Mechanisms of amino acid sensing by the mTORC1 pathway

by

Robert A. Saxton

B.A., Molecular and Cell Biology
University of California, Berkeley

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2018

© Robert A. Saxton. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis in whole or in part in any medium now known or hereafter created.

Signature of Author:

Signature redacted

Robert A. Saxton
Department of Biology

Certified by:

Signature redacted

David M. Sabatini
Professor of Biology
Member, Whitehead Institute

Certified by:

Signature redacted

Thomas U. Schwartz
Boris Magasanik Professor of Biology

Accepted by:

Signature redacted

Amy Keating
Professor of Biology
Co-Chair, Biology Graduate Committee
Mechanisms of amino acid sensing by the mTORC1 pathway

By

Robert A. Saxton

Abstract

In order to coordinate growth with organismal physiology, cells must couple anabolic and catabolic metabolism with a variety of environmental cues. In humans, cell growth is controlled primarily by the mechanistic Target of Rapamycin Complex 1 (mTORC1), which senses a diverse array of environmental inputs and in turn regulates nearly all aspects of growth and metabolism. Consistent with its central role in these processes, deregulation of the mTORC1 pathway is also associated with numerous diseases, including epilepsy, diabetes, and cancer. Although decades of study have revealed how growth factor signaling networks regulate mTORC1, the components that detect and signal nutrient availability to mTORC1 have been elusive until the recent discovery of the leucine and arginine sensors, named Sestrin2 and CASTOR1, respectively. However, the molecular mechanisms through which these components detect their respective amino acids have been elusive.

To understand how Sestrin2 senses leucine and signals its presence to mTORC1, we solved the X-ray crystal structure of Sestrin2 in complex with leucine to 2.7 Å resolution. This structure revealed how Sestrin2 recognizes leucine through a highly specific binding pocket, and utilizes a unique “lid-latch” mechanism to sense leucine. This structural data also allowed us to identify a highly conserved GATOR2-binding site on the surface of Sestrin2, in close proximity to the leucine-binding pocket, providing a plausible allosteric mechanism for the leucine-dependent dissociation of Sestrin2 from GATOR2.

Using a similar approach, we then solved the crystal structure of arginine-bound CASTOR1 to 1.8 Å resolution. This structure showed that CASTOR1 forms a homodimeric complex that engages arginine at the interface of two domains. Combining this structural information with various biochemical assays, we found that arginine acts as “molecular glue” that bridges these two domains together to allosterically modulate the adjacent GATOR2 binding site, thereby explaining the molecular basis for arginine-dependent mTORC1 activation.

Together, the structural and biochemical characterization of Sestrin2 and CASTOR1 presented here have revealed for the first time the mechanisms of mammalian amino acid sensing in atomic detail. These structures have also paved the way for the development of novel compounds targeting mTORC1 through the Sestrin2 and CASTOR1 amino acid-binding pockets, which may have important clinical implications in the context of cancer, diabetes, neurological disorders, and longevity.

Thesis co-supervisor: David M. Sabatini
Title: Professor of Biology

Thesis co-supervisor: Thomas U. Schwartz
Title: Boris Magasanik Professor of Biology
Acknowledgements

First and foremost I would like to thank my thesis advisors David Sabatini and Thomas Schwartz, who have provided me with incredible mentorship and given me the support and opportunities necessary to succeed. As both teachers and role models, David and Thomas have shown me the traits needed to do interesting, important, and rigorous science, which I hope to carry with me throughout my scientific career. I am also grateful to all of the lab members in both the Sabatini and Schwartz lab, who have created excellent research environments. In particular, Tim Wang, Kevin Knockenhauer, Lynne Chantranupong, Rachel Wolfson, and Michael Pacold have all contributed enormously to my graduate career as both scientific mentors and friends. Finally, I would like to thank my parents, grandparents, and Natasha, who have provided me with endless support and encouragement throughout graduate school.
# Table of Contents

Title page 1

Abstract 3

Acknowledgements 5

Chapter 1: Introduction 7

Chapter 2: Structural basis for leucine sensing by the Sestrin2-mTORC1 pathway 52

Chapter 3: Analysis of reported “apo-Sestrin2” structures 84

Chapter 4: Mechanism of arginine sensing by CASTOR1 upstream of mTORC1 97

Chapter 5: Summary and future directions 127
Chapter 1

Introduction

Parts of this chapter were first published as:


Section 1. Introduction

A fundamental challenge for all living organisms is to appropriately adapt their metabolism and behavior to changing environmental conditions. In particular, organisms must alter their energy expenditure, nutrient uptake, and growth rate depending on the presence of energy sources, nutrients, or noxious stressors. In multicellular eukaryotes, individual cells must also integrate information from both the external world and various tissues in the body, and adjust their metabolic behavior accordingly. Failures of cellular communication systems leading to a disruption of metabolic coordination is observed in a wide range of human medical conditions, including diabetes, metabolic syndrome, cancer, and aging. Understanding how organisms efficiently coordinate their growth and metabolism with diverse environmental signals is therefore a major question in biology with significant biomedical implications.

Studies into the mechanism of action of the macrolide antibiotic Rapamycin in the early 1990s led to the identification of a highly conserved protein kinase now known as the mechanistic Target of Rapamycin (mTOR). Subsequent studies of the cellular functions of mTOR have revealed that it lies at the center of a complex signaling network that plays a central role in coordinating cell growth and metabolism with changing environmental and physiological conditions. The mTOR signaling network senses and integrates a diverse range of nutrients and hormonal cues, and in turn regulates nearly all aspects of cellular and organismal metabolism. Given these key biological roles, the questions of how the mTOR pathway senses and
integrates environmental signals, and how these pathways can be therapeutically targeted in the context of human disease, have been the subject of intense study.

The two most well established inputs that regulate mTOR signaling are growth factors, such as insulin, and nutrients like glucose and amino acids. In particular, the amino acids leucine and arginine have been shown to be very potent activators of mTORC1 signaling in a variety of tissues and organisms, though how these molecules are specifically sensed upstream of mTOR has been mysterious.

In this thesis, I describe our work to structurally and biochemically characterize the mechanisms underlying leucine and arginine sensing by the mTOR pathway in human cells. In order to provide appropriate context for these findings, in this chapter I first introduce the mTOR signaling network, describe its cellular and physiological roles, and provide an overview of the nutrient sensing pathway upstream of mTOR.

Section II. The mTOR Signaling Network

A. Discovery of mTOR

In 1964 a Canadian expedition to the isolated South Pacific island of Rapa Nui (also known as Easter Island) collected a set of soil samples with the goal of identifying novel antimicrobial agents. In bacteria isolated from one of these samples, Sehgal and colleagues discovered a compound with remarkable antifungal, immunosuppressive, and antitumor properties (Eng et. al. 1984; Martel et. al. 1977; Vezina et. al 1975). Further analysis of this compound, named rapamycin after its site of discovery (clinically referred to as sirolimus), revealed that it acts in part by forming a gain of function complex with the peptidyl-prolyl-isomerase FKBP12 to inhibit signal transduction pathways required for cell growth and proliferation (Chung et. al., 1992).

Despite these insights, the full mechanism of action of rapamycin remained elusive until 1994 when biochemical studies identified the mechanistic (formerly "mammalian") Target of
Rapamycin (mTOR) as the direct target of the rapamycin-FKBP12 complex in mammals (Brown et. al. 1994; Sabatini et. al. 1994; Sabers et. al 1995), and revealed it to be the homolog of the yeast TOR/DRR genes that had previously been identified in genetic screens for rapamycin resistance (Cafferkey et. al. 1993; Heitman et. al. 1991; Kunz et. al. 1993).

B. mTORC1 and mTORC2

mTOR is a serine/threonine protein kinase in the PI3K-related kinase (PIKK) family that forms the catalytic subunit of two distinct protein complexes, known as mTOR Complex 1 (mTORC1) and 2 (mTORC2) (Fig. 1A). mTORC1 is defined by its three core components: mTOR, Raptor (regulatory protein associated with mTOR), and mLST8 (mammalian lethal with Sec13 protein 8, also known as GβL) (Fig. 1B, Kim et al. 2002; Hara et al. 2002; Kim et al. 2003). Raptor facilitates substrate recruitment to mTORC1 through binding to the TOR signaling (TOS) motif found on several canonical mTORC1 substrates (Nojima et al., 2003; Schalm et al., 2003), and, as described later, is required for the correct subcellular localization of mTORC1. mLST8 by contrast associates with the catalytic domain of mTORC1 and may stabilize the kinase activation loop (Yang et al. 2013), though genetic studies suggest it is dispensible for the essential functions of mTORC1 (Guertin et al., 2006). In addition to these three core components, mTORC1 also contains the two inhibitory subunits PRAS40 (proline-rich Akt substrate of 40 kDa) (Sancak et al. 2007; Vander Haar et al. 2007; Wang et al. 2007) and DEPTOR (DEP domain containing mTOR interacting protein) (Peterson et al. 2009).

Structural studies of mTORC1 have yielded significant insights into its assembly, function, and perturbation by rapamycin. Cryo-EM reconstructions of both mTORC1 and yeast
Figure 1. mTORC1 and mTORC2

(A) The mTORC1 and mTORC2 signaling pathways. (B) mTORC1 subunits and respective binding sites on mTOR. The FKBP12-rapamycin. The 5.9 Å cryo-EM structure of mTORC1 (without DEPTOR and PRAS40, PDB ID: 5FLC) is depicted as a space filling model and colored by subunit. (C) mTORC2 subunits and respective binding sites on mTOR.

TORC1 have revealed that the complex forms a 1 mDa “lozenge”-shaped dimer, with the dimerization interface comprised of contacts between the mTOR HEAT repeats as well as between Raptor and mTOR (Fig. 1B, Aylett et al. 2016; Baretic et al., 2016; Yip et al. 2010). In addition, a crystal structure of the mTOR kinase domain bound to mLST8 showed that the rapamycin-FKBP12 complex binds to the FRB domain of mTOR to narrow the catalytic cleft and partially occlude substrates from the active site (Yang et al., 2013).

While the rapamycin-FKBP12 complex directly inhibits mTORC1, mTORC2 is characterized by its insensitivity to acute rapamycin treatment. Like mTORC1, mTORC2 also contains mTOR and mLST8 (Fig. 1C). Instead of Raptor however, mTORC2 contains Rictor (rapamycin insensitive companion of mTOR), an unrelated protein that likely serves an
analogous function (Jacinto et al. 2004; Sarbassov et al. 2004). mTORC2 also contains DEPTOR (Peterson et al. 2009), as well as the regulatory subunits mSin1 (Frias et al. 2006; Jacinto et al. 2006; Yang et al. 2006) and Protor1/2 (Pearce et al. 2007; Thedieck et al. 2007; Woo et al., 2007). Although rapamycin-FKBP12 complexes do not directly bind or inhibit mTORC2, prolonged rapamycin treatment does abrogate mTORC2 signaling, likely due to the inability of rapamycin-bound mTOR to incorporate into new mTORC2 complexes (Lamming et al., 2012).

C. Downstream of mTORC1

In order to grow and divide, cells must increase production of proteins, lipids, and nucleotides while also suppressing catabolic pathways such as autophagy. mTORC1 plays a central role in regulating all of these processes, and therefore controls the balance between anabolism and catabolism in response to environmental conditions (Fig. 2, A and B). Here we review the critical substrates and cellular processes downstream of mTORC1 and how they contribute to cell growth. Most of the functions discussed here were identified and characterized in the context of mammalian cell lines, while the physiological context in which these processes are important will be discussed in greater detail later.

Protein synthesis: mTORC1 promotes protein synthesis largely through the phosphorylation of two key effectors, p70S6 Kinase 1 (S6K1) and eIF4E Binding Protein (4EBP) (Fig 2B). mTORC1 directly phosphorylates S6K1 on its hydrophobic motif site, Thr389, enabling its subsequent phosphorylation and activation by PDK1. S6K1 phosphorylates and activates several substrates that promote mRNA translation initiation including eIF4B, a positive regulator of the 5’cap binding eIF4F complex (Holz et al., 2005). S6K1 also phosphorylates and promotes the degradation of PDCD4, an inhibitor of eIF4B (Dorello et al. 2006), and enhances the translation efficiency of spliced mRNAs via its interaction with SKAR, a component of exon-junction complexes (Ma et al., 2008).
The mTORC1 substrate 4EBP is unrelated to S6K1 and inhibits translation by binding and sequestering eIF4E to prevent assembly of the eIF4F complex. mTORC1 phosphorylates 4EBP at multiple sites to trigger its dissociation from eIF4E (Brunn et al., 1997; Gingras et al. 1999), allowing 5’cap-dependent mRNA translation to occur. Although it has long been appreciated that mTORC1 signaling regulates mRNA translation, whether and how it affects specific classes of mRNA transcripts has been debated. Global ribosome foot printing analyses however revealed that while acute mTOR inhibition moderately suppresses general mRNA translation, it most profoundly affects mRNAs containing pyrimidine-rich 5’ TOP or “TOP-like” motifs, which includes most genes involved in protein synthesis (Hsieh et al. 2012; Thoreen et al. 2012).

**Lipid, nucleotide, and glucose metabolism:** Growing cells require sufficient lipids for new membrane formation and expansion. mTORC1 promotes de novo lipid synthesis through the sterol responsive element binding protein (SREBP) transcription factors, which control the expression of metabolic genes involved in fatty acid and cholesterol biosynthesis (Porstmann et al., 2008). While SREBP is canonically activated in response to low sterol levels, mTORC1 signaling can also activate SREBP independently through both an S6K1-dependent mechanism (Duvel et al., 2010) as well as through the phosphorylation of an additional substrate, Lipin1, which inhibits SREBP in the absence of mTORC1 signaling (Peterson et al., 2011).

Recent studies established that mTORC1 also promotes the synthesis of nucleotides required for DNA replication and ribosome biogenesis in growing and proliferating cells. mTORC1 increases the ATF4-dependent expression of MTHFD2, a key component of the mitochondrial tetrahydrofolate cycle that provides one-carbon units for purine synthesis (Ben-Sahra et al., 2016). Additionally, S6K1 phosphorylates and activates carbamoyl-phosphate synthetase (CAD), a critical component of the de novo pyrimidine synthesis pathway (Ben-Sahra et al., 2013; Robataille et al., 2013).
mTORC1 also facilitates growth by promoting a shift in glucose metabolism from oxidative phosphorylation to glycolysis, which likely facilitates the incorporation of nutrients into new biomass. mTORC1 increases the translation of the transcription factor HIF1α (Fig. 2C), which drives the expression of several glycolytic enzymes such as phospho-fructo kinase (PFK) (Duvel et al., 2010). Furthermore, mTORC1-dependent activation of SREBP leads to increased flux through the oxidative pentose phosphate pathway (PPP), which utilizes carbons from glucose to generate NADPH and other intermediary metabolites needed for proliferation and growth.

**Regulation of protein turnover:** In addition to the various anabolic processes outlined above, mTORC1 also promotes cell growth by suppressing protein catabolism (Fig. 1B), most notably autophagy. An important early step in autophagy is the activation of ULK1, a kinase that forms a complex with ATG13, FIP2000, and ATG101 and drives autophagosome formation. Under nutrient replete conditions, mTORC1 phosphorylates ULK1, thereby preventing its activation by AMPK, a key activator or autophagy (Kim et al., 2011). Thus, the relative activity of mTORC1 and AMPK in different cellular contexts largely determines the extent of autophagy induction. mTORC1 also regulates autophagy in part by phosphorylating and inhibiting the nuclear translocation of the transcription factor TFEB, which drives the expression of genes for lysosomal biogenesis and the autophagy machinery (Martina et al., 2012; Rożniewski-Ferguson et al., 2012; Settembre et al., 2012).

