be employed to search for proteins that interact with Glk1 in either the monomer or polymer state. Finding filament binding proteins could be accomplished with ultracentrifugation of cell lysate after the addition of ligand. Monomer binding proteins would be evaluated through co-immuno precipitation using an affinity tag. Understanding what proteins bind specifically to these different states of Glk1 may lead to insights regarding other conditions where Glk1 polymerization is important.

4.5. The Evolution of Metabolism in Yeast

In Chapter 3.4 I showed that the polymerization of Glk1 arose in the last common ancestor of all of the saccharomycetaceae, that is, in a yeast that existed 150-200 million years ago. This group of yeast also lacks Respiratory Complex 1 from their mitochondria, suggesting that these two changes happened at a similar point in yeasts' evolutionary history [135]. The lack of Respiratory Complex 1 reduces the respiratory capacity within these yeasts and has been suggested to be an early step towards development of the Crabtree effect [160]. This change predates the yeast whole genome duplication, which gave rise to *S. cerevisiae* and its closest relatives. Thus, when glucokinase polymerization arose, it was in a context where cells only contained one hexokinase and one glucokinase. This is still the case for a majority of yeast species.

What might have led to the evolution of polymerization-limited glucokinase activity? One possible explanation for this adaptation lies in an adaptation to the reduced respiratory capacity caused by the loss of Respiratory Complex 1. When a sudden pulse of a nutrient occurs, it can lead to overflow metabolism. In this case, overflow is caused by glycolysis outpacing respiration, resulting in short term fermentation in response to a glucose pulse. Interestingly, amongst the non-Crabtree members of the saccharomycetacea (all of whom did not undergo a genome duplication) only the kluyveromycetes show overflow metabolism in response to a glucose pulse [161]. The kluyveromycetes are the only lineage in which glucokinase polymerization has been lost. In the future, I would like to see if elimination of

glucokinase polymerization results in overflow metabolism in other non-Crabtree members of the sacharomycetacea.

The Saccharomycetaceae represent a group of yeast with diverse metabolism and lifestyle. These range from aerobic fermenters like *S. cerevisiae*, to commensal yeast like *C. glabrata*, and filamentous fungi like *E. gossypii*. Given the evolutionary distance of these organisms and their varied carbon metabolisms, it is plausible that glucokinase regulation by polymerization holds a different biological function in different yeast species. It would be interesting to eliminate glucokinase polymerization in each of the Saccharomycetaceae. This would contribute to our understanding of what the original function of glucokinase polymerization might have been.

4.6. Actin-family Protein Polymers

In Chapter 3.3 I show and discuss the structure of Glk1 and of Glk1 polymers. The polymer structure of all other known actin-fold protein polymers have some key similarities; notably, within each strand, the geometry of monomer-monomer interaction is conserved [29,30]. This geometry is violated in Glk1 polymers. This appears to be caused by the constraint that glucokinases and hexokinases have a dimerization interface [145]. The dimers involve contacts between subdomain IB and IIB. This constraint means that within individual strands of Glk1 filaments, the primary contact is between subdomain IIB and subdomain IA (see Appendix A-1)

The conservation of the geometry between monomers in individual strands of actin-related protein filaments has been explained in two different ways. One explanation is that these contacts are structural constraints on polymerization of an

actin fold protein. Under this model, actin-related filaments may have arisen independently from different actin-fold proteins. The other explanation is that the conservation of these intra-strand contacts is evidence that all extant actin-related polymers are descended from a single ancestral, polymer-forming protein. Once stable contacts had evolved along strands they deviated little, and filament diversity arose by changing inter-strand contacts.

The structure of Glk1 filaments demonstrates that the first model is not true. IA/IB and IIA/IIB contacts are not constraints for filament formation of an actin-fold protein. This supports the model that actin-related filaments are all descended from an ancestral polymerizing protein. Taken together with the fact that all polymer-forming actin-fold proteins form a monophyletic group (along with Hsp70) in a tree of the actin superfamily [29] makes it highly likely that polymerization has evolved twice within the actin superfamily; once in the last universal common ancestor and once 150-200 million years ago in the last common ancestor of the Saccharomycetaceae.

4.7. Glucokinase Becoming a Polymer

In Chapter 3.5 I demonstrated that Glk1 polymerization relies on the presence of solvent-exposed phenylalanine at the end of its N-terminal helix to form polymer. It has been demonstrated that the addition of hydrophobic residues to specific solvent-exposed regions in other proteins can lead to the ability to form polymers. This has been observed in natural systems (e.g. sickle-cell hemoglobin) and has also been engineered in symmetric protein complexes [162,163]. These engineered polymers demonstrate that properly spaced, solvent-exposed hydrophobic patches are a general

feature of protein oligomerization. Other features that I found to be important for Glk1 polymerization are an extended loop 438-444 in the Glk1 sequence and that the C-terminal helix ends within the protein core, creating a hydrophobic patch for the N-terminal phenylalanine to bind.

While I have good phylogenetic and structural evidence for the importance of these features in glucokinase polymerization, it is unclear how much of a contribution each of these individual factors makes. I would like to ask several questions about what causes glucokinase polymerization. How important is the sequence of loop 438-444? The addition of protein tags to the C-terminus still support polymerization as long as there is a disordered linker region. What types of C-terminal extensions can polymerization tolerate? Can adding loop 438-444, and the N-terminal, and C-terminal motifs to non-polymerizing glucokinases generate polymerizing proteins? Are there other differences between hexokinases and glucokinases that make these features sufficient in a glucokinase but not in a hexokinase? Emi2 does not polymerize and has a mutation in its ATP binding site. Does this mutation eliminate Glk1 polymerization? These questions should be answered by studying the chimeric proteins *in vitro*.

4.8. Fitness Effects of Glk1 Polymerization

In Chapter 3.5 I describe two fitness defects associated with a lack of Glk1 polymerization.

1) Cells expressing non-polymerizing Glk1 have reduced viability upon glucose refeeding.

2) Cells expressing non-polymerizing Glk1 show reduced competitive fitness when growing on sugars in mixed cultures.

The magnitude of these two effects is not equal, and therefore, Glk1 polymerization likely has effects beyond limiting glucokinase activity during glucose refeeding. I have not established what the mechanism of these effects is, but Glk1 polymerization can be observed during other states than glucose refeeding. I describe these conditions below and describe how I think questioning should progress.

Polymerization During Approach to Diauxy

As a yeast culture depletes the available supply of glucose, glucose repression is alleviated before glucose concentration reaches zero. When this occurs, Glk1 expression increases leading to a transient period where polymers are present (see Appendix A-2) before cells fully deplete glucose and fully switch to respiration. Another observation that I was never able to explain is that cells expressing Glk1(F3S) reach saturation at a lower density than wild-type cells. I speculate that this, combined with the observation that Glk1 polymers appear as cell approach diauxy, indicates a potential role of Glk1 polymerization in preparing cells appropriately for respiration. Without Glk1 polymerization, cells respond to diauxy inappropriately.