The second major pathway responsible for protein turnover is the ubiquitin-proteasome system (UPS), through which proteins are selectively targeted for degradation by the 20S proteasome following covalent modification with ubiquitin. Two recent studies found that acute mTORC1 inhibition rapidly increases proteasome-dependent proteolysis through either a general increase in protein ubiquitylation, or an increased abundance of proteasomal chaperones via inhibition of Erk5 (Fig. 2B, Rousseau et al., 2016, Zhao et al., 2015). However, another study found that genetic hyper-activation of mTORC1 signaling also increases
proteasome activity, through elevated expression of proteasome subunits downstream of Nrf1 (Zhang et al., 2014). One possible explanation for this discrepancy is that while acute mTORC1 inhibition promotes proteolysis to restore free amino acid pools, prolonged mTORC1 activation also triggers a compensatory increase in protein turnover to balance the increased rate of protein synthesis. Given that the UPS is responsible for the majority of protein degradation in human cells, precisely how mTORC1 regulates this process is an important question going forward.

D. Downstream of mTORC2

While mTORC1 regulates cell growth and metabolism, mTORC2 instead controls proliferation and survival primarily by phosphorylating several members of the AGC (PKA/PKG/PKC) family of protein kinases (Fig. 2D). The first mTORC2 substrate to be identified was PKCa, a regulator of the actin cytoskeleton (Jacinto et al., 2004, Sarbassov et al., 2004). More recently, mTORC2 has also been shown to phosphorylate several other members of the PKC family, including PKCδ (Gan et al., 2012), PKCζ (Li and Gao, 2014), as well as PKCy and PKCc (Thomanetz et al., 2013), all of which regulate various aspects of cytoskeletal remodeling and cell migration.

The most important role of mTORC2 however is likely the phosphorylation and activation of Akt, a key effector of insulin/PI3K signaling (Sarbassov et al., 2005). Once active, Akt promotes cell survival, proliferation, and growth through the phosphorylation and inhibition of several key substrates including the FoxO1/3a transcription factors, the metabolic regulator GSK3β, and the mTORC1 inhibitor TSC2. However while mTORC2-dependent phosphorylation is required for Akt to phosphorylate some substrates in vivo, such as FoxO1/3a, it is dispensable for the phosphorylation of others including TSC2 (Guertin et al., 2006; Jacinto et al., 2006). Finally, mTORC2 also phosphorylates and activates SGK1, another AGC-kinase that regulates ion transport as well as cell survival (Garcia-Martinez and Alessi, 2008).
E. Upstream of mTORC1

The mTORC1-dependent shift towards increased anabolism should only occur in the presence of pro-growth endocrine signals as well as sufficient energy and chemical building blocks for macromolecular synthesis. In mammals, these inputs are largely dependent on diet, such that mTORC1 is activated following feeding to promote growth and energy storage in tissues such as the liver and muscle, but inhibited during fasting conserve limited resources. Here we discuss the cellular pathways upstream of mTORC1 and the mechanisms through which they control mTORC1 activation.

Growth Factors: Studies of rapamycin in the early 1990s revealed that mTORC1 is a downstream mediator of several growth factor and mitogen-dependent signaling pathways, all of which inhibit a key negative regulator of mTORC1 signaling known as the Tuberous Sclerosis Complex (TSC) complex. TSC is a heterotrimeric complex comprised of TSC1, TSC2, and TBC1D7 (Dibble et al., 2012), and functions as a GTPase activating protein (GAP) for the small GTPase Rheb (Inoki et al., 2003; Tee et al., 2003), which directly binds and activates mTORC1 (Long et al., 2005; Sancak et al., 2007). Although Rheb is an essential activator of mTORC1, exactly how it stimulates mTORC1 kinase activity remains unknown.

Numerous growth factor pathways converge on TSC (Fig. 2A), including the insulin/insulin-like growth factor-1 (IGF-1) pathway, which triggers the Akt-dependent multisite phosphorylation of TSC2 (Inoki et al., 2002; Manning et al., 2002). This phosphorylation inhibits TSC by dissociating it from the lysosomal membrane, where at least some fraction of cellular Rheb localizes (Menon et al., 2014). Similarly, receptor tyrosine kinase-dependent Ras signaling activates mTORC1 via the MAP Kinase Erk and its effector p90RSK, both of which also phosphorylate and inhibit TSC2 (Ma et al., 2005; Roux et al., 2004). It is unclear however whether these inputs also control the localization of TSC, or rather inhibit its GAP activity through a distinct mechanism. Additional growth factor pathways upstream of TSC include Wnt and the inflammatory cytokine TNFα, both of which activate mTORC1 through the inhibition of...
TSC1 (Inoki et al., 2006; Lee et al., 2007). Precisely how the TSC complex integrates these numerous signals and their relative impact on mTORC1 activity in various contexts however remains an open question.

**Energy, oxygen, and DNA damage:** mTORC1 also responds to intracellular and environmental stresses that are incompatible with growth such as low ATP levels, hypoxia, or DNA damage. A reduction in cellular energy charge, for example during glucose deprivation, activates the stress responsive metabolic regulator AMPK, which inhibits mTORC1 both indirectly, through phosphorylation and activation of TSC2, as well as directly through the phosphorylation of Raptor (Gwinn et al., 2008; Inoki et al., 2003b; Shaw et al., 2004). Interestingly, glucose deprivation also inhibits mTORC1 in cells lacking AMPK, through inhibition of the Rag GTPases, suggesting that mTORC1 senses glucose through more than one mechanism (Efeyan et al., 2013; Kalender et al., 2010). Similarly, hypoxia inhibits mTORC1 in part through AMPK activation, but also through the induction of REDD1 (Regulated in DNA damage and development 1), which activates TSC (Brugarolas et al., 2004). Finally, the DNA damage-response pathway inhibits mTORC1 through the induction of p53 target genes including the AMPK regulatory subunit (AMPKβ), PTEN, and TSC2 itself, all of which increase TSC activity (Feng et al., 2007).

**Amino Acids:** In addition to glucose-dependent insulin release, feeding also leads to an increase in serum amino acid levels due to the digestion of dietary proteins. As amino acids are not only essential building blocks of proteins but also sources of energy and carbon for many other metabolic pathways, mTORC1 activation is tightly coupled to diet-induced changes in amino acid concentrations.

A breakthrough in the understanding of amino acid sensing by mTORC1 came with the discovery of the heterodimeric Rag GTPases as components of the mTORC1 pathway (Kim et al., 2008; Sancak et al., 2008). The Rags are obligate heterodimers of RagA or RagB with RagC or RagD, and are tethered to the lysosomal membrane through their association with the
pentameric Ragulator complex comprised of MP1, p14, p18, HBXIP and c7ORF59 (Sancak et al., 2010; Bar-Peled et al., 2012). Amino acid stimulation converts the Rags to their active nucleotide-bound state, allowing them to bind Raptor and recruit mTORC1 to the lysosomal surface, where Rheb is also located. This pathway architecture therefore forms an “AND-gate”, whereby mTORC1 signaling is only on when both the Rags and Rheb are activated, explaining why both growth factors and amino acids are required for mTORC1 activation.

Despite these insights, the identities of the direct amino acid sensors upstream of mTORC1 have been elusive until very recently. It is now clear that mTORC1 senses both intra-lysosomal and cytosolic amino acids through distinct mechanisms. Amino acids inside the lysosomal lumen alter the Rag nucleotide state through a mechanism dependent on the lysosomal v-ATPase, which interacts the Ragulator-Rag complex to promote the guanine-nucleotide exchange factor (GEF) activity of Ragulator towards RagA/B (Zoncu et al., 2011; Bar-Peled et al., 2012). The lysosomal amino acid transporter SLC38A9 interacts with the Rag-Ragulator-v-ATPase complex and is required for arginine to activate mTORC1, making it a promising candidate to be a lysosomal amino acid sensor (Jung et al., 2015; Rebsamen et al., 2015; Wang et al., 2015).

Cytosolic leucine and arginine signal to mTORC1 through a distinct pathway comprised of the GATOR1 and GATOR2 complexes (Bar-Peled et al., 2013). GATOR1 consists of DEPDC5, Nprl2, and Nprl3, and inhibits mTORC1 signaling by acting as a GAP for RagA/B. The recently identified KICSTOR complex (consisting of Kaptin, ITFG2, c12orf66, and SZT2) tethers GATOR1 to the lysosomal surface and is necessary for the appropriate control of the mTORC1 pathway by nutrients (Wolfson et al., 2017). GATOR2 by contrast is a pentameric complex comprised of Mios, WDR24, WDR59, Seh1L and Sec13, and is a positive regulator of mTORC1 signaling that interacts with GATOR1 at the lysosomal membrane (Bar-Peled et al., 2013).

An important insight into the mechanism of cytosolic amino acid sensing came with the identification of Sestrin2 as a GATOR2 interacting protein that inhibits mTORC1 signaling under
amino acid deprivation (Chantranupong et al., 2014; Parmigiani et al., 2014). Subsequent biochemical and structural analyses established that Sestrin2 is a direct leucine sensor upstream of mTORC1 that binds and inhibits GATOR2 function in the absence of leucine, and dissociates from it upon leucine binding (Wolfson et al., 2016). It remains to be seen whether and in what tissues leucine concentrations fluctuate within the relevant range to be sensed by Sestrin2 in vivo, as the levels of interstitial or cytosolic leucine are unknown. Interestingly, another recent study found that Sestrin2 is transcriptionally induced upon prolonged amino acid starvation via the stress-responsive transcription factor ATF4 (Ye et al., 2015), suggesting that Sestrin2 functions as both an acute leucine sensor as well as an indirect mediator of prolonged amino acid starvation.

Cytosolic arginine also activates mTORC1 through the GATOR1/2-Rag pathway by directly binding the recently identified arginine sensor CASTOR1 (Cellular Arginine Sensor for mTORC1). Much like Sestrin2, CASTOR1 binds and inhibits GATOR2 in the absence of arginine, and dissociates upon arginine binding to enable the activation of mTORC1 (Chantranupong et al., 2016). Thus, both leucine and arginine stimulate mTORC1 activity at least in part by releasing inhibitors from GATOR2, establishing GATOR2 as a central node in the signaling of amino acids to mTORC1. Importantly however, the molecular function of GATOR2 and the mechanisms through which Sestrin2 and CASTOR1 regulate it are unknown.

Several additional mechanisms through which amino acids regulate mTORC1 signaling have also recently been reported, including the identification of the Folliculin-FNIP2 complex as a GAP for RagC/D that activates mTORC1 in the presence of amino acids (Petit et al., 2013; Tsun et al., 2013). Another study found that the amino acid glutamine, which is utilized as a nitrogen and energy source by proliferating cells, activates mTORC1 independently of the Rag GTPases through the related Arf family GTPases (Jewell et al., 2015). Finally, a recent report found that the small polypeptide SPAR associates with the v-ATPase-Ragulator complex to
suppress mTORC1 recruitment to lysosomes, though how this occurs is unclear (Matsumoto et al., 2016).

F. Upstream of mTORC2

In contrast to mTORC1, mTORC2 primarily functions as an effector of insulin/PI3K signaling (Fig. 2A). Like most PI3K regulated proteins, the mTORC2 subunit mSin1 contains a phosphoinositide-binding PH domain that is critical for the insulin-dependent regulation of mTORC2 activity. The mSin1 PH domain inhibits mTORC2 catalytic activity in the absence of insulin, and this autoinhibition is relieved upon binding to PI3K-generated PIP₃ at the plasma membrane (Liu et al., 2015). mSin1 can also be phosphorylated by Akt, suggesting the existence of a positive-feedback loop whereby partial activation of Akt promotes the activation of mTORC2, which in turn phosphorylates and fully activates Akt (Yang et al., 2015). Another study found that PI3K promotes the association of mTORC2 with ribosomes to activate its kinase activity, although the mechanistic basis for this is unclear (Zinzalla et al., 2011).

Unexpectedly, mTORC2 signaling is also regulated by mTORC1, due to the presence of a negative feedback loop between mTORC1 and insulin/PI3K signaling. mTORC1 phosphorylates and activates Grb10, a negative regulator of insulin/IGF-1 receptor signaling upstream of Akt and mTORC2, (Hsu et al., 2011; Yu et al., 2011), while S6K1 also suppresses mTORC2 activation through the phosphorylation-dependent degradation of insulin receptor substrate 1 (IRS1) (Harrington et al., 2004; Shah et al., 2004). This negative feedback regulation of PI3K and mTORC2 signaling by mTORC1 has numerous implications for the pharmacological targeting of mTOR in disease, discussed below.
Figure 2. The mTOR Signaling Network

(A) The signaling pathways upstream of mTORC1 and mTORC2. Positive regulators of mTORC1 signaling are shown in yellow, while negative regulators are shown in blue. mTORC1 and mTORC2 are shown in green and red, respectively. (B) The major signaling pathways downstream of mTORC1 signaling in mRNA translation, metabolism, and protein turnover. (C) mTORC1 controls the activity of several transcription factors that can also be independently regulated by cell stress. (D) The major signaling pathways downstream of mTORC2 signaling.
Section III. Evolutionary conservation of the TOR pathway

One remarkable feature of the TOR pathway is its conservation as a major growth regulator in virtually all eukaryotes. Like mammals, *S. cerevisiae* also have two distinct TOR containing complexes, TORC1 and TORC2 (reviewed in Loewith and Hall, 2011) as well as homologs of Raptor (Kog1), mLST8 (Lst8), Rictor (Avo3), and mSin1 (Avo1), although several additional components are yeast or mammal specific. Furthermore yeast TORC1 also primarily controls cell growth and anabolic metabolism, including the activation of protein synthesis and inhibition of autophagy, while yeast TORC2 primarily functions to activate AGC family kinases such as YPK1, the homologue of mammalian SGK1.

As with mTORC1, yeast TORC1 also senses and responds to a diverse array of environmental stimuli, although the specific inputs and upstream signaling components differ in several respects, as one would expect given the vastly different environmental conditions that are relevant for these organisms (Fig. 3A). For example, hormone and growth factor receptor signaling are developments specific to multicellular organisms, and the mTORC1 regulator TSC is not found in *S. cerevisiae*. Instead, yeast TORC1 appears to be primarily sensitive to direct biosynthetic inputs such as carbon, nitrogen and phosphate sources.

Unlike mammals, yeast are able to synthesize all 20 amino acids, and starvation of individual amino acids like leucine or arginine does not inhibit TORC1 signaling in wild-type strains. Leucine deprivation does inhibit TORC1 in leucine-auxotrophs however (Binda et al., 2009), suggesting there may be a mechanism for signaling amino acid levels to TORC1, although this could also be due to the sensing of nitrogen sources that are perturbed in this context. Both the Rag GTPases and the GATOR1/2 complexes are present in *S. cerevisiae* in the form of Gtr1/2 and the SEACIT/SEACAT complexes, respectively (Fig. 3A, Panchaud et al., 2009), while the yeast EGO Complex is a structural homolog of Ragulator that interacts with Gtr1/2 and likely serves an analogous function (Powis et al., 2015; Zhang et al., 2012). In contrast to mammals however, amino acids do not affect the localization of yeast TORC1, which
is constitutively bound to the Gtr-Ego complex at the vacuolar membrane (Binda et al., 2009), suggesting an alternative sensing mechanism exists. Consistent with this, the mammalian amino acid sensors SLC38A9, Sestrin2, and CASTOR1 all lack clear homologs in yeast.

Figure 3: Evolutionary Conservation of the TOR Pathway

(A) The nutrient sensing pathway upstream of mammalian mTORC1 (left) and yeast TORC1 (right). (B) Phylogenetic tree depicting the presence (green box) of key mTORC1 regulators in various model organisms.

While most of the TORC1 pathway components are also well conserved in other multicellular model organisms, the direct amino acid sensors appear to have diverged (Fig. 3B). For example, although both SLC38A9 and CASTOR1 are conserved throughout many metazoan lineages, they are absent in D. melanogaster, suggesting that this organism either
does not sense arginine or does so through a distinct mechanism. Both *D. melanogaster* and *C. elegans* do have a Sestrin homolog however, and dmSestrin also binds leucine (Wolfson et al., 2016). Interestingly, both ceSestrin and dmSestrin contain subtle differences in their leucine binding pockets predicted to reduce their affinity for leucine relative to human Sestrin2, likely enabling the sensing of physiologically relevant leucine levels in these organisms (Saxton et al., 2016a).

**Section IV. Physiological Roles of mTOR**

Changes in available energy sources following fasting or feeding require alterations in whole body metabolism to maintain homeostasis. Under starvation, levels of nutrients and growth factors drop, inducing a catabolic state in which energy stores are mobilized to maintain essential functions (Fig. 4A). Alternatively, high levels of nutrients in the fed-state trigger a switch towards anabolic growth and energy storage. Consistent with its role in coordinating anabolic and catabolic metabolism at the cellular level, physiological studies in mice have revealed that mTOR signaling is essential for proper metabolic regulation at the organismal level as well. Importantly, however, constitutive mTOR activation is also associated with negative physiological outcomes, indicating that the proper modulation of mTOR signaling in response to changing environmental conditions is crucial (Fig. 4B).

**A. Glucose homeostasis**

When blood glucose levels drop, the liver activates a compensatory response involving the induction of autophagy, gluconeogenesis, and the release of alternative energy sources in the form of ketone bodies. Several lines of evidence suggest that regulation of mTORC1 signaling is crucial for the response of the liver to diet. For example, mice with liver specific deletion of TSC1, which have constitutively activated mTORC1 signaling, fail to generate ketone bodies during fasting due to sustained mTORC1-dependent suppression of PPARα, a
transcriptional activator of ketogenic genes (Sengupta et al., 2010). The importance of inhibiting mTORC1 in the liver during fasting has also been observed through the generation mice with a while-body knock-in of a constitutively active allele of RagA (RagAGTP). Although they develop normally, these mice die rapidly after birth due to an inability to maintain blood glucose levels during the perinatal fasting period (Efeyan et al., 2013). Further analysis of these mice revealed that sustained mTORC1 activity during this fasting period prevents the induction of autophagy in the liver, which is critical for supplying free amino acids for gluconeogenesis. As a result, RagAGTP mice display fatal hypoglycemia in response to fasting, consistent with a similar phenotype in autophagy deficient mice (Kuma et al., 2004).

mTORC1 signaling also plays an important role in glucose homeostasis by regulating pancreatic β-cell function. Studies using β-cell specific TSC2 knock out (β-TSC2KO) mice revealed that hyperactivation of mTORC1 has a biphasic effect on β-cell function, with young β-TSC2KO mice exhibiting increased β-cell mass, higher insulin levels, and improved glucose tolerance (Mori et al., 2009; Shigeyama et al., 2008). This effect is reversed in older β-TSC2KO mice, which more rapidly develop reduced β-cell mass, lower insulin levels, and hyperglycemia. Thus, high mTORC1 activity in the pancreas is initially beneficial for glucose tolerance, but also leads to a faster decline in β-cell function over time (Fig. 4C).

This biphasic effect of mTORC1 signaling is reminiscent of diet-induced (type 2) diabetes progression, in which pancreatic β-cells initially expand and produce more insulin to compensate for an increased glycemic load, but eventually undergo exhaustion. Indeed, obese or high fat diet (HFD) treated mice have high mTORC1 signaling in many tissues, including the pancreas, likely due to increased levels of circulating insulin, amino acids, and pro-inflammatory cytokines (Khamzina et al., 2005). Increased mTORC1 signaling in these tissues also contributes to peripheral insulin resistance due to enhanced feedback inhibition of insulin/PI3K/Akt signaling, which is prevented in mice lacking S6K1/2 (Fig. 4D, Um et al., 2004).
That mTORC1 hyperactivation from genetic or dietary manipulation results in insulin resistance has led many to speculate that mTORC1 inhibitors could improve glucose tolerance and protect against type 2 diabetes. Paradoxically, however, chronic pharmacological inhibition of mTORC1 using rapamycin has the opposite effect, causing insulin resistance and impaired glucose homeostasis (Fig. 4D, Cunningham et al., 2007). This result is explained at least in part by the fact that prolonged rapamycin treatment also inhibits mTORC2 signaling \textit{in vivo} (Lamming et al., 2012). As mTORC2 directly activates Akt downstream of insulin/PI3K signaling, it is not surprising that mTORC2 inhibition disrupts the physiological response to insulin. Consistent with this, liver specific Rictor knockout mice have severe insulin resistance and glucose intolerance (Hagiwara et al., 2012; Yuan et al., 2012), as do mice lacking Rictor in the muscle or fat (Kumar et al., 2008 & 2010).

**B. Muscle mass and function**

Although the importance of mTOR signaling in promoting muscle growth is well appreciated by basic scientists and bodybuilders alike, the mechanisms underlying this process are still poorly understood, in part due to the difficulty of genetically manipulating multinucleate myocytes \textit{in vivo}. Nonetheless, early studies of mTOR signaling in the muscle revealed that mTORC1 activation is associated with muscle hypertrophy (Bodine et al., 2001) and that both IGF-1 and leucine promote hypertrophy through the activation of mTORC1 signaling in cultured cells and in mice (Anthony et al., 2000; Rommel et al., 2001). Moreover, muscle specific Raptor knockout mice display severe muscle atrophy and reduced body weight leading to early death (Bentzinger et al., 2008). This dramatic phenotype is also observed in muscle specific mTOR knockout mice, but not Rictor deficient mice, suggesting that the critical functions of mTOR in skeletal muscle are through mTORC1 (Bentzinger et al., 2008; Risson et al., 2009).

While acute activation of mTORC1 signaling \textit{in vivo} does promote muscle hypertrophy in the short-term (Bodine et al., 2001) chronic mTORC1 activation in the muscle through loss of
TSC1 also results in severe muscle atrophy, low body mass, and early death, primarily due to a lack the inability to induce autophagy in this tissue (Fig. 4D, Castets et al, 2013). Considering that turnover of old or damaged tissue plays a critical role in muscle growth, these results suggest that alternating periods of high and low mTORC1 activity, as occurs with normal feeding and fasting cycles, is essential for maintaining optimal muscle health and function (Fig. 4F).

An accumulating body of evidence suggests that muscle contraction also activates mTORC1 in the muscle; potentially explaining at least in part how increased muscle use promotes anabolism (Baar and Esser, 1999). A recent study found that mechanical stimuli activate mTORC1 signaling by inducing the phosphorylation of Raptor (Frey et al., 2014), although how this occurs is not clear. Understanding how mTORC1 can integrate the distinct signals from insulin, amino acids, and mechanical force in the muscle will be an important goal going forward and may inform approaches for treating muscle wasting disorders such as those associated with disuse and aging.

C. Adipogenesis and lipid homeostasis

Many studies over the last two decades also reveal a role for mTOR in promoting adipocyte formation and lipid synthesis in response to feeding and insulin (Fig. 4D, reviewed in Lamming and Sabatini, 2013). mTORC1 promotes adipogenesis and enhanced lipogenesis in cell culture and in vivo, consistent with adipocyte-specific raptor knock out (Ad-RapKO) mice displaying lipodystrophy and hepatic steatosis (Lee et al., 2016). However, the role of mTORC1 in adipose is complicated by the fact that Ad-RapKO mice are also resistant to diet induced obesity due to reduced adipogenesis, suggesting mTORC1 inhibition in this tissue can have both positive and negative effects (Polak et al., 2008).

Similarly, the loss of mTORC2 activity in adipocytes primarily results in insulin resistance due to reduced Akt activity (Kumar et al., 2010), but also in less lipid synthesis in part due to

Chapter 1 - Introduction
reduced expression of ChREBPβ, a master transcription factor for lipogenic genes (Tang et al., 2016). mTORC2 has also been shown to promote lipogenesis in the liver as well, suggesting a general role for mTORC2 in lipid synthesis (Hagiwara et al., 2012; Yuan et al., 2012). Thus, both mTORC1 and mTORC2 play important roles in multiple aspects of adipocyte function and lipid metabolism.

D. Immune function

Early studies into the biological properties of rapamycin revealed a role in blocking lymphocyte proliferation, leading to its eventual clinical approval as an immunosuppressant for kidney transplants in 1999. The immunosuppressive action of rapamycin is largely attributed to its ability to block T-cell activation, a key aspect of the adaptive immune response (reviewed in Powell et al., 2012). Mechanistically, mTORC1 facilitates the switch towards anabolic metabolism that is required for T-cell activation and expansion, and lies downstream of several activating signals present in the immune microenvironment including IL-2, the co-stimulatory receptor CD28, as well as amino acids. Interestingly, mTORC1 inhibition during antigen presentation results in T-cell anergy, whereby cells fail to activate upon subsequent antigen exposure (Zheng et al., 2007). As the induction of T-cell anergy via nutrient depletion or other inhibitory signals is a mechanism utilized by tumors in immune evasion, these data suggest that promoting mTORC1 activation in immune cells may actually be beneficial in some contexts, such as cancer immunotherapy.

Recent studies have also found a role for mTORC1 in influencing T-cell maturation, as rapamycin promotes the differentiation and expansion of CD4+FoxP3+ Regulatory T-cells and CD8+ memory T-cells while suppressing CD8+ and CD4+ effector T-cell populations (Araki et al., 2009; Haxhanisto et al., 2008), consistent with the metabolic profiles of these cell types. Indeed, a recent report found that during the asymmetric division of activated CD8+ T-cells, high mTORC1 activity is high in the “effector-like” daughter cell, but low in the “memory-like”
daughter cell, due to the asymmetric partitioning of amino acid transporters (Polizzi et al., 2016; Verbist et al., 2016). Thus, the role of mTOR signaling in the immune system is clearly more complex than previously thought. Given the current clinical use of mTOR inhibitors in both immunosuppression and cancer, a more comprehensive understanding of how mTOR signaling influences overall immune responses in vivo will be a critical goal going forward.

**E. Brain function**

mTOR has also emerged as an important regulator of numerous neurological processes including neural development, circuit formation, and the neural control of feeding (reviewed in Lipton and Sahin, 2014). The deletion of either Raptor or Rictor in neurons causes reduced neuron size, and early death, suggesting that signaling by both mTORC1 and mTORC2 is important for proper brain development. Conversely, the impact of hyperactive mTORC1 signaling in the brain is best observed in human patients with Tuberous Sclerosis Complex (TSC), who exhibit a range of debilitating neurological disorders including epilepsy, autism, and the presence of benign brain tumors.

The fact that mTORC1 hyperactivation in TSC patients correlates with a high occurrence of epileptic seizures (90% of TSC patients) and autistic traits (50%) suggests that deregulated mTORC1 signaling may also be involved in epilepsy and autism more generally. Indeed, mTORC1 hyperactivation in mice through neural loss of Tsc1 or Tsc2 leads to severe epileptic seizures that are prevented by rapamycin treatment (Zeng et al., 2008), and mutations in components of the GATOR1 and KICSTOR complexes have been linked to epilepsy in humans (Basel-Vanagaite et al., 2013; Ricos et al., 2016).

The importance of mTORC1 in this tissue stems in part from its role in promoting activity-dependent mRNA translation near synapses, a critical step in neuronal circuit formation. Consistent with this, the NMDA receptor antagonist ketamine acutely activates mTORC1 signaling in mouse neurons, which coincides with an increased translation of synaptic proteins.
(Li et al., 2010). The role of mTORC1 in regulating autophagy is likely also important, as autophagy dysfunction is strongly implicated in the pathogenesis of neurodegenerative disorders including Parkinson's disease and Alzheimer's disease (AD). Inhibiting mTOR signaling has beneficial effects on mouse models of AD (Spilman et al., 2010), and it remains to be seen whether similar results will be seen in humans.

Section V. mTOR signaling in disease

A. mTOR in cancer

As discussed above, mTORC1 functions as a downstream effector for many frequently mutated oncogenic pathways including the PI3K/Akt pathway as well as the Ras/Raf/Mek/Erk (MAPK) pathway, resulting in mTORC1 hyperactivation in a high percentage of human cancers (Fig. 5A). Furthermore, the common tumor suppressors TP53 and LKB1 are negative regulators of mTORC1 upstream of TSC1 and TSC2, which are also tumor suppressors originally identified through genetic analysis of the familial cancer syndrome TSC. Several components of the nutrient sensing input to mTORC1 have also been implicated in cancer progression, including all three subunits of the GATOR1 complex, which are mutated with low frequency in glioblastoma (Bar-Peled et al., 2013), as well as RagC, which was recently found to be mutated at high frequency (~18%) in follicular lymphoma (Okosun et al., 2015). Additionally, mutations in the gene encoding folliculin (FLCN) are the causative lesion in the Birt-Hogg-Dube hereditary cancer syndrome (Nickerson et al., 2002), which manifests similarly to TSC. Finally, mutations in MTOR itself are also found in a variety of cancer subtypes, consistent with a role for mTOR in tumorigenesis (Grabiner et al., 2014; Sato et al., 2010).
Figure 4: Physiological Roles of mTOR

(A) mTORC1 controls the balance between anabolism and catabolism in response to fasting and feeding. (B) The effect of cumulative mTORC1 activity on overall health. (C) The effect of mTORC1 hyperactivation in pancreatic β-cells on glucose tolerance over time. (D) The normal functions of mTORC1 in the liver, muscle, pancreas, and adipose tissue (left), and the consequences of chronic mTORC1 inhibition (middle) or activation (right). (E) Deregulation of mTORC1 signaling in insulin resistance/diabetes, and the effect of rapamycin or a theoretical mTORC1 specific inhibitor. (F) Constitutive activation (red) or inhibition (purple) of mTORC1.
signaling in the muscle leads to atrophy, whereas optimal muscle growth and function requires alternating periods of high and low mTORC1 activity (blue).

mTORC2 signaling is also implicated in cancer largely due to its role in activating Akt, which drives pro-proliferative processes such as glucose uptake and glycolysis while also inhibiting apoptosis. Indeed, at least some PI3K/Akt driven tumors appear to rely on mTORC2 activity, as Rictor is essential in mouse models of prostate cancer driven by PTEN loss, as well as in human prostate cancer cell lines that lack PTEN (Guertin et al., 2009; Hietakangas et al., 2008).

While many mTORC1-driven processes likely contribute to tumorigenesis, the translational program initiated by the phosphorylation of 4EBP is likely the most critical, at least in mouse models of Akt-driven prostate cancer and T-cell lymphoma (Hsieh et al., 2010 & 2012). Consistent with this, a variety of Akt and Erk driven cancer cell lines are dependent on 4EBP phosphorylation, and the ratio of 4EBP to eIF4E expression correlates well with their sensitivity to mTOR inhibitors (Alain et al., 2012; She et al., 2010).

The first mTOR inhibitors approved for use in cancer were a class of rapamycin derivatives known as “rapalogs”. The rapalog temsirolimus (Pfizer) was first approved for treatment of advanced renal cell carcinoma in 2007, followed by everolimus (Novartis) in 2009. Although a small number of “extraordinary responders” have been reported, these rapalogs have been less successful in the clinic than anticipated from pre-clinical cancer models.