Polymerization During the Cell Cycle

I have observed that in a batch culture Glk1-GFP tends to form polymers more frequently in budded cells than it does in unbudded cells. This difference is most pronounced when cells are respiring ethanol after undergoing diauxy. Amongst unbudded cells there are no observable filaments, however, some budded cells have Glk1-GFP polymers despite the lack of glucose or other sugars in the environment.

One possibility for why these polymers appear is that storage carbohydrates are released in a cell-cycle dependent manner during S/M phase [164,165].

Cells expressing non-polymerizing Glk1 appear to die slowly after diauxy. Given that Glk1-GFP polymers are observed in some budded cells after diauxy implies that Glk1 polymerization may be important for fitness after this transition. Investigating this could lead to valuable insights about the interaction between cell cycle and carbon metabolism and would increase our understanding of the biological significance of Glk1 polymerization.

4.9. What can we learn from enzyme redundancy?

It is problematic that for many organisms, and more broadly, cell types, that we do not know precisely to what range and frequency of environments they are adapted. Without this basic understanding, studying cell physiology is of limited value. Being blind to the organism's natural environmental conditions, of which the organism is a product, often leads to measurements that are aphysiological. This can lead to incorrect conclusions about the essentiality of genes amongst other, more subtle problems regarding our understanding of cell systems. The key way to solving this problem is to build an understanding of the types of conditions to which the organism is adapted.

Throughout biology there are enzymes that appear to be redundant with each other; they perform the same chemistry in a single species. Hxk1, Hxk2, and Glk1 are one such set of 'redundant' enzymes. Frequently, studies of the biochemical properties of these enzymes reveal striking differences. These differences include different pH

optima, different temperature optima, different rates and substrate affinities, and different allosteric regulation. Investigation of the differences between redundant enzymes can lead to insights about the environments that the organism-in-question has seen on a regular basis during its evolutionary history. Understanding the differences between enzymes with similar functions, when these different enzymes are expressed, and with what other proteins they are expressed is the key to understanding the environments to which organisms are adapted. This information is crucial for our meaningful understating of biological systems.

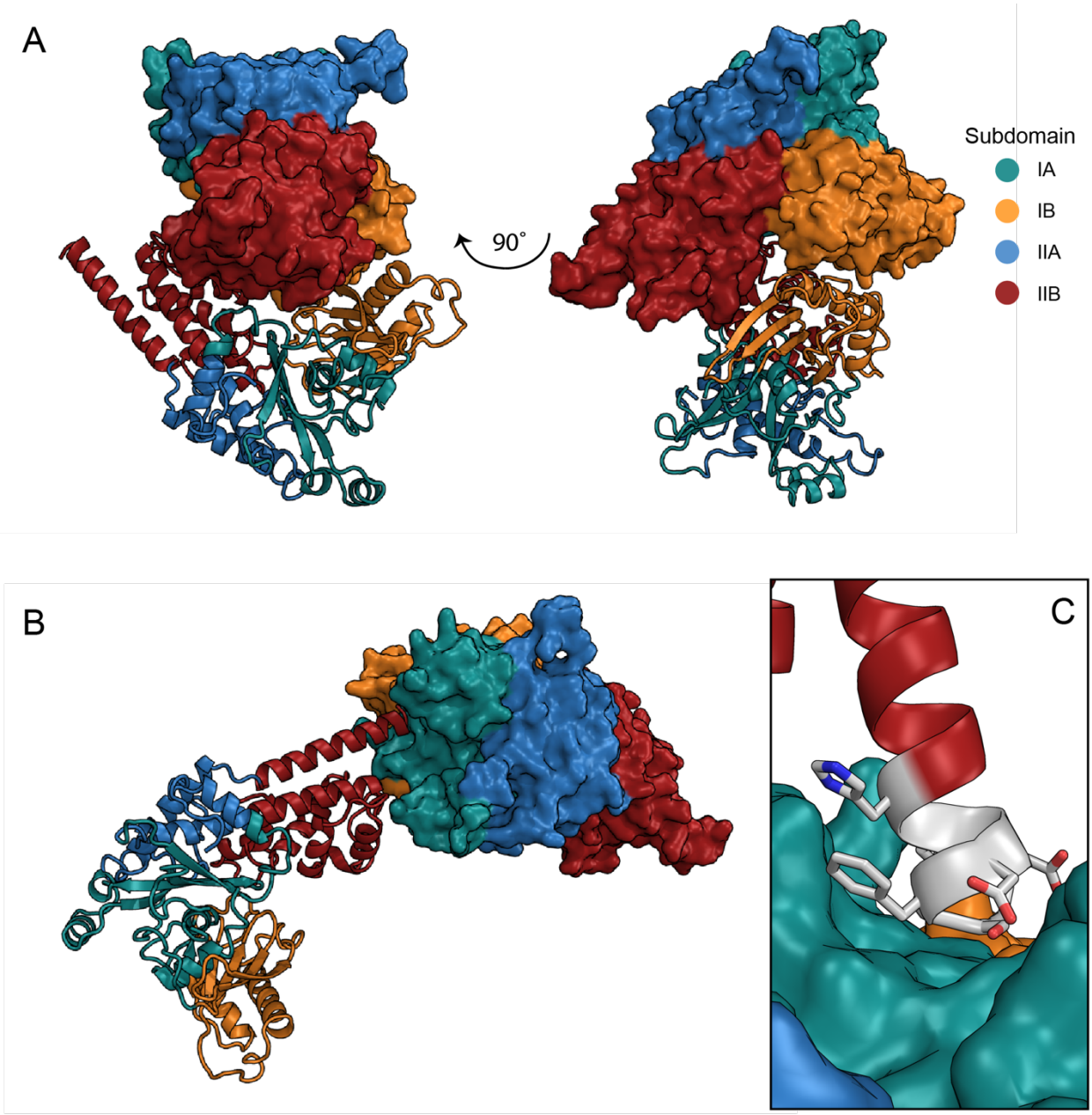
Appendix A

Additional Information

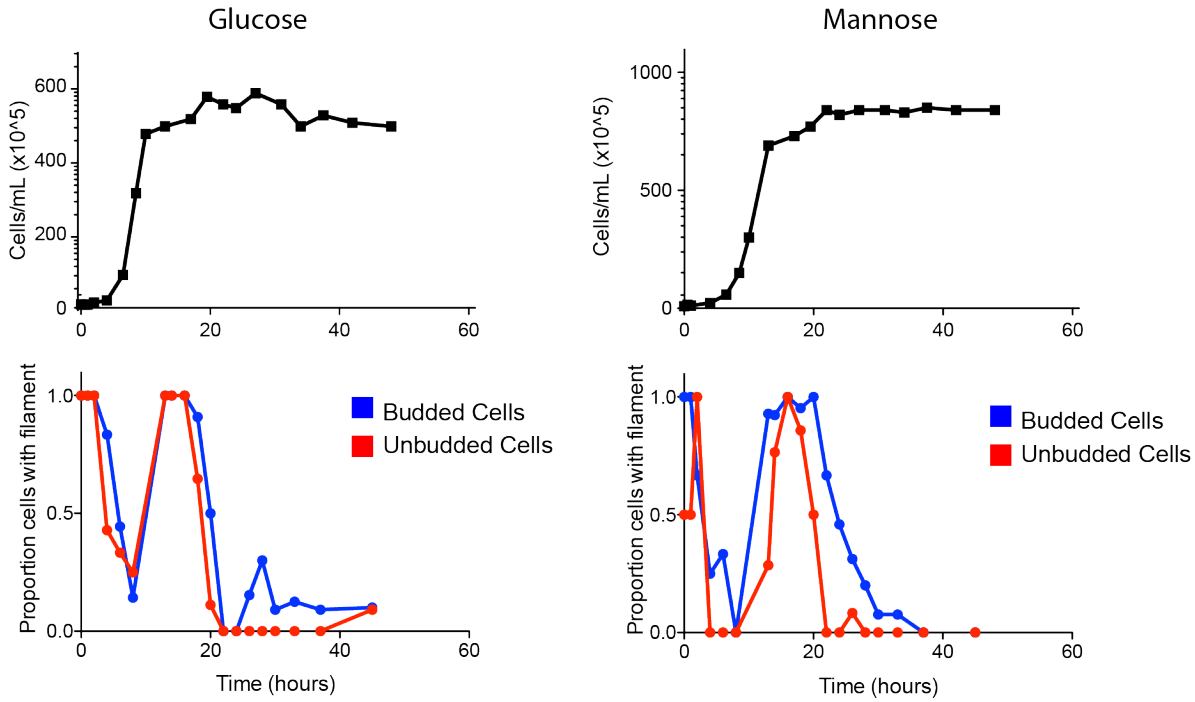
A-1 Glk1 Filament Contacts

A-2 Glk1 Filaments in Other Conditions

A-3 Table of Strains



Appendix A-1: Glk1 Filament Contacts
A: Domain IB and IIB are involved in interstrand contacts, and therefore are unavailable for the typical intrastrand contacts observed in other actin-fold protein polymers. This interface is similar to the dimerization interface seen in other glucokinases and hexokinases.
B: The Glk1 filament longitudinal contact is between subdomain IA and subdomain IIB.
C: Closeup of the N-terminal helix contacting subdomain IA.



Apendix A-2: Glk1 Filaments in Other Conditions

A: Glk1-GFP filaments are observed in the entire population as glucose is exhaausted and in some budded cells when growing on accumulated ethanol. This may imply a role for Glk1 polymerization during these growth phases.

Table A-3. Strains used.

Strain	Genotype
yPS003	W303 (*ura3 ade2-1 his3-11,15 leu2-3,112 trp1-1 Mat*α BUD4 GLK1-sfGFP::spHIS5*)*
yPS031	W303 (*ura3 ade2-1 mCherry::HIS leu2-3,112 trp1-1 Mat*α BUD4*)*
yPS033	W303 (*ura3 ade2-1 Citrine::HIS leu2-3,112 trp1-1 Mat*α BUD4*)*
yPS041	W303 (*ura3 ade2-1 Citrine::HIS leu2-3,112 trp1-1 Mat*α BUD4 glk1Δ::HPH*)*
yPS106	W303 (*ura3 ade2-1 his3-11,15 leu2-3,112 trp1-1 Mat*α BUD4 GLK1(F3S)*)*
yPS108	W303 (*ura3 ade2-1 his3-11,15 leu2-3,112 trp1-1 Mat*α BUD4 GLK1(F3S,K182A)*)*
yPS109	W303 (*ura3 ade2-1 his3-11,15 leu2-3,112 trp1-1 Mat*α BUD4 GLK1(K182A)*)*
yPS110	W303 (*ura3 ade2-1 his3-11,15 leu2-3,112 trp1-1 Mat*α BUD4 GLK1(F3S)-GFP::URA*)*
yPS111	W303 (*ura3 ade2-1 his3-11,15 leu2-3,112 trp1-1 Mat*α BUD4 GLK1(F3S,K182A)-GFP::URA*)*
yPS112	W303 (*ura3 ade2-1 his3-11,15 leu2-3,112 trp1-1 Mat*α BUD4 GLK1(K182A)-GFP::URA*)*
yPS113	W303 (*ura3 ade2-1 Citrine::HIS leu2-3,112 trp1-1 Mat*α BUD4 GLK1(F3S)*)*
yPS116	W303 (*ura3 ade2-1 Citrine::HIS leu2-3,112 trp1-1 Mat*α BUD4 GLK1(F3S,K182A)*)*
yPS117	W303 (*ura3 ade2-1 Citrine::HIS leu2-3,112 trp1-1 Mat*α BUD4 GLK1(K182A)*)*

Appendix B

An Archaeal Tubulin Homolog Forms Tubules

B-1 Introduction

B-2 Odin Tubulin Polymerizes

B-3 Odin Tubulin Forms Tubules

B1: Introduction

Lokiarchaeota are a recently discovered group of archaeal prokaryotes who were first found in metagenomic assemblies from deep sea sediment near geothermal vents.[28] More species have also been found in other environments, though no example has been isolated and brought into culture.[166] These species form a monophyletic group with eukaryotes in the tree of known eukaryotes, eubacteria, and archaea. The genomes of these organisms contain many putative eukaryotic-type proteins that have not been previously observed within prokaryotes, some of which are cytoskeletal proteins and known cytoskeletal effector proteins. One clade within the lokiarchaeota, dubbed 'Odin,' possesses a tubulin homolog that is more similar to eukaryotic tubulin than it is to tubulin homologs previously observed in eubacteria, viruses, and archaea.[97,167,168]

B2: Odin Tubulin Polymerizes

No species of lokiarchaea have been cultured, and thus, cell biological experiments are impossible at this point. To study the Odin tubulin homolog, I cloned the synthetic gene into an expression vector and purified the protein. I found that Odin tubulin oligomerizes in a GTP dependent fashion (Figure B-2-1), which I observed by ultracentrifugation. Polymerization kinetics can be observed by light scattering, and increased temperature causes faster and more complete polymerization (Figure B-2-2 A). Polymerization can also be observed down to 1.5μM Odin tubulin (Figure B-2-2 B).

B3: Odin Tubulin Forms Tubules

To determine the structure of Odin tubulin oligomers I took negative stain images of the polymers, and they form slinky-like tubules that look to grow quite differently than microtubules (Figure B-3-1). The tubules have an external diameter of approximately 100nm and an internal diameter of approximately 70nm. The tubules have thick walls that are approximately four times the width of a tubulin monomer (~16nm). Both side on and end on views of the tubules reveal this thickness. Globular objects with the same width as the wall thickness were also visible in the field. The strand thickness was approximately 4nm, however, which is the expected width of a tubulin monomer. It appears that a disc shaped multimer may be the unit of addition and the filaments grow in a spiral 'slinky-like' fashion.