Several explanations for this lack of efficacy have been suggested. The first came with the realization that, as allosteric inhibitors, the rapalogs block the phosphorylation of some but not all mTORC1 substrates (Fig. 5B, Choo et al., 2008; Feldman et al., 2009; Kang et al., 2013; Thoreen et al., 2009). In particular, the phosphorylation of 4EBP is largely insensitive to rapamycin. Second, inhibiting mTORC1 releases the negative feedback on insulin/PI3K/Akt signaling, and therefore may paradoxically promote cell survival and prevent apoptosis in some...
contexts. Indeed, increased Akt signaling has been observed in biopsies of cancer patients following everolimus treatment, and may help explain why rapalogs have largely cytostatic, but not cytotoxic, effects on tumors (Tabernero et al., 2008). Finally, mTORC1 inhibition also induces autophagy, which can help maintain cancer cell survival poorly vascularized, nutrient poor microenvironments such as in pancreatic tumors. Indeed, mTORC1 inhibition also promotes macropinocytosis, whereby extracellular proteins are internalized and degraded to provide amino acids for nutrient-starved tumors (Palm et al., 2015). These data suggest that combining rapalogs with autophagy inhibitors may improve efficacy, consistent with a recent phase 1 clinical trial using temsirolomus and the autophagy inhibitor hydrochloroquine in melanoma patients, which showed an improvement over temsirolomus alone (Rangwala et al., 2014).

In order to address some of the drawbacks of the rapalogs, “second generation,” ATP-competitive catalytic inhibitors against mTOR have also been developed and are now in clinical trials. Unlike rapamycin, these compounds directly inhibit the catalytic activity of mTOR and therefore fully suppress both mTORC1 and mTORC2 (Fig. 5B), making them more effective than rapalogs in a variety of preclinical cancer models. Although these second-generation mTOR inhibitors initially suppress Akt signaling due to inhibition of mTORC2, the release of negative feedback on Insulin/PI3K signaling eventually overcomes this blockade and Akt is reactivated following long-term treatment (Rodrik-Outmezguine et al., 2011). One possible solution to this problem may be the utilization of dual PI3K/mTOR inhibitors, which inhibit the closely related catalytic domains of both PI3K and mTOR, thereby more fully blocking phosphorylation of both Akt and 4EBP (Fig. 5B). These inhibitors have also shown some promise in preclinical and early clinical trial data, but have raised concerns over dose-limiting toxicities.

The cases where rapalogs have had the most success to date have generally involved mutations in the mTOR pathway itself, such as in TSC1 or MTOR (Iyer et al., 2012; Wagle et al.,
Even in these cases, however, an exquisite initial response has been followed by additional mutations in the kinase or FRB domains of MTOR, leading to acquired resistance (Wagle et al., 2014b). One creative way to overcome these resistance mutations may be with the recently described “third-generation” mTOR inhibitor called “RapaLink,” in which the ATP-competitive mTOR inhibitor is chemically linked to rapamycin, enabling inhibition of mTOR mutants that are resistant to either MLN0128 or rapamycin alone (Rodrik-Outmezguine et al., 2016).

B. mTOR in aging

mTOR signaling is strongly implicated in the aging process of diverse organisms including yeast, worms, flies, and mammals. This was first observed through studies in the nematode C. elegans, which found that reduced expression of the homologs of mTOR (ceTOR, formerly let-363) or Raptor (daf-15) extend life span (Vellai et al., 2003; Jia et al., 2004). Subsequent genetic studies found that reduced TOR signaling also promotes longevity in formerly let-363) or Raptor (daf-15) extend life span (Vellai et al., 2003; Jia et al., 2004). Subsequent genetic studies found that reduced TOR signaling also promotes longevity in Drosophila (Kapahi et al., 2004), budding yeast, (Kaeberlin et al., 2005) as well as mice (Lamming et al., 2012; Wu et al., 2013). Consistent with this, the mTOR inhibitor rapamycin is currently the only pharmacological treatment proven to extend life span in all of these model organisms (Bjedov et al., 2010; Harrison et al., 2009; Powers et al., 2006; Robida-Stubbs et al., 2012).

The only other intervention shown to extend lifespan in such a wide range of organisms is caloric restriction (CR), defined as a reduction in nutrient intake without malnutrition. Given the critical role of mTORC1 in sensing nutrients and insulin, this has led many to speculate that the beneficial effects of CR on life span are also due to reduced mTORC1 signaling. Indeed, CR-like regimens do not further extend life span in yeast, worms, or flies with reduced mTOR

Chapter 1 - Introduction
Figure 5: mTOR in Cancer and Aging

(A) The common tumor suppressors and oncogenes upstream of mTORC1 leading to increased mTORC1 signaling in a wide variety of cancers. (B) The varying effects of rapalogs, catalytic mTOR inhibitors, a combination of an mTOR inhibitor and autophagy inhibitor on cancer proliferation and survival. (C) The role of mTORC1 signaling in aging.

signaling, suggesting an overlapping mechanism (Kaeberlin et al., 2005; Hansen et al., 2007; Kapahi et al., 2004).

While there is now a general consensus that mTOR signaling plays a key role in mammalian aging, the mechanism through which this occurs is still unclear. Several lines of
evidence suggest that the general reduction in mRNA translation during mTORC1 inhibition slows aging by reducing the accumulation of proteotoxic and oxidative stress, consistent with the observation that loss of the mTORC1 substrate S6K1 also extends life span in mammals (Selman et al., 2009). A related possibility is that inhibition of mTORC1 slows aging by increasing autophagy, which helps clear damaged proteins and organelles such as mitochondria, the accumulation of which are also associated with aging and aging-related diseases. Finally, another model suggests that the attenuation of adult stem cells in various tissues plays a central role in organismal aging, and mTOR inhibition boosts the self-renewal capacity of both hematopoietic and intestinal stem cells in mice (Chen et al., 2009; Yilmaz et al., 2012). Ultimately, the importance of mTORC1 signaling in aging likely reflects its unique capacity to regulate such a wide variety of key cellular functions (Fig 5C).

The observation that mTOR inhibition extends life span and delays the onset of age-associated diseases in mammals has led many to speculate that mTOR inhibitors could be used to enhance longevity in humans. The major drawback of prolonged rapamycin treatment in humans however is the potential for side effects such as immunosuppression and glucose intolerance. There is reason for optimism however, as a trial in healthy elderly humans using everolimus showed safety and even improved immune function (Mannick et al., 2014), and alternative dosing regimens have been proposed that can promote longevity with reduced side effects (Arriola Apelo et al., 2016). Given that many of the negative metabolic side effects associated with mTOR inhibitors are due to inhibition of mTORC2, while the anti-aging effects are due to inhibition of mTORC1, the development of mTORC1 specific inhibitors would be particularly beneficial.

Section VI. Contributions of this thesis
A. Structural basis for leucine sensing by the Sestrin2-mTORC1 pathway (Chapter 2)
In order to understand the mechanism through which the leucine sensor Sestrin2 detects leucine and signals its presence to downstream components of the mTORC1 pathway, we purified recombinant, human Sestrin2 from *E. coli*, crystallized it in the presence of leucine, and solved the structure to 2.7 Å resolution (Saxton et al., 2016a). This structure revealed the overall architecture of Sestrin2, which is unique among human proteins with known structures and unrelated to any known amino-acid binding proteins. This structure also revealed the molecular basis for leucine recognition by Sestrin2, which contains a single pocket with high specificity for leucine and a “lid-latch” mechanism that facilitates ligand binding and likely mediates the allosteric, leucine-dependent regulation of the Sestrin2-GATOR2 interaction. In addition we identified a highly conserved GATOR2-binding site on the surface of Sestrin2 in close proximity to the leucine-binding pocket, shedding further light on how leucine triggers the dissociation of Sestrin2 from GATOR2 in order to enable the activation of mTORC1. Together, these data showed the structural basis for amino acid sensing by the mTOR pathway, revealing for the first time how human cells sense the key nutrient leucine and enabling the development of novel mTORC1-specific therapeutics.

B. Systematic analysis of reported “apo-Sestrin2” structures (Chapter 3)

A key prediction of our proposed structural model of leucine sensing by Sestrin2 is that leucine binding induces a conformational change in Sestrin2 that modulates the conserved GATOR2-binding site. However, another group subsequently published a reported “apo” structure of human Sestrin2, which showed the protein in a nearly identical conformation as our leucine-bound structure. Based on this observation, it was argued that Sestrin2 is unlikely to be a true cellular leucine-sensor for the mTORC1 pathway. In order to understand the basis for this discrepancy, we computationally re-analyzed the published “apo”-Sestrin2 structure, which we showed to in fact contain electron density likely corresponding to leucine or a chemically similar ligand bound to the leucine-binding pocket (Saxton et al., 2016b). In addition, we re-purified and
crystallized human Sestrin2 in the absence of exogenous leucine, and upon solving this structure revealed that it too contained a bound leucine molecule, indicating that the probability of inadvertently crystallizing the leucine-bound form of Sestrin2 while trying to obtain the apo-form is high. Together, these results demonstrate that the structure of apo-Sestrin2 is still elusive and remains an important challenge for the field.

C. Mechanism of arginine sensing by CASTOR1 upstream of mTORC1 (Chapter 4)

In addition to leucine, the proteogenic amino acid arginine is also known to exert a range of important physiological effects in humans in large part through its ability to activate mTORC1. Previous work demonstrated that arginine signals to mTORC1 both through the lysosomal amino acid transporter SLC38A9, and through the GATOR2-interacting protein CASTOR1 located in the cytosol. However, the molecular mechanism of arginine sensing had been mysterious in both cases. In order to understand how CASTOR1 senses cytosolic arginine and signals its availability to mTORC1, we solved the crystal structure of arginine-bound CASTOR1 to 1.8 Å resolution (Saxton et al., 2016c). This structure showed how CASTOR1 binds arginine through a pocket at the interface of two ACT (Aspartate Kinase, Chorismate mutase, TyrA) domains. Our subsequent biochemical analysis revealed that arginine acts as "molecular glue," stabilizing the association of the two ACT domains, thereby allosterically altering the adjacent GATOR2 binding site. In addition, our structure revealed that, despite very low sequence homology, CASTOR1 is structurally very similar to bacterial Aspartate kinase, suggesting an evolutionary model of how the mTORC1 pathway acquired arginine-sensing capabilities.
REFERENCES


Chapter 1 - Introduction


Chapter 1 - Introduction


Chapter 1 - Introduction


Chapter 2

Structural basis for leucine sensing by the Sestrin2-mTORC1 pathway

Parts of this chapter were first published as:


Abstract

Eukaryotic cells coordinate growth with the availability of nutrients through mTOR complex 1 (mTORC1), a master growth regulator. Leucine is of particular importance and activates mTORC1 via the Rag GTPases and their regulators GATOR1 and GATOR2. Sestrin2 interacts with GATOR2 and is a leucine sensor. We present the 2.7-Å crystal structure of Sestrin2 in complex with leucine. Leucine binds through a single pocket that coordinates its charged functional groups and confers specificity for the hydrophobic side chain. A loop encloses leucine and forms a lid-latch mechanism required for binding. A structure-guided mutation in Sestrin2 that decreases its affinity for leucine leads to a concomitant increase in the leucine concentration required for mTORC1 activation in cells. These results provide a structural mechanism of amino acid sensing by the mTORC1 pathway.
Introduction

The mechanistic target of rapamycin complex 1 (mTORC1) protein kinase is a major growth-regulator that coordinates cell anabolism and catabolism with the availability of key nutrients like amino acids (Dibble and Manning, 2013; Jewell et al., 2013; Zoncu et al., 2011b). Among the amino acids, leucine is of particular interest due to its ability to promote important physiological phenomena, including muscle growth and satiety (Greiwe et al., 2001; Nair et al., 1992; Potier et al., 2009), in large part through activation of mTORC1 (Fox et al., 1998; Lynch et al., 2000). However, the biochemical mechanism of leucine sensing by the mTORC1 pathway has remained elusive.

While growth factors, energy, and other inputs signal to mTORC1 primarily through the Tuberous Sclerosis Complex (TSC)-Rheb axis (Buerger et al., 2006; Efeyan and Sabatini, 2013; Saito et al., 2005), amino acids act by regulating the nucleotide state of the heterodimeric Rag guanosine triphosphatases (GTPases) and promoting the localization of mTORC1 to lysosomes, its site of activation (Bar-Peled et al., 2012; Sancak et al., 2010; Sancak et al., 2008). Lysosomal amino acids including arginine are thought to signal to the Rags through a lysosomal membrane associated complex consisting of the v-ATPase (Zoncu et al., 2011a), Ragulator complex (Sancak et al., 2010), and the putative arginine sensor SLC38A9 (Rebsamen et al., 2015; Wang et al., 2015). Cytosolic leucine, however, signals to the Rags through a distinct pathway consisting of a pentameric protein complex of unknown function called GATOR2, and GATOR1, the GTPase-Activating protein (GAP) for RagA and RagB (Bar-Peled et al., 2013; Wolfson et al., 2016).

Proteomic studies have identified the Sestrins as GATOR2-interacting proteins that inhibit mTORC1 only in the absence of amino acids (Chantranupong et al., 2014; Parmigiani et al., 2014). Subsequent in vitro studies demonstrated that the Sestrin2-GATOR2 interaction is sensitive specifically to leucine, which binds Sestrin2 with a dissociation constant ($K_d$) of approximately 20 $\mu$M (Wolfson et al., 2016). Human Embryonic Kidney (HEK)-293T cells
expressing a Sestrin2 mutant that cannot bind leucine fail to activate mTORC1 in response to leucine, suggesting that Sestrin2 is the primary leucine sensor for the mTORC1 pathway in these cells (Wolfson et al., 2016). However, Sestrin2 shares no sequence similarity with known amino acid binding domains, raising the question of how this protein can detect leucine and signal its presence to mTORC1.

Here, we present the structure of human Sestrin2 in complex with leucine, revealing in atomic detail the mechanism of leucine sensing by the mTORC1 pathway.

**Structure of leucine-bound Sestrin2**

To understand how Sestrin2 detects leucine, we expressed and purified full-length human Sestrin2 from *E. coli* and verified binding to leucine in vitro by differential scanning fluorimetry (DSF) (Fig. S1) (Niesen et al., 2007). Although we were unable to obtain crystals of Sestrin2 alone, incubation of the protein with leucine allowed formation of crystals containing leucine-bound Sestrin2 that diffracted to 2.7-Å resolution. We solved the structure using single-wavelength anomalous dispersion (SAD) with selenomethionine-derivatized protein and refined the model against the native data to a final Rwork/Rfree of 19.6%/22.3% (Table S1). Sestrin2 crystallized in a cubic space group containing five copies per asymmetric unit.

Sestrin2 is a 55 kDa, monomeric, all α-helical, globular protein that contains distinct N-terminal [NTD, residues 66-220] and C-terminal [CTD, residues 339-480] domains connected by a partially disordered, partially helical linker region [Linker, residues 221-338] (Fig. 1A). The N-terminal 65 residues of the protein appear disordered and were not observed in our structure. Electron density map analysis revealed the presence of a single leucine molecule bound to Sestrin2 in the C-terminal domain (Fig 2A).

The N- and C-terminal domains of Sestrin2 appear to be structurally similar and superpose well, with a root mean square deviation (rmsd) of ~3.0 Å over 55 aligned Ca positions, despite a low sequence identity of 10.9% (Fig. 1B). Furthermore, the two domains
make extensive contacts with each other, primarily through the two core hydrophobic helices N9 and C7, burying 1,872 Å² of surface area (Fig. 1A).

A small region in the N terminus of Sestrin2 contains weak sequence similarity to the bacterial alkylhydroperoxidase AhpD (Budanov et al., 2004). Analysis of our structure with the NCBI Vector Alignment Search Tool (VAST) (Gibrat et al., 1996) showed that Sestrin2 shares a common fold with the carboxymucolactone decarboxylase (CMD) protein family, consisting of bacterial γ-CMD as well as AhpD (pfam: PF02627). Despite low sequence similarity, Sestrin2 strongly resembles an AhpD homodimer, with each half of Sestrin2 matching a single AhpD molecule (Fig. 1C, S2A). The N- and C-terminal domains both superpose well with *R. eutropha* AhpD, with rmsd’s of ~2.0 Å over 129 and 101 Ca’s, respectively. Thus, Sestrin2 structurally resembles an intra-molecular homo-dimer of two CMD-like domains, despite extensive divergence in the primary sequence.

To test the importance of the intra-molecular contacts between the two domains of Sestrin2, we expressed the FLAG-tagged N- and C-terminal halves either alone or together as separate polypeptides and performed co-immunoprecipitation analysis. Although neither domain alone bound GATOR2, the separated halves, when expressed together, bound strongly to both each other and to GATOR2 (Fig. 1D). Similarly, although neither half of Sestrin2 alone bound to leucine, the co-expressed halves did bind leucine (Fig. 1E). Therefore, the N- and C-terminal domains of Sestrin2 interact stably with each other and are both required for the interactions with GATOR2 as well as leucine.

In addition to its role as a leucine sensor and GATOR2 inhibitor, Sestrin2 has been reported to have peroxiredoxin-reductase activity, based in large part on its weak sequence similarity to bacterial AhpD, which does have this activity (Budanov et al., 2004). However, other groups have failed to reproduce this finding (Woo et al., 2009). The active site of AhpD contains two cysteines, both of which are required for its catalytic activity (Koshkin et al., 2003; Woo et al., 2009). Superposing our Sestrin2 structure with AhpD confirms previous reports.
### Table S1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sestrin2 + Leu native</th>
<th>Sestrin2 + Leu SeMet derivative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organism</td>
<td>Homo sapiens</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>PDB ID</td>
<td>5DJ4</td>
<td></td>
</tr>
</tbody>
</table>

#### Data collection
- **Space group**: l23
- **a, b, c (Å)**: 293.03, 293.03, 293.03
- **α, β, γ (°)**: 90.0, 90.0, 90.0
- **Wavelength (Å)**: 0.9792
- **Resolution range (Å)**: 92.67 – 2.70 (2.75 – 2.70)
- **Total reflections**: 4,575,519
- **Unique reflections**: 114,283
- **Completeness (%)**: 100 (100)
- **Redundancy**: 40.0 (39.4)
- **Anomalous completeness (%)**: 99.9
- **Rsym (%)**: 15.3 (>100)
- **R_{p.i.m.} (%)**: 3.8 (59.7)
- **I/σ**: 24.5 (1.4)
- **CC_{1/2} (%)**: 99.8 (65.6)

#### Refinement
- **Resolution range (Å)**: 92.66 – 2.70 (3.45-2.70)
- **R_{work} (%)**: 19.6
- **R_{free} (%)**: 22.3
- **Number of Reflections**: 114,277
- **Total R_{free} reflections**: 2,001
- **Number of non-hydrogen atoms**: 14,945
- **Protein atoms**: 14,830
- **Water atoms**: 115
- **R.m.s. deviations**: 0.011
- **Average B factors (Å^2)**:
  - **Protein**: 57.6
  - **Leu**: 49.6
  - **Water**: 45.6
- **Wilson B-factor (Å)**: 57.3
- **Ramachandran (%):**
  - **Favored (%)**: 97.0
  - **Allowed (%)**: 2.5
  - **Outlier (%)**: 0.5
- **Molprobity percentile**: 92<sup>rd</sup>
Figure 1: Structure of leucine-bound Sestrin2

(A) Two views of human Sestrin2 are shown as ribbon diagrams, with the N-terminal (NTD), Linker, and C-terminal (CTD) domains colored in blue, gray, and teal respectively. The bound leucine molecule is shown in orange. Disordered regions not present in the crystal structure (1-65, 242-255, 272-280, 296-309) are shown as dashed lines. (B) Structural superposition of Sestrin2 NTD (blue, residues 66-220) and CTD (teal, residues 339-480). (C) Structural superposition on Sestrin2 NTD (blue) and CTD (teal) with a *R. eutropha* AhpD dimer (pink, PDB ID: 2PRR). (D) Immunoprecipitation of N- and C-terminal fragments of Sestrin2. HEK-293T cells transiently transfected with FLAG-metap2, FLAG-Sestrin2 full length (FL), FLAG-Sestrin2-NTD (N-terminal domain, 1-220), FLAG-Sestrin2-CTD+L (C-terminal domain plus Linker, 220-480) or both Flag-Sestrin2-NTD and HA-Sestrin2-CTD+L were starved for amino acids for 50 minutes. FLAG-immunoprecipitates were prepared from cell lysates. Immunoprecipitates and lysates from one representative experiment were analyzed by immunoblotting for indicated proteins. WDR24 and Mios were used as representative GATOR2 components. (E) [3H]Leucine binding assay using N- and C-terminal fragments of Sestrin2. FLAG-immunoprecipitates prepared from HEK-293T cells transiently expressing indicated proteins were used as described in the methods. Unlabeled leucine was used as a competitor where indicated. Values are Mean ± SD for 3 technical replicates from one representative experiment. Two-tailed t tests were used for comparison between two groups.
Figure S1

(A) Differential Scanning Fluorimetry with bacterially purified Sestrin2. Purified Sestrin2 was incubated with Sypro orange dye with or without leucine or arginine. Upon heating the sample the change in fluorescence was captured and used to calculate melting temperatures (Tm) under the indicated conditions. Values are Mean ± SD from 3 replicates. The protein was analyzed by SDS-PAGE followed by Coomassie blue staining.

(Budanov et al., 2004; Woo et al., 2009) that only one of these active site residues is present in the N-terminal half of Sestrin2 (Cys125) whereas both are absent from the C-terminal half (Fig. S2B), suggesting that Sestrin2 either does not reduce peroxiredoxins, or does so through an entirely different mechanism than does AhpD.

Sestrin2 is also reported to inhibit the mTORC1 pathway by directly acting as a guanine-nucleotide dissociation inhibitor (GDI) for RagA and RagB through a motif consisting of Arg419, Lys422, and Lys426 (Peng et al., 2014). However, in our structure two of these three residues (Lys422 and Lys426) are buried (Fig. S2C), and Sestrin2 shows no structural similarity to known GDI proteins.

Chapter 2 - Structure of leucine-bound Sestrin2
Figure S2

(A) Side-by-side comparison of the human Sestrin2 structure from this study and *R. eutropha* AhpD (PDB ID: 2PRR). (B) Close up view of the active site cysteine residues in *R. eutropha* AhpD and the corresponding regions in the N- and C- terminal domains of human Sestrin2. Helix numbers and relevant residues are labeled. (C) Solvent exposed surface view of Sestrin2, with the reported “RKK” motif highlighted. Arg419 (yellow) is exposed while Lys422 and Lys426 (red) are buried.
Recognition of leucine by Sestrin2

Sestrin2 binds leucine through a single pocket formed at the intersection of helices C2, C3, and C7 in the C-terminal domain. Charged residues Glu451 and Arg390 form two sides of the pocket and anchor leucine in place through salt bridges with the free amine and carboxyl groups, respectively (Fig. 2, A and B). In addition, helix L in the Linker region packs against side of the pocket via residue Leu261 (Fig. S3A), sealing the side of the pocket. This is consistent with mutagenesis studies that identified Glu451 and Leu261 as critical for leucine binding (Wolfson et al., 2016). Meanwhile, the isopropyl side chain of the bound leucine points down towards the hydrophobic base of the pocket, forming extensive van der Waals contacts with residues Leu389, Trp444, and Phe447 (Fig. 2, A and B). The depth and overall hydrophobicity of this pocket floor make it well suited to accommodate leucine (Fig. 2B).

To test the importance of these protein-ligand interactions, we generated a series of Sestrin2 leucine-pocket mutants. Disrupting the electrostatic coordination of the free amine by switching a single oxygen atom in Glu451 to nitrogen (E451Q) resulted in a complete loss of leucine binding, as did eliminating the interaction between Arg390 and the free carboxyl (R390A, Fig. 2C). In addition, although leucine readily triggered dissociation of the wild-type Sestrin2-GATOR2 complex, both the E451Q and R390A mutants remained constitutively bound to GATOR2 even in the presence of leucine (Fig. 2D). The hydrophobic integrity of the pocket floor is also critical, as insertion of a single charged residue (W444E) was sufficient to abolish any detectable interaction with leucine (Fig. 2, C and D). Consistent with an essential role for these residues in leucine sensing, a multiple sequence alignment of Sestrin homologs showed that both Glu451 and Arg390 are strictly conserved in Sestrin proteins across phylogenetically diverse organisms, as is the hydrophobic nature of the pocket floor (Fig. 2E).

These results provide a molecular explanation for how the Sestrin2-mTORC1 pathway specifically detects leucine and not other amino acids. Although Glu451 and Arg390 likely
interact with any amino acid containing free amine and carboxyl groups, the hydrophobic base of the pocket excludes all charged and polar amino acids. Furthermore, large hydrophobic residues can be seen in Figure 2.

**Figure 2**

(A) Close-up view of the leucine binding pocket in Sestrin2, focusing on the bound leucine (shown in orange) together with its 2Fo-Fc electron density map calculated and contoured at 1.5σ from an omit map lacking leucine and all pocket residues. Predicted hydrogen bonds or salt-bridges are shown as black dashed lines. Helix numbers are labeled as in 1A. (B) Surface representation of leucine-bound Sestrin2, focusing on the leucine binding pocket. The bound leucine is represented as a stick model (orange). Residues 373-387 are omitted to allow visibility of the pocket. Residue Glu451, which contacts the amine of leucine is shown in red, and Arg390 which contacts the carboxyl of leucine is shown in blue. The domains of Sestrin2 are colored as in 1A. (C) Binding of the E451Q, R390A and W444E mutants of Sestrin2 to leucine. HA-immunoprecipitates prepared from HEK-293T cells transiently expressing indicated HA-tagged proteins were used in binding assays with [3H]Leucine. Binding was analyzed as in 1E. (D) Effect of leucine on the interactions of Sestrin2 E451Q, R390A or W444E with GATOR2. FLAG-immunoprecipitates were prepared from cells stably expressing FLAG-WDR24 and transiently expressing the indicated HA-tagged Sestrin2 constructs. The immunoprecipitates were treated with the indicated concentrations of leucine and analyzed by immunoblotting for...
the indicated proteins. (E) Multiple Sequence Alignment of Sestrin2 homologs from various organisms. Positions of residues contacting leucine are indicated with orange dots. Positions are colored white to blue according to increasing sequence identity.

residues such as phenylalanine will clash with Trp444 in the pocket floor, whereas small aliphatic amino acids such as alanine or valine will fail to make favorable van der Waals contacts. Thus, only leucine and the structurally similar amino acids isoleucine and methionine interact appreciably. This is consistent with the finding that only leucine and to a much lesser extent isoleucine and methionine disrupt the Sestrin2-GATOR2 interaction in vitro (Wolfson et al., 2016).

Interestingly, the corresponding region in the NTD of Sestrin2 is filled by protein side chains and cannot accommodate leucine (Fig. S3B). However, the position of key residues including Trp444 and Glu451 are conserved in the NTD "pocket" (Trp189 and Glu193), and a leucine side chain contributed by Leu107 occupies the same position as the bound leucine in the CTD (Fig S3B).

A lid-latch mechanism is required for leucine binding by Sestrin2

In addition to contacting the charged sides and hydrophobic base of the pocket, a "lid" formed by a loop connecting helices C2 and C3 encloses the top of the leucine such that it is completely buried within the structure (Fig. 2A). Three highly conserved threonine residues (Thr374, Thr377 and Thr386) are positioned directly above the leucine and help coordinate the free amine and carboxyl groups, locking the ligand in place (Fig. 2E, 3A). The side chain hydroxyl groups of Thr374 and Thr386 make hydrogen bond contacts with the carboxyl group of leucine, whereas the free amine donates a hydrogen bond to the backbone carbonyl of Thr377 (Fig. 2A, 3A).
To analyze the importance of these lid interactions for leucine detection by Sestrin2, we generated mutants predicted to eliminate the critical contacts between the lid and leucine. Mutation of either Thr374 or Thr386 (T374A or T386A) abolished the interaction with leucine and resulted in a constitutive interaction with GATOR2, (Fig. 3, C and D), demonstrating a crucial role for the lid in leucine binding.

Although our in vitro binding data demonstrated a requirement for both the N- and C-terminal halves of Sestrin2 (Fig. 1D), the structural model shows that the bound leucine only
Figure S4

Chapter 2 - Structure of leucine-bound Sestrin2
Figure S4

(A) Expanded Multiple Sequence Alignment of Sestrin2 homologues from various organisms. Positions are colored white to blue according to increasing sequence identity. Residues involved in GATOR2 binding (red) and leucine binding (orange) are indicated.

makes direct contacts with residues in the C-terminal domain (Fig. 2A). Further structural analysis however revealed that the lid residue Tyr375 forms a tight hydrogen bond with the N-terminal residue His86, located in a loop between helices N2 and N3 adjacent to the leucine-binding pocket (Fig. 3B). This interaction appears to form a “latch,” which locks the lid in place over the bound leucine. Indeed, this inter-domain contact appears to be critical for the Sestrin2-leucine interaction, as specifically eliminating this hydrogen-bonded latch with either a Y375F or H86A mutation abolished leucine binding (Fig. 3, C and D). Both Tyr375 and His86 are also
highly conserved in Sestrin proteins across organisms (Fig. 2E, S4). The requirement for His86 to maintain the latch interaction likely explains why the N-terminal domain of the protein is also essential for the interaction with leucine.

Figure 3

(A) Top-down view of the leucine bound pocket, focusing on the "lid" residues Thr374, Thr377, and Thr386 which form hydrogen bonds with the amine and carboxyl groups of leucine (indicated by black dashed lines). Leucine is represented as a stick model (orange). Helix numbers are labeled as in 1A. (B) Orthogonal view of the leucine binding pocket, focusing on the "latch" formed by the predicted hydrogen bond between Tyr375 and His86, which locks the lid in place over the bound leucine (orange). Helix numbers are labeled as in 1A. (C) Binding of Sestrin2 T374A, T386A, Y375F and H86A mutants of Sestrin2 to leucine. Binding assays were performed and immunoprecipitates analyzed as in Figure 2C. (D) Effect of leucine on the interactions of Sestrin2 T386A, Y375F or H86A with GATOR2 in vitro. FLAG immunoprecipitates were prepared from cells stably expressing FLAG-WDR24 and transiently expressing the indicated HA-tagged Sestrin2 constructs and analyzed as in 2D.
Altering the leucine sensitivity of the mTORC1 pathway in cells

One prediction for a bona fide cellular leucine sensor is that its affinity for leucine should in part determine the sensitivity of the mTORC1 pathway to leucine. We tested this hypothesis directly by generating a mutant of Sestrin2 with lower affinity for leucine. We predicted that deepening the hydrophobic base of the pocket by mutating Trp444 to Leu (W444L) would reduce the van der Waals contacts with the bound leucine side chain, thereby weakening but not eliminating the interaction (Fig. 4A). Indeed, the Sestrin2 W444L mutant bound one-sixth to one-eighth the amount of leucine that bound to wild-type Sestrin2 (Fig. 4B). Furthermore, addition of \(~10\) to 15 fold more leucine was required to fully dissociate the W444L mutant from GATOR2 compared to wild type Sestrin2 (Fig. 4C).