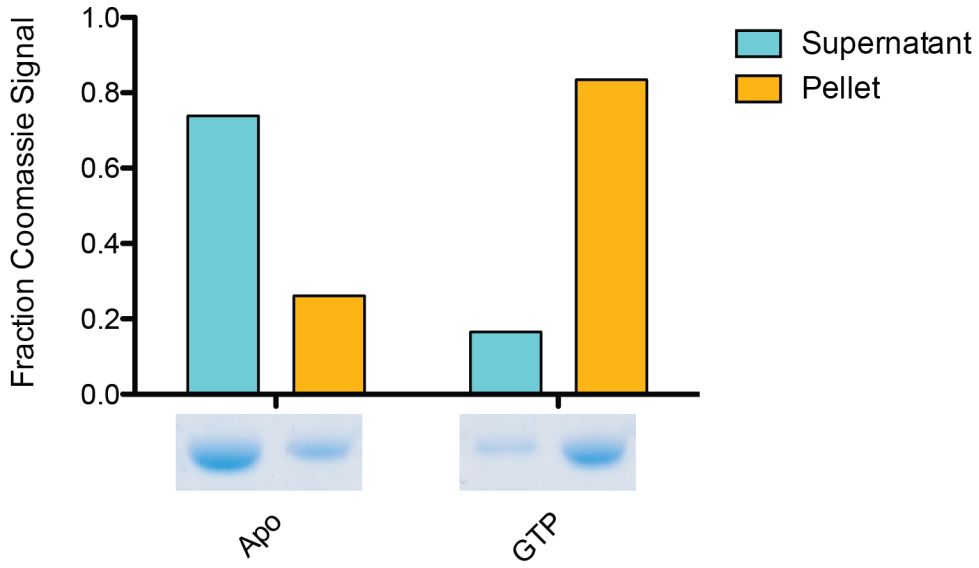


Figure B-2-1. Odin Tubulin Oligomerizes With GTP
Odin Tubulin was ultracentrifuged in the presence or absence of GTP. GTP is required for Odin tubulin oligomerization.

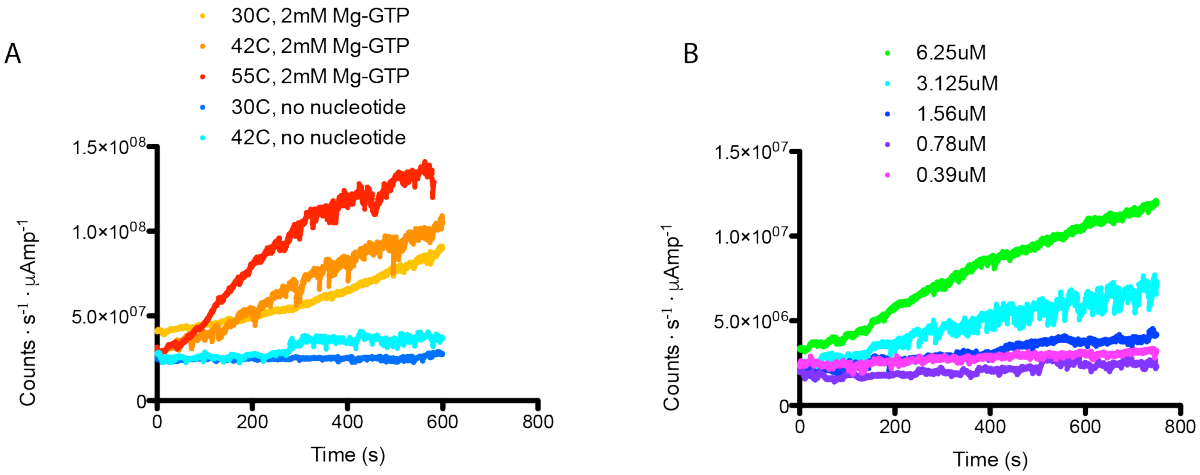


Figure B-2-2. Odin Tubulin Polymerization Can be Observed by Light Scattering
Odin Tubulin polymerization is stimulated by increasing temperature (A). Polymerization can be observed down to ~1.5uM at 42C (B).

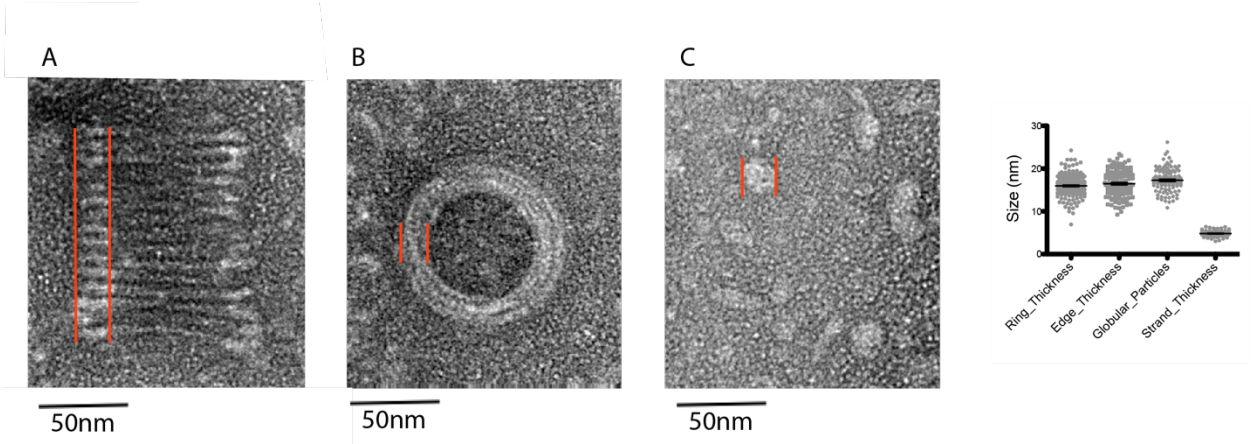


Figure B-3-1. Odin Tubulin forms slinky-like tubules.
The tubules are wider and shorter than microtubules. The walls of the tubules are thicker than the expected width of a single monomer (A and B), while the strand thickness is the expected width of a monomer. This suggests that a ring or disc multimer may be theunit of addition, perhaps as the object in C.