To test the effect of this mutation on mTORC1 signaling in cells, we used a HEK-293T cell line in which Sestrin1, 2 and 3 were knocked-out the CRISPR-Cas9 system (Sestrin TKO cells). The mTORC1 signaling in these cells is fully resistant to leucine deprivation, and re-introduction of wild-type Sestrin2 restores normal signaling with half-maximal mTORC1 activity occurring upon addition of \(~20\) to 50 \(\mu\)M leucine (20, Fig. 4D). Expression of Sestrin2 W444L in the Sestrin TKO lines however shifted the dose response of mTORC1 to leucine, such that addition of \(~250\) to 500 \(\mu\)M was required to achieve half-maximal activation of the pathway (Fig. 4D). Thus, the affinity of Sestrin2 for leucine is a major determinant of the sensitivity of the mTORC1 pathway to leucine in human cells.

Although the overall hydrophobicity of the pocket floor is well conserved, the specific residues present at the W444 and F447 positions vary across organisms, and some organisms, including Drosophila, carry the corresponding W444L mutation (Fig. 2E). These differences may alter the shape and depth of the leucine pocket, leading to different affinities or specificities for leucine in different organisms. This may represent an evolutionary adaptation to enable efficient sensing of leucine concentrations that are physiologically relevant in these organisms.
Figure 4: Altering the leucine sensitivity of the mTORC1 pathway in cells

(A) Close-up view of Sestrin2-bound leucine (orange) and the pocket floor residues Phe447 and W444, with the W444L mutant (red) overlaid onto the wild-type protein (teal). Both residues are represented as stick models. Numbers indicate distance from leucine to residue 444 in Sestrin2 WT and Sestrin2 W444L. (B) Leucine binding by Sestrin2 W444L. Binding assays were performed and immunoprecipitates analyzed as in Figure 2C. (C) Higher concentrations of leucine are required to dissociate Sestrin2 W444L from GATOR2 compared to Sestrin2 WT. FLAG immunoprecipitates were prepared from cells stably expressing FLAG-WDR24 and transiently expressing the indicated HA-tagged Sestrin2 constructs and analyzed as in figure 2D. (D) Sensitivity of the mTORC1 pathway to leucine in Sestrin1-3 triple-knock-out (TKO) cells expressing Sestrin2 wild type or W444L. HEK-293T cells generated with the CRISPR/Cas9 system expressing the indicated proteins via lentiviral transduction. Cells were starved of leucine for 50 minutes then re-stimulated with the indicated amount of leucine for 10 minutes. Cell lysates from one representative experiment were prepared and analyzed via immunoblotting.
Characterizing the GATOR2 binding site on Sestrin2

To better understand how leucine binding triggers dissociation of Sestrin2 from GATOR2, we sought to structurally characterize the GATOR2 binding interface of Sestrin2. Mutagenesis studies identified residue S190 in the NTD as required for GATOR2 binding (Wolfson et al., 2016), however this site is distal to the leucine-binding pocket. Mapping electrostatic potential onto the solvent-exposed surface of Sestrin2 revealed a region in close proximity to the leucine-binding site containing the highly conserved charged residues Asp406 and Asp407 (Fig. 5A, S5, A and B). Mutating these residues to alanine (DD406-7AA) completely eliminated GATOR2 binding without affecting leucine binding (Fig. 5B, S5C), suggesting that this region is required for the Sestrin2-GATOR2 interaction. Thus, Sestrin2 may make multiple contacts with GATOR2 through both the NTD and CTD (Fig. 5C, 1B), consistent with both halves of Sestrin2 being required for GATOR2 binding (Fig. 1D).

Conclusions

Our results provide a structural model of leucine sensing by the Sestrin2-mTORC1 pathway and shed light into the mechanism through which mTORC1 couples cell growth to leucine availability. The structure shows that Sestrin2 contains an evolutionarily unique leucine-binding pocket consisting of a hydrophobic floor that determines specificity for the side-chain of leucine, with adjacent glutamate and arginine residues that coordinate the free amine and carboxyl groups, respectively. An additional “lid-latch” mechanism helps lock the ligand in place and is required for binding. Our structure also reveals a highly conserved GATOR2 binding site in close proximity to the leucine pocket, suggesting possible mechanisms for how leucine binding can cause dissociation of Sestrin2 from GATOR2. The key residues for the GATOR2 interaction, Asp406 and Asp407, are located on a loop separated from the “lid” of the pocket by the 15-residue helix C3 (Fig. 5A). It is therefore conceivable that a conformational change in the
lid, corresponding to leucine binding or release, could transmit a conformational change to the GATOR2 binding site via movement of helix C3 (Fig 5D). Alternatively, a segment of the

Figure 5

Figure 5: Identification of GATOR2 binding site and model of leucine sensing by Sestrin2

(A) View highlighting the conserved surface aspartates Asp406/407 and their position relative to the bound leucine (orange). (B) Co-immunoprecipitation of GATOR2 with Sestrin2 wild type or DD406-7AA. FLAG immunoprecipitates were prepared from cells stably expressing FLAG-WDR24 and transiently expressing the indicated HA-tagged Sestrin2 constructs and analyzed as in 2D. (C) Surface view of Sestrin2 highlighting the GATOR2 binding sites (red) and their position relative to the leucine-binding pocket (orange). Domains are colored as in 1A. (D) Model of leucine sensing by Sestrin2. Binding of leucine (orange) causes closing of the lid-latch, resulting in a conformational change altering the position of the GATOR2 binding site in the

Chapter 2 - Structure of leucine-bound Sestrin2
CTD. This leads to dissociation of Sestrin2 from GATOR2, enabling GATOR2 to activate the mTORC1 pathway.

Figure S5

(A) Two views of Sestrin2 with the solvent exposed surface colored by electrostatic potential. The location of the leucine-binding pocket (orange circle) and a nearby negative charge cluster (black circle) are indicated. (B) Multiple sequence alignment of showing the conserved residues Asp406 and Asp407. (C) Sestrin2 DD407-7AA still binds leucine. Binding assays were performed and immunoprecipitates analyzed as in Figure 2C. The indicated proteins purified from HEK-293T cells and used in the [3H]-leucine binding assay were analyzed by SDS-PAGE followed by Coomassie blue staining. Equal volumes of each elution were loaded.

partially disordered Linker domain, which contacts the leucine pocket via Leu261 in helix L1 (Fig. S3A), is also in close proximity to the GATOR2 binding site in our structure (Fig. 5C).
Therefore, changes in the leucine binding state of Sestrin2 could potentially alter the position of the Linker, thereby affecting the availability of the GATOR2 binding site.

Despite these insights, several important questions remain. Fully understanding how leucine binding causes dissociation of Sestrin2 from GATOR2 will likely require the structure of either apo-Sestrin2 or the Sestrin2-GATOR2 complex. Furthermore, understanding the exact mechanism by which Sestrin2 inhibits the mTORC1 pathway awaits the elucidation of the molecular function of GATOR2.

Finally, as a critical regulator of cell growth, mTORC1 is mis-regulated in various human diseases including cancer, diabetes, and aging (Laplante and Sabatini, 2012; Zoncu et al., 2011b). By revealing the mechanism through which a natural small molecule regulates this pathway, our results may enable the identification of compounds to pharmacologically target the nutrient-sensing pathway upstream of mTORC1 in vivo.
Acknowledgements

We thank all members of the Sabatini and Schwartz labs for helpful insights. We also thank Cell Signalling Technologies (CST) for providing many antibodies. The X-ray crystallography was conducted at the Advanced Photon Source Northeastern Collaborative Access Team (APS NE-CAT) beamlines, which are supported by award GM103403 from the US National Institutes of Health (NIH). Use of the APS is supported by the US Department of Energy, Office of Basic Energy Science under contract DE-AC02-06CH11357. This work was supported in part by the NIH Pre-Doctoral Training Grant T32GM007287. This work has also been supported by grants from the US NIH (R01CA103866 and AI47389) the Department of Defense (W81XWH-07-0448) to D.M.S. Fellowship support was provided by the NIH to R.L.W. (T32 GM007753 and F30 CA189333), L.C. (F31 CA180271) T.W. (F31 CA189437). T.W. is also supported by an award from the MIT Whitaker Health Sciences Fund. M.E.P. is supported by the Sally Gordon Fellowship of the Damon Runyon Cancer Research Foundation (DRG-112-12) and a Department of Defense Breast Cancer Research Program Postdoctoral Fellowship (BC120208). D.M.S is an investigator of the Howard Hughes Medical Institute.

Materials and Methods

Materials

Reagents were obtained from the following sources: HRP-labeled anti-rabbit secondary antibody from Santa Cruz Biotechnology; antibodies to phospho-T389 S6K1, S6K1, Sestrin2, Mios and the FLAG epitope from Cell Signaling Technology; antibodies to the HA epitope from Bethyl laboratories; antibody to raptor from Millipore. FLAG M2 affinity gel and amino acids from Sigma Aldrich; RPMI without leucine, arginine, or lysine from Pierce; DMEM from SAFC Biosciences; XtremeGene9 and Complete Protease Cocktail from Roche; Inactivated Fetal Calf Serum (IFS) and SimplyBlue SafeStain from Invitrogen; amino acid-free RPMI from US
Biologicals; [3H]-labeled amino acids from American Radiolabeled Chemicals. The Mios antibody was generously provided by Jianxin Xie (Cell Signaling Technology).

**Protein production and purification**

Full-length, codon-optimized human Sestrin2 was N-terminally fused with a human rhinovirus 3C protease–cleavable His$_{10}$-Arg$_8$-ScSUMO tag and cloned into a PET-Duet-1 bacterial expression vector. This vector was transformed into *Escherichia coli* LOBSTR (DE3) cells (Kerafast)(Andersen et al., 2013). Cells were grown at 37 °C to 0.6 OD, then protein production was induced with 0.2 mM IPTG at 18 °C for 12–14 h. Cells were collected by centrifugation at 6,000g, resuspended in lysis buffer (50 mM potassium phosphate, pH 8.0, 500 mM NaCl, 30 mM imidazole, 3 mM β-mercaptoethanol (βME) and 1 mM PMSF) and lysed with a cell disruptor (Constant Systems). The lysate was cleared by centrifugation at 10,000g for 20 min. The soluble fraction was incubated with Ni-Sepharose 6 Fast Flow beads (GE Healthcare) for 30 min on ice. After washing of the beads with lysis buffer, the protein was eluted in 250 mM imidazole, pH 8.0, 150 mM NaCl and 3 mM βME. The Ni eluate was diluted 1:1 with 10 mM potassium phosphate, pH 8.0, 0.1 mM EDTA and 1 mM dithiothreitol (DTT), and was subjected to cation-exchange chromatography on a 5 ml SP sepharose fast flow column (GE Healthcare) with a linear NaCl gradient. The eluted Sestrin2 was then incubated with 3C protease and dialyzed overnight at 4 °C into 10 mM potassium phosphate, pH 8.0, 150 mM NaCl, 0.1 mM EDTA and 1 mM DTT, followed by a second cation-exchange chromatography run on an SP sepharose fast flow column (GE Healthcare) with a linear NaCl gradient. The protein was further purified via size-exclusion chromatography on a Superdex S200 16/60 column (GE Healthcare) equilibrated in running buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1 mM EDTA and 1 mM DTT). Selenomethionine (SeMet)-derivatized Sestrin2 was prepared as described previously (Brohawn et al., 2008) and purified as the native version, except that the reducing-agent concentration (βME and DTT) was 5 mM in all buffers.

Chapter 2 - Structure of leucine-bound Sestrin2
Crystallization

Purified Sestrin2 was concentrated to 10 mg/ml and incubated in 2 mM leucine for >1 hour prior to setting crystal trays. Crystals were grown at 18 °C by hanging-drop vapor diffusion with 1 µl of protein at 10 mg/ml mixed with an equal volume of reservoir solution containing 100 mM MES, pH 6.0, 1.2 M disodium malonate, and 1% (v/v) Jeffamine ED 2001. 1% (v/v) glycerol as an additive slowed nucleation and improved the morphology of the crystals. Selenomethionine-derivatized Sestrin2 crystallized under identical conditions. Crystals were cryoprotected in mother liquor supplemented with 18% (v/v) glycerol.

Data collection and structure determination

Data collection was performed at the Advanced Photon Source end station 24-IDC at Argonne National Lab. All data-processing steps were carried out with programs provided through SBgrid (Morin et al., 2013). Data reduction was performed with HKL2000 (Otwinowski and Minor, 1997). A complete native dataset was collected to 2.7-Å and a complete SeMet dataset, at the selenium peak wavelength, was collected to 3.0-Å. The phase problem was solved using single-wavelength anomalous dispersion (SAD) and selenium positions (60 total) were determined in HYSS, run as part of the PHENIX AutoSol program (Adams et al., 2010), for the SeMet dataset (space group I23, 5 molecules per asymmetric unit). An interpretable 3.0-Å experimental electron density map was obtained after solvent modification with Parrot from the CCP4 suite (Winn et al., 2011), and manual model building was carried out in Coot (Emsley et al., 2010). Subsequent refinement was carried out with the superior 2.7-Å native data set using phenix.refine.

Chapter 2 - Structure of leucine-bound Sestrin2
Structural analysis

Protein-protein and protein-ligand interfaces were analyzed using PDBePISA (Krissinel and Henrick, 2007). NCBI’s Vector Alignment Search Tool (VAST) (Gibrat et al., 1996) was used to identify structurally related proteins in the PDB. The multiple sequence alignment (MSA) was generated in Jalview (Waterhouse et al., 2009) with the T-Coffee alignment algorithm (Notredame et al., 2000). Sequences of Sestrin2 homologs were obtained via NCBI BLAST searches (Altschul et al., 1990). All structure figures were made in PyMol (Schrodinger, LLC., 2010).

Cell lysis and immunoprecipitation

Cells were rinsed one time with ice-cold PBS and immediately lysed with Triton lysis buffer (1% Triton, 10 mM β-glycerol phosphate, 10 mM pyrophosphate, 40 mM Hepes pH 7.4, 2.5 mM MgCl2 and 1 tablet of EDTA-free protease inhibitor [Roche] (per 25 ml buffer). The cell lysates were cleared by centrifugation at 13,000 rpm at 4°C in a microcentrifuge for 10 minutes. For anti-FLAG-immunoprecipitations, the FLAG-M2 affinity gel was washed 3 times with lysis buffer. 30 µl of a 50/50 slurry of the affinity gel was then added to clarified cell lysates and incubated with rotation for 2 hours at 4°C. Following immunoprecipitation, the beads were washed 4 times with lysis buffer containing 500 mM NaCl. Immunoprecipitated proteins were denatured by the addition of 50 µl of sample buffer and boiling for 5 minutes as described (Kim et al., 2002), resolved by 8%–16% SDS-PAGE, and analyzed by immunoblotting.