Works Cited

1. Hollien, J. Evolution of the unfolded protein response. *Biochim. Biophys. Acta* **1833,** 2458–2463 (2013).
2. Mager, W. H. & Siderius, M. Novel insights into the osmotic stress response of yeast. *FEMS Yeast Res.* **2,** 251–257 (2002).
3. Calcott, P. H. & Postgate, J. R. On substrate-accelerated death in Klebsiella aerogenes. *J. Gen. Microbiol.* **70,** 115–122 (1972).
4. Postma, E., Verduyn, C., Kuiper, A., Scheffers, W. A. & van Dijken, J. P. Substrate-accelerated death of Saccharomyces cerevisiae CBS 8066 under maltose stress. *Yeast* **6,** 149–158 (1990).
5. Calcott, P. H., Montague, W. & Postgate, J. R. The levels of cyclic AMP during substrate-accelerated death. *J. Gen. Microbiol.* **73,** 197–200 (1972).
6. Fraenkel, D. G. *Yeast Intermediary Metabolism*. (2011).
7. Rodríguez, A., La Cera, De, T., Herrero, P. & Moreno, F. The hexokinase 2 protein regulates the expression of the GLK1, HXK1 and HXK2 genes of Saccharomyces cerevisiae. *Biochem. J.* **355,** 625–631 (2001).
8. Teusink, B., Walsh, M. C., van Dam, K. & Westerhoff, H. V. The danger of metabolic pathways with turbo design. *Trends Biochem. Sci.* **23,** 162–169 (1998).
9. Bonini, B. M., Van Dijck, P. & Thevelein, J. M. Uncoupling of the glucose growth defect and the deregulation of glycolysis in Saccharomyces cerevisiae Tps1 mutants expressing trehalose-6-phosphate-insensitive hexokinase from Schizosaccharomyces pombe. *Biochim. Biophys. Acta* **1606,** 83–93 (2003).
10. van Heerden, J. H. *et al.* Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. *Science* **343,** 1245114–1245114 (2014).
11. Hurley, J. H. The sugar kinase/heat shock protein 70/actin superfamily: implications of conserved structure for mechanism. *Annu Rev Biophys Biomol Struct* **25,** 137–162 (1996).
12. Ettema, T. J. G., Lindås, A.-C. & Bernander, R. An actin-based cytoskeleton in archaea. *Mol. Microbiol.* **80,** 1052–1061 (2011).
13. Murakami, K. *et al.* Structural Basis for Actin Assembly, Activation of ATP Hydrolysis, and Delayed Phosphate Release. *Cell* **143,** 275–287 (2010).
14. Toro-Nahuelpan, M. *et al.* Segregation of prokaryotic magnetosomes organelles is driven by treadmilling of a dynamic actin-like MamK filament. *BMC Biol.* **14,** 88 (2016).
15. Møller-Jensen, J. *et al.* Bacterial mitosis: ParM of plasmid R1 moves plasmid DNA by an actin-like insertional polymerization mechanism. *Mol. Cell* **12,** 1477–1487 (2003).
16. Komeili, A., Li, Z., Newman, D. K. & Jensen, G. J. Magnetosomes are cell membrane invaginations organized by the actin-like protein MamK. *Science* **311,** 242–245 (2006).

17. Schmitz, H.-P., Kaufmann, A., Köhli, M., Laissue, P. P. & Philippsen, P. From function to shape: a novel role of a formin in morphogenesis of the fungus Ashbya gossypii. *Mol. Biol. Cell* **17,** 130–145 (2006).
18. Yin, H., Pruyne, D., Huffaker, T. C. & Bretscher, A. Myosin V orientates the mitotic spindle in yeast. *Nature* **406,** 1013–1015 (2000).
19. Rousso, T., Shewan, A. M., Mostov, K. E., Schejter, E. D. & Shilo, B.-Z. Apical targeting of the formin Diaphanous in Drosophila tubular epithelia. *Elife* **2,** e00666 (2013).
20. Westermann, B. Mitochondrial inheritance in yeast. *Biochim. Biophys. Acta* **1837,** 1039–1046 (2014).
21. Jégou, A., Carlier, M.-F. & Romet-Lemonne, G. Formin mDia1 senses and generates mechanical forces on actin filaments. *Nat Commun* **4,** 1883 (2013).
22. Garner, E. C., Campbell, C. S., Weibel, D. B. & Mullins, R. D. Reconstitution of DNA segregation driven by assembly of a prokaryotic actin homolog. *Science* **315,** 1270–1274 (2007).
23. Chang, F. & Huang, K. C. How and why cells grow as rods. *BMC Biol.* **12,** 2102 (2014).
24. Wong, K. C. Y. *et al.* Importance of a myosin II-containing progenitor for actomyosin ring assembly in fission yeast. *Curr. Biol.* **12,** 724–729 (2002).
25. Miller, A. L. The contractile ring. *Curr. Biol.* **21,** R976–8 (2011).
26. Behrmann, E. *et al.* Structure of the rigor actin-tropomyosin-myosin complex. *Cell* **150,** 327–338 (2012).
27. Gittes, F., Mickey, B., Nettleton, J. & Howard, J. Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. *J. Cell Biol.* **120,** 923–934 (1993).
28. Spang, A. *et al.* Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521,** 173–179 (2015).
29. Stoddard, P. R., Williams, T. A., Garner, E. & Baum, B. Evolution of polymer formation within the actin superfamily. *Mol. Biol. Cell* **28,** 2461–2469 (2017).
30. van den Ent, F., Izoré, T., Bharat, T. A., Johnson, C. M. & Löwe, J. Bacterial actin MreB forms antiparallel double filaments. *Elife* **3,** e02634 (2014).
31. Bharat, T. A. M., Murshudov, G. N., Sachse, C. & Löwe, J. Structures of actin-like ParM filaments show architecture of plasmid-segregating spindles. *Nature* **523,** 106–110 (2015).
32. Löwe, J., He, S., Scheres, S. H. W. & Savva, C. G. X-ray and cryo-EM structures of monomeric and filamentous actin-like protein MamK reveal changes associated with polymerization. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 13396–13401 (2016).
33. Söding, J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21,** 951–960 (2005).
34. Halliburton, W. D. On Muscle-Plasma. *J. Physiol. (Lond.)* **8,** 133–202 (1887).
35. Straub, F. B. Actin. *stud. inst. med. chem. univ. szeged* 1–15 (1942).
36. HANSON, J. & HUXLEY, H. E. Structural basis of the cross-striations in muscle. *Nature* **172,** 530–532 (1953).
37. Pollard, T. D., Shelton, E., Weihing, R. R. & Korn, E. D. Ultrastructural characterization of F-actin isolated from Acanthamoeba castellanii and

identification of cytoplasmic filaments as F-actin by reaction with rabbit heavy meromyosin. *Journal of Molecular Biology* **50,** 91–97 (1970).
38. Møller-Jensen, J., Jensen, R. B., Löwe, J. & Gerdes, K. Prokaryotic DNA segregation by an actin-like filament. *EMBO J.* **21,** 3119–3127 (2002).
39. Bork, P., Sander, C. & Valencia, A. An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc. Natl. Acad. Sci. U.S.A.* **89,** 7290–7294 (1992).
40. Mishra, M., Huang, J. & Balasubramanian, M. K. The yeast actin cytoskeleton. *FEMS Microbiol. Rev.* **38,** 213–227 (2014).
41. McDowell, J. M., Huang, S., McKinney, E. C., An, Y. Q. & Meagher, R. B. Structure and evolution of the actin gene family in Arabidopsis thaliana. *Genetics* **142,** 587–602 (1996).
42. Zhang, J. & Yang, J.-R. Determinants of the rate of protein sequence evolution. *Nat. Rev. Genet.* **16,** 409–420 (2015).
43. Kabsch, W., Mannherz, H. G. & Suck, D. Three-dimensional structure of the complex of actin and DNase I at 4.5 A resolution. *EMBO J.* **4,** 2113–2118 (1985).
44. Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. & Holmes, K. C. Atomic structure of the actin:DNase I complex. *Nature* **347,** 37–44 (1990).
45. Steitz, T. A., Anderson, C. M. & Stenkamp, R. E. SEQUENCING A PROTEIN BY X-RAY CRYSTALLOGRAPHY. II. REFINEMENT OF YEAST HEXOKINASE B CO-ORDINATES AND SEQUENCE AT 2.1 ANGSTROMS RESOLUTION. (1978). doi:10.2210/pdb2yhx/pdb
46. Korn, E., Carlier, M. & Pantaloni, D. Actin polymerization and ATP hydrolysis. *Science* **238,** 638–644 (1987).
47. Orlova, A. & Egelman, E. H. Structural basis for the destabilization of F-actin by phosphate release following ATP hydrolysis. *Journal of Molecular Biology* **227,** 1043–1053 (1992).
48. Ozyamak, E., Kollman, J. M. & Komeili, A. Bacterial Actins and Their Diversity. *Biochemistry* **52,** 6928–6939 (2013).
49. Szwedziak, P., Wang, Q., Freund, S. M. V. & Löwe, J. FtsA forms actin-like protofilaments. *EMBO J.* **31,** 2249–2260 (2012).
50. Barry, R. M. *et al.* Large-scale filament formation inhibits the activity of CTP synthetase. *Elife* **3,** e03638 (2014).
51. Noree, C., Sato, B. K., Broyer, R. M. & Wilhelm, J. E. Identification of novel filament-forming proteins in Saccharomyces cerevisiae and Drosophila melanogaster. *J. Cell Biol.* **190,** 541–551 (2010).
52. Lynch, E. & Kollman, J. The Structural Basis of Enzyme Regulation by CTP Synthase Metabolic Filaments. *Biophysical Journal* **110,** 26a–33 (2016).
53. Shen, Q.-J. *et al.* Filamentation of Metabolic Enzymes in Saccharomyces cerevisiae. *J Genet Genomics* **43,** 393–404 (2016).
54. Schuh, M. An actin-dependent mechanism for long-range vesicle transport. *Nat. Cell Biol.* **13,** 1431–1436 (2011).
55. Salje, J., van den Ent, F., de Boer, P. & Löwe, J. Direct membrane binding by bacterial actin MreB. *Mol. Cell* **43,** 478–487 (2011).

56. Bergeron, J. R. C. *et al.* Structure of the magnetosome-associated actin-like MamK filament at subnanometer resolution. *Protein Sci.* **26,** 93–102 (2017).
57. Jiang, S. *et al.* Novel actin filaments from Bacillus thuringiensis form nanotubules for plasmid DNA segregation. *Proc. Natl. Acad. Sci. U.S.A.* **113,** E1200–5 (2016).
58. Takatsuki, H., Bengtsson, E. & Månsson, A. Persistence length of fascin-cross-linked actin filament bundles in solution and the in vitro motility assay. *Biochim. Biophys. Acta* **1840,** 1933–1942 (2014).
59. Jansen, S. *et al.* Mechanism of actin filament bundling by fascin. *J. Biol. Chem.* **286,** 30087–30096 (2011).
60. Pollard, T. D., Blanchoin, L. & Mullins, R. D. Actin dynamics. *J. Cell. Sci.* **114,** 3–4 (2001).
61. Mullins, R. D., Heuser, J. A. & Pollard, T. D. The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proc. Natl. Acad. Sci. U.S.A.* **95,** 6181–6186 (1998).
62. Jacinto, A. & Baum, B. Actin in development. *Mech. Dev.* **120,** 1337–1349 (2003).
63. Lal, A. A., Korn, E. D. & Brenner, S. L. Rate constants for actin polymerization in ATP determined using cross-linked actin trimers as nuclei. *J. Biol. Chem.* **259,** 8794–8800 (1984).
64. Garner, E. C., Campbell, C. S. & Mullins, R. D. Dynamic instability in a DNA-segregating prokaryotic actin homolog. *Science* **306,** 1021–1025 (2004).
65. Polka, J. K., Kollman, J. M., Agard, D. A. & Mullins, R. D. The structure and assembly dynamics of plasmid actin AlfA imply a novel mechanism of DNA segregation. *J. Bacteriol.* **191,** 6219–6230 (2009).
66. Skillman, K. M. *et al.* The unusual dynamics of parasite actin result from isodesmic polymerization. *Nat Commun* **4,** 2285 (2013).
67. Vahokoski, J. *et al.* Structural differences explain diverse functions of Plasmodium actins. *PLoS Pathog.* **10,** e1004091 (2014).
68. Carlier, M. F., Pantaloni, D. & Korn, E. D. Evidence for an ATP cap at the ends of actin filaments and its regulation of the F-actin steady state. *J. Biol. Chem.* **259,** 9983–9986 (1984).
69. Pollard, T. D. & Borisy, G. G. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* **112,** 453–465 (2003).
70. Kirschner, M. W. Implications of treadmilling for the stability and polarity of actin and tubulin polymers in vivo. *J. Cell Biol.* **86,** 330–334 (1980).
71. Derman, A. I. *et al.* Phylogenetic analysis identifies many uncharacterized actin-like proteins (Alps) in bacteria: regulated polymerization, dynamic instability and treadmilling in Alp7A. *Mol. Microbiol.* **73,** 534–552 (2009).
72. Mitchison, T. & Kirschner, M. Dynamic instability of microtubule growth. *Nature* **312,** 237–242 (1984).
73. Husson, C., Renault, L., Didry, D., Pantaloni, D. & Carlier, M.-F. Cordon-Bleu uses WH2 domains as multifunctional dynamizers of actin filament assembly. *Mol. Cell* **43,** 464–477 (2011).

74. Drew, K. R. P. & Pogliano, J. Dynamic instability-driven centering/segregating mechanism in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 11075–11080 (2011).
75. Polka, J. K., Kollman, J. M. & Mullins, R. D. Accessory factors promote AlfA-dependent plasmid segregation by regulating filament nucleation, disassembly, and bundling. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 2176–2181 (2014).
76. Goley, E. D., Rodenbusch, S. E., Martin, A. C. & Welch, M. D. Critical conformational changes in the Arp2/3 complex are induced by nucleotide and nucleation promoting factor. *Mol. Cell* **16,** 269–279 (2004).
77. Dalhaimer, P., Pollard, T. D. & Nolen, B. J. Nucleotide-mediated conformational changes of monomeric actin and Arp3 studied by molecular dynamics simulations. *Journal of Molecular Biology* **376,** 166–183 (2008).
78. Ti, S.-C., Jurgenson, C. T., Nolen, B. J. & Pollard, T. D. Structural and biochemical characterization of two binding sites for nucleation-promoting factor WASp-VCA on Arp2/3 complex. *Proc. Natl. Acad. Sci. U.S.A.* **108,** E463–71 (2011).
79. Choi, C. L., Claridge, S. A., Garner, E. C., Alivisatos, A. P. & Mullins, R. D. Protein-nanocrystal conjugates support a single filament polymerization model in R1 plasmid segregation. *J. Biol. Chem.* **283,** 28081–28086 (2008).
80. Theriot, J. A. *Why are bacteria different from eukaryotes? BMC biology* **11,** 119 (BioMed Central, 2013).
81. Boldogh, I. R. & Pon, L. A. Interactions of mitochondria with the actin cytoskeleton. *Biochim. Biophys. Acta* **1763,** 450–462 (2006).
82. Massarwa, R., Schejter, E. D. & Shilo, B.-Z. Apical secretion in epithelial tubes of the Drosophila embryo is directed by the Formin-family protein Diaphanous. *Dev. Cell* **16,** 877–888 (2009).
83. Gayathri, P. *et al.* A bipolar spindle of antiparallel ParM filaments drives bacterial plasmid segregation. *Science* **338,** 1334–1337 (2012).
84. Knoblach, B. & Rachubinski, R. A. Sharing the cell's bounty - organelle inheritance in yeast. *J. Cell. Sci.* **128,** 621–630 (2015).
85. Jiang, S., Ghoshdastider, U., Narita, A., Popp, D. & Robinson, R. C. Structural complexity of filaments formed from the actin and tubulin folds. *Commun Integr Biol* **9,** e1242538 (2016).
86. Grosse, R. & Vartiainen, M. K. To be or not to be assembled: progressing into nuclear actin filaments. *Nat. Rev. Mol. Cell Biol.* **14,** 693–697 (2013).
87. Gordon, D. J., Eisenberg, E. & Korn, E. D. Characterization of cytoplasmic actin isolated from Acanthamoeba castellanii by a new method. *J. Biol. Chem.* **251,** 4778–4786 (1976).
88. Clerc, P. & Sansonetti, P. J. Entry of Shigella flexneri into HeLa cells: evidence for directed phagocytosis involving actin polymerization and myosin accumulation. *Infect. Immun.* **55,** 2681–2688 (1987).
89. Nemethova, M., Auinger, S. & Small, J. V. Building the actin cytoskeleton: filopodia contribute to the construction of contractile bundles in the lamella. *J. Cell Biol.* **180,** 1233–1244 (2008).
90. Dmitrieff, S. & Nédélec, F. Amplification of actin polymerization forces. *J. Cell Biol.* **212,** 763–766 (2016).
91. Theriot, J. A. The polymerization motor. *Traffic* **1,** 19–28 (2000).