For co-transfection experiments in HEK-293T cells, 2.5 million cells were plated in 10 cm culture dishes. Twenty-four hours later, cells were transfected using the polyethylenimine method (Boussif et al., 1995) with the pRK5-based cDNA expression plasmids indicated in the following amounts: 200 ng FLAG-Sestrin2, 200 ng FLAG-Sestrin2-Nterm, 200 ng FLAG-Sestrin2-Cterm, 200 ng HA-Sestrin2-Cterm, and 600 ng metap2. For in vitro dissociation experiments, 20 ng of HA-Sestrin2 wild-type or mutant were transfected into cells stably
expressing FLAG-WDR24. The total amount of plasmid DNA in each transfection was normalized to 5 μg with empty pRK5. Thirty-six hours after transfection, cells were lysed as described above.

For experiments that required amino acid starvation or restimulation, cells were treated as previously described (Tsun et al., 2013). Briefly, cells were incubated in leucine free RPMI for 50 minutes and then restimulated with leucine for 10 minutes.

**Leucine binding assay**

5 million HEK-293T cells were plated in a 15 cm plate four days prior to the experiment. Forty-eight hours after plating, the cells were transfected via the polyethylenimine method with the pRK5-based cDNA expression plasmids indicated in the figures in the following amounts: 12 μg FLAG-metap2, 5 μg FLAG-Sestrin2, 5 μg FLAG-Sestrin2-Nterm, 5 μg FLAG-Sestrin2-Cterm, 5 μg HA-Sestrin2-Cterm, 12 μg HA-GST-Rap2A, 7 μg HA-Sestrin2 wild-type or mutant. The total amount of plasmid DNA in each transfection was normalized to 20 μg total DNA with empty-pRK5. Forty-eight hours after transfection cells were lysed as previously described. If multiple samples of the same type were represented in the experiment, the cell lysates were combined, mixed, and evenly distributed amongst the relevant tubes.

Anti-FLAG or Anti-HA magnetic beads (Pierce) were blocked by rotating in 1 μg/μl bovine serum albumin (BSA) for 20 minutes at 4 °C, then washed twice in lysis buffer and resuspended in an equal volume of lysis buffer. 30 μl of bead slurry was added to each of the clarified cell lysates and incubated as previously described. Post-IP, the beads were washed as previously and incubated for one hour on ice in cytosolic buffer (0.1% Triton, 40 mM HEPES pH 7.4, 10 mM NaCl, 150 mM KCl, 2.5 mM MgCl2) with the appropriate amount of [3H]-labeled amino acids and cold amino acids. At the end of one hour, the beads were aspirated dry and rapidly washed three times with cytosolic buffer. The beads were aspirated dry again and resuspended in 80 μl of cytosolic buffer. Each sample was mixed well and three 10 μl aliquots
were separately quantified using a TriCarb scintillation counter (PerkinElmer). This process was repeated in pairs for each sample, to ensure similar incubation and wash times for all samples analyzed across different experiments.

For each sample, an immunoprecipitation was performed in parallel. After washing four times as previously described, the proteins were eluted in lysis buffer + 500 mM NaCl and 1 mg/ml HA peptide (for 1 hour at 30 °C) or FLAG-peptide (1 hour at 4°C). The eluted proteins were denatured by the addition of sample buffer and boiling for 5 minutes as described, then resolved by 8%-16% SDS-PAGE and analyzed with Coomassie blue stain.

**In vitro Sestrin2-GATOR2 dissociation assay**

HEK-293T cells stably expressing FLAG-WDR24 were transfected with HA-Sestrin constructs as described above. 48 hours after transfection, cells were starved for all amino acids for 50 minutes, lysed and subjected to anti-FLAG immunoprecipitation as described previously. The GATOR2-Sestrin2 complexes immobilized on the FLAG beads were washed twice in lysis buffer with 500 mM NaCl, then incubated for 10 minutes in 1 mL of cytosolic buffer with the indicated concentrations of leucine. The amount of GATOR2 and Sestrin2 that remained bound was assayed by SDS-PAGE and immunoblotting as described previously.

**Differential Scanning Fluorimetry**

The DSF (a.k.a Thermofluor) assays were performed according to the LightCycler 480 instruction manual. Briefly, 5X Sypro orange dye and Sestrin2 at 4 μM were combined with or without leucine or arginine (at the indicated concentrations) in thermal shift buffer (100 mM Tris pH 7.4, 100 mM NaCl, and 1 mM DTT) in a volume of up to 10 μl in one well of a LightCycler Multiwell 384-well plate. Each condition was tested in triplicate. The plate was subjected to a protocol in which the temperature increased from 20° to 85°C at 0.06°C/second. Fluorescence
was recorded and plotted over time, and melting temperatures were calculated as described in the LightCycler 480 instruction manual. Briefly, the negative first derivative of the curve shown (change in fluorescence/change in temperature) was plotted against the temperature. The peak (i.e., lowest point on this curve) reflects the melting temperature. Each reported melting temperature is the mean ± SD for three replicates from one experiment.

Cell lines and tissue culture

HEK-293T cells were cultured in DMEM 10% IFS supplemented with 2 mM glutamine. All cell lines were maintained at 37°C and 5% CO2.

Plasmid preparation

For production of cell lines stably expressing wild-type and mutant Sestrin2 at levels close to endogenous expression, constructs were cloned into the pLJC5 lentiviral construct containing the UBC promoter as previously described.

Generation of CRISPR/Cas9 genetically modified cells

HEK-293T cells with lacking of all three Sestrins were generated as previously described (Wolfson et al., 2016). Briefly, guide RNAs were cloned into the pX330 vector and transiently transfected into HEK-293T cells transfected with 250 ng shGFP pLKO, 1 ug of the pX330 guide construct, and 500 ng of empty pRK5 using XtremeGene9. Cells were single cell sorted with a flow cytometer into the wells of a 96-well plate containing 150 μl of DMEM supplemented with 30% IFS. Cells were grown for two weeks and the resultant colonies were trypsinized and expanded. Clones were validated for loss of the relevant protein via immunoblotting. Sestrin1-3 triple null cells were generated by sequential knock-out of the individual Sestrins.

Lentivirus production and lentiviral transduction
Lentiviruses were produced by transfection of viral HEK-293T cells with either pLJM1-GFP or pLJC5-FLAG-Sestrin2 (wild-type or mutant) constructs in combination with the VSV-G envelope and CMV ΔVPR packaging plasmids. Twenty-four hours after transfection, the media was changed to DMEM with 20% IFS. Forty-eight hours after transfection, the virus-containing supernatant was collected from the cells and passed through a 0.45 μm filter. Target cells were plated in 6-well plates containing DMEM 10% IFS with 8 μg/mL polybrene and infected with virus containing media. Twenty-four hours later, the media was changed to fresh media containing puromycin for selection.

Statistical analysis

Two-tailed t tests were used for comparison between two groups. All comparisons were two-sided, and P values of less than 0.001 were considered to indicate statistical significance.

References


Chapter 2 - Structure of leucine-bound Sestrin2


Peng, M., Yin, N., and Li, M.O. (2014). Sestrins function as guanine nucleotide dissociation inhibitors for Rag GTPases to control mTORC1 signaling. Cell 159, 122-133.


Chapter 2 - Structure of leucine-bound Sestrin2
Chapter 3

Analysis of reported "apo-Sestrin2" structures

Parts of this chapter were first published as:


Abstract

Sestrin2 is a GATOR2 interacting protein that directly binds leucine and is required for the inhibition of mTORC1 under leucine deprivation, indicating that it is a leucine sensor for the mTORC1 pathway. We recently reported the structure of Sestrin2 in complex with leucine (PDB ID: 5DJ4), and demonstrated that mutations in the leucine-binding pocket alter the affinity of Sestrin2 for leucine and result in a corresponding change in the leucine sensitivity of mTORC1 in cells. A lower resolution structure of human Sestrin2 (PDB ID: 5CUF), which was crystallized in the absence of exogenous leucine, showed Sestrin2 to be in a nearly identical conformation as the leucine-bound structure. Based on this observation, it has been argued that leucine binding does not affect the conformation of Sestrin2, and suggested that Sestrin2 may not be a sensor for leucine. Here, we show that simple analysis of the reported "apo"-Sestrin2 structure reveals the clear presence of prominent, unmodeled electron density in the leucine-binding pocket that exactly accommodates the leucine observed in the higher resolution structure. Refining the reported "apo"-structure with leucine eliminates the large Fo-Fc difference density at this position and improves the working and free R-factors of the model. Consistent with this, our own structure of Sestrin2 crystallized in the absence of exogenous leucine also contains electron density that is best explained by leucine. Thus, the structure of apo-Sestrin2 remains elusive.

Introduction
The mechanistic target of rapamycin complex 1 (mTORC1) couples cell growth with the availability of biosynthetic inputs such as amino acids (Dibble and Manning, 2013; Laplante and Sabatini, 2012). Leucine in particular robustly activates mTORC1, and leucine deprivation inhibits mTORC1 signaling in a wide variety of experimental systems (Anthony et al., 2000; Kim et al., 2002; Lynch et al., 2000). Recently, we identified Sestrin2 as a key leucine sensor in mammalian cells (Saxton et al., 2016; Wolfson et al., 2016). Sestrin2 directly binds leucine in vitro, and genetic loss of Sestrin2 and its closely related homologs renders mTORC1 signaling resistant to inhibition by leucine starvation. Mechanistically, Sestrin2 inhibits mTORC1 in the absence of leucine by binding to GATOR2, an upstream activator of mTORC1 (Chantranupong et al., 2014; Parmigiani et al., 2014; Saxton et al., 2016; Wolfson et al., 2016). The addition of leucine triggers the dissociation of Sestrin2 from GATOR2 both in vitro and in cells, suggesting that leucine binding induces a conformational change in Sestrin2 that disrupts this interaction (Saxton et al., 2016; Wolfson et al., 2016). Consistent with this, leucine substantially increases the thermal stability of purified Sestrin2, as is often observed for ligand-receptor complexes (Niesen et al., 2007; Saxton et al., 2016; Wolfson et al., 2016).

We solved the crystal structure of human Sestrin2 in complex with leucine at 2.7 Å resolution, revealing insights into the mechanism of leucine sensing (Saxton et al., 2016). However, Sestrin2 failed to crystallize in the absence of leucine, and the structure of apo-Sestrin2 remained elusive. Subsequently, a 3.5 Å structure of Sestrin2 crystallized without the addition of exogenous leucine was reported, in the same crystal form, showing Sestrin2 in a nearly identical conformation as the leucine-bound structure (Kim et al., 2015). On the basis of this observation, Lee et al. argued that the similarities between the reported "apo" structure of Sestrin2 and the leucine-bound structure are inconsistent with Sestrin2 being a sensor for leucine (Ho et al., 2016; Lee et al., 2016).

To reconcile our structural, biochemical, and cell biological data with the claims made by Lee et al., we reanalyzed their reported "apo"-Sestrin2 structure (PDB ID: 5CUF), revealing the
clear presence of a ligand bound in the leucine-binding pocket. In addition, we report our own 3.0 Å structure of Sestrin2 that we obtained without adding exogenous leucine during the purification process, in which leucine is also present.

Results

Using the corresponding structure factors and atomic model deposited in the Protein Data Bank, we calculated the 2F₀-Fᵣ and F₀-Fᵣ electron density maps for the protein PDB ID: 5CUF. Simple inspection of the F₀-Fᵣ difference map, which highlights discrepancies between experimental data and the structural model, revealed the presence of substantial unexplained electron density (positive 7.5σ peak) at the exact location of the leucine-binding pocket observed in the leucine-bound Sestrin2 (PDB ID: 5DJ4, Fig. 1A). This unmodeled density is easily observed in all 5 crystallographically independent copies of Sestrin2 in the asymmetric unit and is the prime location where modeled density and observed density differ. Furthermore, the unexplained electron density in PDB ID: 5CUF resembles the density corresponding to leucine in the higher resolution PDB ID: 5DJ4 structure (Fig. 1B). To test whether this density may represent leucine, we refined the 5CUF model either with or without leucine built into it. As with PDB ID: 5DJ4, refining PDB ID: 5CUF with leucine eliminated the F₀-Fᵣ map peak at the binding site and improved the R-factors of the model (Fig. 1, B, C and E), consistent with the presence of bound leucine in this structure.

We also analyzed the possibility that one of the buffer components (MES, Tris, sodium, chloride, malonate, TCEP) of the reported crystallization condition of PDB ID: 5CUF may be contributing to the large difference density at the leucine-binding site (Kim et al., 2015). The only ingredient that could possibly be entertained at the modest resolution of 3.5Å as a leucine-substitute is malonate (Fig. 1D). It is the main precipitant at 1.15 M concentration and it could at least mimic the interaction of the carboxyl group of leucine. However, the contacts formed by the hydrophobic leucine sidechain as well as the cationic amine group would still be
unexplained entirely. Nonetheless, refining PDB ID: 5CUF with malonate also eliminated the Fo-
Fc map peak and improved the R-factors of the model (Fig. 1, D and E).

In our own efforts to obtain the structure of apo-Sestrin2, we purified Sestrin2 from E.
coli and crystallized it without the addition of leucine following bacterial cell lysis. While pre-
incubating Sestrin2 with leucine yielded large, high quality crystals within 24-48 hours of setting
drops, the crystals that formed in absence of added leucine were smaller, took much longer to
crystalize (7-14 days), and formed in a background of a heavy protein precipitate. Nonetheless,
we were able to obtain high quality crystals from which we obtained a structure at 3.0 Å
resolution (Table 1). Inspection of the Fc-Fc difference map for this structure again revealed a
large positive peak in the location of the leucine-binding pocket that strongly resembles the
electron density of leucine (Fig. 2A). As with PDB ID: 5CUF and PDB ID: 5DJ4, refining the
structure with leucine eliminated the Fo-Fc peak and improved the model statistics (Fig. 2, B and
D). Furthermore, at this resolution leucine is clearly the best candidate to fit into this density, as
malonate does not fully explain the difference density at this position and did not improve the
Rfree as substantially as when the structure was refined with leucine (Fig. 2, C and D).

Discussion

The leucine-binding site of Sestrin2 consists of a shallow hydrophobic pocket flanked on
each side by opposing positive and negative charges contributed by Arg^{390} and Glu^{451},
respectively, making it well suited to specifically accommodate zwitterionic α-amino acids with
short aliphatic sidechains (Saxton et al., 2016). Among the amino acids, Sestrin2 has the
highest affinity for leucine, with approximately 15-fold and 30-fold lower affinity for the closely
related amino acids methionine and isoleucine, respectively, and undetectable affinity for the
Figure 1: Identification of ligand in the “apo”-Sestrin2 structure reported by Lee et al.

(1A) Ribbon diagrams of the reported “apo”-Sestrin2 structure (PDB ID: 5CUF, pink, left) and the leucine-bound structure (PDB ID: 5DJ4, light blue, right). The largest peak in the 3.5 Å 5CUF Fo-Fc difference map, contoured at 3σ, is shown as a green mesh in the 5CUF structure. The bound leucine built into the 5DJ4 structure is shown in orange. (1B) Close up view of the leucine-binding pocket of Sestrin2 in the 3.5 Å 5CUF (left) and 2.7 Å 5DJ4 structures (right), refined without leucine built into the model. The Fo-Fc (green mesh) and 2Fo-Fc (blue mesh) electron density maps were contoured at 3σ and 1σ, respectively, and shown in the location of the bound leucine. (1C) The PDB ID: 5CUF and PDB ID: 5DJ4 structures were refined with...
leucine built into the model and shown in the same view as in 1B. The F_o-F_c and 2F_o-F_c electron density maps are depicted as in 1B. Hydrogen bonds and electrostatic interactions are shown as black dashed lines. (1D) The PDB ID: 5CUF structure was refined with malonate built into the model and shown in the same view as in 1B. The F_o-F_c and 2F_o-F_c electron density maps are depicted as in 1B. (1E) The R_w/R_f (%) for the PDB ID: 5CUF structure refined with no ligand, leucine, or malonate modeled.