92. Balasubramanian, M. K., Srinivasan, R., Huang, Y. & Ng, K.-H. Comparing contractile apparatus-driven cytokinesis mechanisms across kingdoms. *Cytoskeleton (Hoboken)* **69,** 942–956 (2012).
93. HUXLEY, A. F. & NIEDERGERKE, R. Structural changes in muscle during contraction; interference microscopy of living muscle fibres. *Nature* **173,** 971–973 (1954).
94. Lenz, M. Geometrical Origins of Contractility in Disordered Actomyosin Networks. *Physical Review X* **4,** 041002 (2014).
95. Jacinto, A. *et al.* Dynamic Analysis of Actin Cable Function during Drosophila Dorsal Closure. *Current Biology* **12,** 1245–1250 (2002).
96. Ma, X., Ehrhardt, D. W. & Margolin, W. Colocalization of cell division proteins FtsZ and FtsA to cytoskeletal structures in living Escherichia coli cells by using green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.* **93,** 12998–13003 (1996).
97. Bisson-Filho, A. W. *et al.* Treadmilling by FtsZ filaments drives peptidoglycan synthesis and bacterial cell division. *Science* **355,** 739–743 (2017).
98. Loose, M. & Mitchison, T. J. The bacterial cell division proteins FtsA and FtsZ self-organize into dynamic cytoskeletal patterns. *Nat. Cell Biol.* **16,** 38–46 (2014).
99. Egan, A. J. F. & Vollmer, W. The stoichiometric divisome: a hypothesis. *Front Microbiol* **6,** 455 (2015).
100. Otegui, M. & Staehelin, L. A. Cytokinesis in flowering plants: more than one way to divide a cell. *Curr. Opin. Plant Biol.* **3,** 493–502 (2000).
101. Hayles, J. *et al.* A genome-wide resource of cell cycle and cell shape genes of fission yeast. *Open Biol* **3,** 130053–130053 (2013).
102. Urnavicius, L. *et al.* The structure of the dynactin complex and its interaction with dynein. *Science* **347,** 1441–1446 (2015).
103. Gill, S. R. *et al.* Dynactin, a conserved, ubiquitously expressed component of an activator of vesicle motility mediated by cytoplasmic dynein. *J. Cell Biol.* **115,** 1639–1650 (1991).
104. Plamann, M., Minke, P. F., Tinsley, J. H. & Bruno, K. S. Cytoplasmic dynein and actin-related protein Arp1 are required for normal nuclear distribution in filamentous fungi. *J. Cell Biol.* **127,** 139–149 (1994).
105. de Lanerolle, P. Nuclear actin and myosins at a glance. *J. Cell. Sci.* **125,** 4945–4949 (2012).
106. Illergård, K., Ardell, D. H. & Elofsson, A. Structure is three to ten times more conserved than sequence--a study of structural response in protein cores. *Proteins* **77,** 499–508 (2009).
107. Remmert, M., Biegert, A., Hauser, A. & Söding, J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods* **9,** 173–175 (2011).
108. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30,** 772–780 (2013).

109. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25,** 1972–1973 (2009).
110. Nguyen, L.-T., Schmidt, H. A., Haeseler, von, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32,** 268–274 (2015).
111. Quang, L. S., Gascuel, O. & Lartillot, N. Empirical profile mixture models for phylogenetic reconstruction. *Bioinformatics* **24,** 2317–2323 (2008).
112. Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25,** 1307–1320 (2008).
113. Yang, Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* **39,** 306–314 (1994).
114. Braun, T. *et al.* Archaeal actin from a hyperthermophile forms a single-stranded filament. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 9340–9345 (2015).
115. Zaremba-Niedzwiedzka, K. *et al.* Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541,** 353–358 (2017).
116. Izoré, T., Duman, R., Kureisaite-Ciziene, D. & Löwe, J. Crenactin from Pyrobaculum calidifontis is closely related to actin in structure and forms steep helical filaments. *FEBS Lett.* **588,** 776–782 (2014).
117. Izoré, T., Kureisaite-Ciziene, D., McLaughlin, S. H. & Löwe, J. Crenactin forms actin-like double helical filaments regulated by arcadin-2. *Elife* **5,** 213 (2016).
118. Doolittle, R. F. The origins and evolution of eukaryotic proteins. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **349,** 235–240 (1995).
119. Ingerson-Mahar, M., Briegel, A., Werner, J. N., Jensen, G. J. & Gitai, Z. The metabolic enzyme CTP synthase forms cytoskeletal filaments. *Nat. Cell Biol.* **12,** 739–746 (2010).
120. Edelstein, S. J., Telford, J. N. & Crepeau, R. H. Structure of fibers of sickle cell hemoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **70,** 1104–1107 (1973).
121. Italiano, J. E., Stewart, M. & Roberts, T. M. How the assembly dynamics of the nematode major sperm protein generate amoeboid cell motility. *Int. Rev. Cytol.* **202,** 1–34 (2001).
122. Pollard, T. D. & Cooper, J. A. Actin, a central player in cell shape and movement. *Science* **326,** 1208–1212 (2009).
123. Franzmann, T. M. *et al.* Phase separation of a yeast prion protein promotes cellular fitness. *Science* **359,** eaao5654 (2018).
124. Aughey, G. N. & Liu, J.-L. Metabolic regulation via enzyme filamentation. *Crit. Rev. Biochem. Mol. Biol.* **51,** 282–293 (2015).
125. Petrovska, I. *et al.* Filament formation by metabolic enzymes is a specific adaptation to an advanced state of cellular starvation. *Elife* **3,** 6063 (2014).
126. Maitra, P. K. A glucokinase from Saccharomyces cerevisiae. *J. Biol. Chem.* **245,** 2423–2431 (1970).
127. CAPUTTO, R. & LELOIR, L. R. The enzymatic transformation of galactose into glucose derivatives. *J. Biol. Chem.* **179,** 497 (1949).
128. Lynch, E. M. *et al.* Human CTP synthase filament structure reveals the active enzyme conformation. *Nat. Struct. Mol. Biol.* **24,** 507–514 (2017).

129. Hilitski, F. *et al.* Measuring cohesion between macromolecular filaments one pair at a time: depletion-induced microtubule bundling. *Phys. Rev. Lett.* **114,** 138102 (2015).
130. Pragay, D. A. & Gergely, J. Effect of tropomyosin on the polymerization of ATP--G-actin and ADP--G-actin. *Arch. Biochem. Biophys.* **125,** 727–733 (1968).
131. Landgraf, D., Okumus, B., Chien, P., Baker, T. A. & Paulsson, J. Segregation of molecules at cell division reveals native protein localization. *Nat. Methods* **9,** 480–482 (2012).
132. Yachdav, G. *et al.* PredictProtein--an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res.* **42,** W337–43 (2014).
133. Dosztányi, Z. Prediction of protein disorder based on IUPred. *Protein Sci.* **27,** 331–340 (2018).
134. Linding, R. *et al.* Protein disorder prediction: implications for structural proteomics. *Structure* **11,** 1453–1459 (2003).
135. Hagman, A., Säll, T., Compagno, C. & Piskur, J. Yeast 'make-accumulate-consume' life strategy evolved as a multi-step process that predates the whole genome duplication. *PLoS ONE* **8,** e68734 (2013).
136. Zhang, J. *et al.* Lys169 of human glucokinase is a determinant for glucose phosphorylation: implication for the atomic mechanism of glucokinase catalysis. *PLoS ONE* **4,** e6304 (2009).
137. Bell, W. *et al.* Composition and functional analysis of the Saccharomyces cerevisiae trehalose synthase complex. *J. Biol. Chem.* **273,** 33311–33319 (1998).
138. Blázquez, M. A., Lagunas, R., Gancedo, C. & Gancedo, J. M. Trehalose-6-phosphate, a new regulator of yeast glycolysis that inhibits hexokinases. *FEBS Lett.* **329,** 51–54 (1993).
139. van Eunen, K. *et al.* Measuring enzyme activities under standardized in vivo-like conditions for systems biology. *FEBS J.* **277,** 749–760 (2010).
140. Shevchuk, N. A. *et al.* Construction of long DNA molecules using long PCR-based fusion of several fragments simultaneously. *Nucleic Acids Res.* **32,** e19–19 (2004).
141. Schiestl, R. H. & Gietz, R. D. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr. Genet.* **16,** 339–346 (1989).
142. Gibson, D. G. Enzymatic assembly of overlapping DNA fragments. *Meth. Enzymol.* **498,** 349–361 (2011).
143. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9,** 671–675 (2012).
144. Malakhov, M. P. *et al.* SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J. Struct. Funct. Genomics* **5,** 75–86 (2004).
145. Kuettner, E. B. *et al.* Crystal structure of hexokinase KlHxk1 of Kluyveromyces lactis: a molecular basis for understanding the control of yeast hexokinase functions via covalent modification and oligomerization. *J. Biol. Chem.* **285,** 41019–41033 (2010).

146. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501 (2010).
147. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68,** 352–367 (2012).
148. Boratyn, G. M. *et al.* BLAST: a more efficient report with usability improvements. *Nucleic Acids Res.* **41,** W29–33 (2013).
149. Chojnacki, S., Cowley, A., Lee, J., Foix, A. & Lopez, R. Programmatic access to bioinformatics tools from EMBL-EBI update: 2017. *Nucleic Acids Res.* **45,** W550–W553 (2017).
150. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25,** 1189–1191 (2009).
151. Zhu, A., Romero, R. & Petty, H. R. An enzymatic colorimetric assay for glucose-6-phosphate. *Anal. Biochem.* **419,** 266–270 (2011).
152. Taylor, T. N., Hass, H. & Kerp, H. The oldest fossil ascomycetes. *Nature* **399,** 648–648 (1999).
153. Conant, G. C. & Wolfe, K. H. Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Mol. Syst. Biol.* **3,** 129 (2007).
154. Espinosa-Cantú, A., Ascencio, D., Barona-Gómez, F. & DeLuna, A. Gene duplication and the evolution of moonlighting proteins. *Front Genet* **6,** 227 (2015).
155. Naseeb, S., Ames, R. M., Delneri, D. & Lovell, S. C. Rapid functional and evolutionary changes follow gene duplication in yeast. *Proc. Biol. Sci.* **284,** 20171393 (2017).
156. Enyenihi, A. H. & Saunders, W. S. Large-scale functional genomic analysis of sporulation and meiosis in Saccharomyces cerevisiae. *Genetics* **163,** 47–54 (2003).
157. Griffith, L. M. & Pollard, T. D. Cross-linking of actin filament networks by self-association and actin-binding macromolecules. *J. Biol. Chem.* **257,** 9135–9142 (1982).
158. Los, G. V. *et al.* HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3,** 373–382 (2008).
159. Krogan, N. J. *et al.* Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440,** 637–643 (2006).
160. Dashko, S., Zhou, N., Compagno, C. & Piskur, J. Why, when, and how did yeast evolve alcoholic fermentation? *FEMS Yeast Res.* **14,** 826–832 (2014).
161. Hagman, A., Säll, T. & Piskur, J. Analysis of the yeast short-term Crabtree effect and its origin. *FEBS J.* **281,** 4805–4814 (2014).
162. Garcia-Seisdedos, H., Empereur-Mot, C., Elad, N. & Levy, E. D. Proteins evolve on the edge of supramolecular self-assembly. *Nature* **548,** 244–247 (2017).
163. Dykes, G. W., Crepeau, R. H. & Edelstein, S. J. Three-dimensional reconstruction of the 14-filament fibers of hemoglobin S. *Journal of Molecular Biology* **130,** 451–472 (1979).
164. Ewald, J. C., Kuehne, A., Zamboni, N. & Skotheim, J. M. The Yeast Cyclin-Dependent Kinase Routes Carbon Fluxes to Fuel Cell Cycle Progression. *Mol. Cell* **62,** 532–545 (2016).

165. Zhao, G., Chen, Y., Carey, L. & Futcher, B. Cyclin-Dependent Kinase Co-Ordinates Carbohydrate Metabolism and Cell Cycle in S. cerevisiae. *Mol. Cell* **62,** 546–557 (2016).
166. Dey, G., Thattai, M. & Baum, B. On the Archaeal Origins of Eukaryotes and the Challenges of Inferring Phenotype from Genotype. *Trends Cell Biol.* **26,** 476–485 (2016).
167. Kraemer, J. A. *et al.* A phage tubulin assembles dynamic filaments by an atypical mechanism to center viral DNA within the host cell. *Cell* **149,** 1488–1499 (2012).
168. Duggin, I. G. *et al.* CetZ tubulin-like proteins control archaeal cell shape. *Nature* **519,** 362–365 (2015).