Figure 2

A

3.0 Å "apo"-Sestrin2 (5TON, this study)

refined with no ligand bound:

B

3.0 Å "apo"-Sestrin2 (5TON, this study)

refined with leucine bound:

C

3.0 Å "apo"-Sestrin2 (5TON, this study)

refined with malonate bound:

D

<table>
<thead>
<tr>
<th>5TON</th>
<th>R_w/R_f (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>apo</td>
<td>19.4/22.7</td>
</tr>
<tr>
<td>leu</td>
<td>18.9/22.5</td>
</tr>
<tr>
<td>mal</td>
<td>19.3/22.7</td>
</tr>
</tbody>
</table>

Figure 2: Sestrin2 crystalizes in complex with leucine without the addition of exogenous leucine.

(2A) Close up view of the leucine-binding pocket in Sestrin2 (PDB ID: 5TON, this study) refined without leucine built into the model. The F_o-F_c (green mesh) and 2F_o-F_c (blue mesh) electron density maps, contoured at 3σ and 1σ, respectively, are shown as in 1B. (2B) Close up view of the leucine-binding pocket in Sestrin2 refined with leucine built into the model. The F_o-F_c and 2F_o-F_c electron density maps are depicted as in 2A. Hydrogen bonds and electrostatic interactions are shown as black dashed lines. (2C) Close up view of the leucine-binding pocket in Sestrin2 refined with malonate built into the model. The F_o-F_c and 2F_o-F_c electron density maps are depicted as in 2A. Hydrogen bonds and electrostatic interactions are shown as black dashed lines. (2D) The R_w/R_f (%) for the PDB ID: 5TON structure (this study) refined with no ligand, leucine, or malonate modeled.
Table 1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Protein</th>
<th>“Pseudo-apo” Sestrin2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organism</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>PDB ID</td>
<td>5T0N</td>
</tr>
</tbody>
</table>

**Data collection**

<table>
<thead>
<tr>
<th>Space group</th>
<th>123</th>
</tr>
</thead>
<tbody>
<tr>
<td>a, b, c (Å)</td>
<td>292.26, 292.26, 292.26</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.0, 90.0, 90.0</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9792</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>200.0 – 3.0 (3.05 – 3.00)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>40.7 (41.9)</td>
</tr>
<tr>
<td>R_meas (%)</td>
<td>28.0 (&gt;100)</td>
</tr>
<tr>
<td>R_p.i.m. (%)</td>
<td>4.4 (37.8)</td>
</tr>
<tr>
<td>I/σ</td>
<td>19.5 (2.0)</td>
</tr>
</tbody>
</table>

**Refinement**

<table>
<thead>
<tr>
<th>Resolution range (Å)</th>
<th>103.3 – 3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_work (%)</td>
<td>18.9</td>
</tr>
<tr>
<td>R_free (%)</td>
<td>22.5</td>
</tr>
</tbody>
</table>

**Number of Reflections:**

- Total: 82205
- R_free reflections: 1637

**Number of non-hydrogen atoms:**

- Protein atoms: 14648
- Water atoms: 0

**R.m.s. deviations:**

- Bond lengths (Å): 0.010
- Bond angles (°): 1.037

**Average B factors (Å²):**

- Protein: 49.0

**Ramachandran (%):**

- Favored (%): 96.4
- Allowed (%): 3.0
- Outlier (%): 0.7
remaining natural amino acids (Wolfson et al., 2016). In addition, while it is impossible to definitively conclude whether the density present in the 3.5 Å PDB ID: 5CUF structure corresponds to leucine or another ligand such as malonate, our own 3.0 Å structure clearly favors leucine as the best candidate to explain the observed density. In any case, these data do conclusively demonstrate that PDB ID: 5CUF represents the ligand-bound form of Sestrin2, not the apo-Sestrin2 conformation as reported by Lee et al.

Although leucine is highly abundant in bacterial cells (Sajed et al., 2016), it may seem surprising that Sestrin2 can remain stably bound to leucine throughout the purification process given the reported 20 µM affinity of Sestrin2 for leucine (Wolfson et al., 2016). One possibility is that the off-rate of the Sestrin2-leucine interaction is substantially lower in the absence of additional protein components or post-translational modifications present only in mammalian cells, such that a small fraction of Sestrin2 remains bound to leucine even following purification. Crystallization is, after all, a purification method and thus perfectly suited to enrich a small fraction of leucine-bound Sestrin2 in a large background of non-crystallizing apo-Sestrin. Because all three Sestrin2 crystal structures crystallize under highly similar conditions, in the same crystal form, and that the largest and best diffracting crystals grow in the presence of leucine, we conclude that all reported structures are best interpreted as being leucine bound.

In summary, we argue that the PDB ID: 5CUF structure depicts Sestrin2 in the ligand-bound conformation, not the apo-conformation as reported by Lee et al (Lee et al., 2016). The leucine-dependent dissociation of Sestrin2 from GATOR2 represents the critical first step in the activation of mTORC1 by leucine, and understanding how leucine-binding alters the conformation of Sestrin2 to disrupt its interaction with GATOR2 remains a major open question in the field.
Acknowledgements

We thank all members of the Sabatini and Schwartz laboratories for helpful insights. This work is based on research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. This work has been supported by grants from NIH (R01CA103866 and AI47389) and the U.S. Department of Defense (W81XWH-07-0448) to D.M.S. D.M.S. is an investigator of the Howard Hughes Medical Institute. R.A.S, T.U.S, and D.M.S designed the research plan. R.A.S performed the crystallographic experiments with input from K.E.K. and T.U.S. R.A.S, K.E.K and T.U.S analyzed the structural data. R.A.S, T.U.S and D.M.S wrote and all authors edited the manuscript.

D.M.S. is a founder, a member of the Scientific Advisory Board, a paid consultant, and a shareholder of Navitor Pharmaceuticals, which is targeting for therapeutic benefit the amino acid sensing pathway upstream of mTORC1. Coordinates and structure factors for the x-ray crystal structure of the "pseudo-apo" Sestrin2 have been deposited in the Protein Data Bank (PDB) with accession code 5TON.

Materials and Methods

Protein production and purification

Human Sestrin2 was expressed and purified as described previously (6). Briefly, full-length, codon-optimized human Sestrin2 was N-terminally fused with a human rhinovirus 3C protease–cleavable His$_{10}$-Arg$_{8}$-ScSUMO tag and cloned into a PET-Duet-1 bacterial expression vector. This vector was transformed into Escherichia coli LOBSTR (DE3) cells.
(Kerafast)(Andersen et al., 2013). Cells were grown at 37°C to 0.6 OD, then protein production was induced with 0.2 mM IPTG at 18 °C for 12–14 h. Cells were collected by centrifugation at 6,000g, resuspended in lysis buffer (50 mM potassium phosphate, pH 8.0, 500 mM NaCl, 30 mM imidazole, 3 mM β-mercaptoethanol (βME) and 1 mM PMSF) and lysed with a cell disruptor (Constant Systems). The lysate was cleared by centrifugation at 10,000g for 20 min. The soluble fraction was incubated with Ni-Sepharose 6 Fast Flow beads (GE Healthcare) for 30 min on ice. After washing of the beads with lysis buffer, the protein was eluted in 250 mM imidazole, pH 8.0, 150 mM NaCl and 3 mM βME. The Ni eluate was diluted 1:1 with 10 mM potassium phosphate, pH 8.0, 0.1 mM EDTA and 1 mM dithiothreitol (DTT), and was subjected to cation-exchange chromatography on a 5 ml SP sepharose fast flow column (GE Healthcare) with a linear NaCl gradient. The eluted Sestrin2 was then incubated with 3C protease and dialyzed overnight at 4 °C into 10 mM potassium phosphate, pH 8.0, 150 mM NaCl, 0.1 mM EDTA and 1 mM DTT, followed by a second cation-exchange chromatography run on an SP sepharose fast flow column (GE Healthcare) with a linear NaCl gradient. The protein was further purified via size-exclusion chromatography on a Superdex S200 16/60 column (GE Healthcare) equilibrated in running buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1 mM EDTA and 1 mM DTT).

Crystallization

Crystals were grown at 18 °C by hanging-drop vapor diffusion with 1 µl of Sestrin2 at 5 mg/ml mixed with an equal volume of reservoir solution containing 100 mM MES, pH 6.0, 1.2 M disodium malonate, and 1% (v/v) Jeffamine ED 2001. Drops were microseeded to obtain well diffracting crystals. Microseeds were obtained as described previously (Luft and DeTitta, 1999). Briefly, a single 2 µl drop containing crystals of Sestrin2 was diluted into a 50 µl solution of mother liquor (100 mM MES, pH 6.0, 1.2 M disodium malonate, and 1% (v/v) Jeffamine ED 2001), vortexed for 5 minutes in the presence of a polystyrene bead, then diluted again into 500 µl of mother liquor. This solution was then diluted at a ratio of 1:1,000 to obtain the final

Chapter 3 - Analysis of reported “apo”-Sestrin2 structures
microseed stock. 0.2 μl of microseed stock was then added to the 2 μl hanging drop described above (for a final total dilution of 1:2,500,000). Crystals were cryoprotected in mother liquor supplemented with 18% (v/v) glycerol.

Data collection and structure determination

Data collection was performed at the Advanced Photon Source end station 24-IDC at Argonne National Lab. A complete native dataset was collected to 3.0 Å. All data-processing steps were carried out with programs provided through SBgrid (Morin et al., 2013). Data reduction was performed with HKL2000 (Otwinowski and Minor, 1997). The phase problem was solved by molecular replacement using our published Sestrin2 structure (PDB ID: 5DJ4) as the search model in the Phaser-MR program, run as part of the Phenix Autosol program (space group I23, 5 molecules per asymmetric unit)(Adams et al., 2010). To reduce potential model bias, we used a search model lacking the bound leucine and surrounding protein residues. An interpretable 3.0-Å electron density map was obtained and model building was carried out in Coot (Emsley et al., 2010). Subsequent refinement was carried out using phenix.refine.

Structure and electron density map analysis

For analysis of published structures 5DJ4 and 5CUF, the F_{o} - F_{c} and 2F_{o} - F_{c} maps were calculated from the models and structure factors deposited in the Protein Data Bank using phenix.maps, and visualized in Coot. Refinement of the models either with or without bound-leucine built in was performed using phenix.refine. XYZ-coordinates and individual B-factors were refined over three cycles. All structure figures were made in PyMol (Schrodinger, LLC, 2010).

Chapter 3 - Analysis of reported “apo”-Sestrin2 structures
References


Chapter 3 - Analysis of reported "apo"-Sestrin2 structures
Chapter 4

Mechanism of arginine sensing by CASTOR1 upstream of mTORC1

Parts of this chapter were first published as:


Abstract

The mechanistic Target of Rapamycin Complex 1 (mTORC1) is a major regulator of eukaryotic growth that coordinates anabolic and catabolic cellular processes with inputs such as growth factors and nutrients, including amino acids. In mammals, arginine is particularly important and promotes diverse physiological effects including immune cell activation, insulin secretion, and muscle growth, largely through activation of mTORC1. Arginine activates mTORC1 upstream of the Rag GTPases, through either the lysosomal amino acid transporter SLC38A9 or the GATOR2-interacting CASTOR1. However, the mechanism by which the mTORC1 pathway detects and transmits the arginine signal has been elusive. Here, we present the 1.8 Å crystal structure of arginine-bound CASTOR1. Homodimeric CASTOR1 binds arginine at the interface of two ACT domains, enabling allosteric control of the adjacent GATOR2-binding site to trigger dissociation from GATOR2 and the downstream activation of mTORC1. Our data reveal that CASTOR1 shares substantial structural homology with the lysine-binding regulatory domain of prokaryotic aspartate kinases, suggesting that the mTORC1 pathway exploited an ancient amino-acid-dependent allosteric mechanism to acquire arginine sensitivity. Together, these results establish a structural basis for arginine sensing by the mTORC1 pathway and provide insights into the evolution of a mammalian nutrient sensor.
Introduction

The mTOR Complex 1 (mTORC1) pathway coordinates eukaryotic cell growth with the availability of growth factors and nutrients, including amino acids (Dibble and Manning, 2013; Laplante and Sabatini, 2012; Saxton and Sabatini, 2017). mTORC1 drives mass accumulation by promoting anabolic processes, such as protein synthesis, while inhibiting catabolic ones, such as autophagy. Due to this critical role in modulating cell growth, mTORC1 and its upstream regulators are frequently deregulated in diseases such as cancer and diabetes (Saxton and Sabatini, 2017).

In mammals, the amino acid arginine is a particularly important input for mTORC1 (Ban et al., 2004). In addition to its role as a fundamental building block of proteins, arginine also functions as a signaling molecule that promotes diverse physiological effects including immune cell activation, insulin secretion, and muscle growth, in large part through activation of mTORC1 (Bronte and Zanovello, 2005; Floyd et al., 1966; Yao et al., 2008). However, the mechanism through which mTORC1 senses arginine is poorly understood.

While growth factors, oxygen, and other inputs regulate mTORC1 catalytic activity through the Tuberous Sclerosis Complex (Menon et al., 2014; Saito et al., 2005), amino acids signal by altering the nucleotide-bound state of the Rag GTPases (Sancak et al., 2008), which recruit mTORC1 to the lysosomal surface where it is activated (Sancak et al., 2010). Amino acids inside the lysosome have been proposed to signal to the Rags through a lysosomal-membrane associated complex comprised of the Ragulator, v-ATPase, and amino acid transporter SLC38A9 (Bar-Peled et al., 2012; Rebsamen et al., 2015; Wang et al., 2015; Zoncu et al., 2011a). Cytosolic amino acids, however, regulate the Rags through a distinct pathway involving the multimeric GATOR1 and GATOR2 complexes (Bar-Peled et al., 2013). The identification of the GATOR2-interacting protein Sestrin2 as a leucine sensor suggests that GATOR2 is a critical signaling node in the amino acid sensing pathway upstream of mTORC1 (Chantranupong et al., 2014; Parmigiani et al., 2014; Saxton et al., 2016; Wolfson et al., 2016).

Chapter 4 - Mechanism of arginine sensing by CASTOR1
Recently, we identified CASTOR1 (Cellular Arginine Sensor upstream of mTORC1), a previously uncharacterized protein that interacts with GATOR2 only in the absence of arginine (Chantranupong et al., 2016). CASTOR1 binds arginine and is required for arginine-dependent regulation of mTORC1 in cells, indicating that it is an arginine sensor for the mTORC1 pathway. Despite these insights however, the mechanism through which CASTOR1 specifically detects arginine and signals its presence to GATOR2 is unknown.

Here, we present the 1.8 Å crystal structure of CASTOR1 in complex with arginine, providing a structural mechanism for arginine sensing by the mTORC1 pathway.

**Results**

To understand the molecular mechanisms through which CASTOR1 detects arginine and signals its presence to mTORC1, we determined the crystal structure of arginine-bound CASTOR1 to 1.8 Å resolution. CASTOR1 forms a rod-shaped homodimer, with the monomers associated in a side-by-side manner and rotated 180° with respect to each other (Fig. 1a). While sequence analysis of CASTOR1 predicted the presence of two ACT (Aspartate Kinase, Chorismate mutase, TyrA) domains (Aravind and Koonin, 1999; Chantranupong et al., 2016), the structure reveals that each monomer actually contains four tandem ACT domains. ACT1 displays the canonical βαββαβ ACT domain topology (Chipman and Shaanan, 2001; Grant, 2006), whereas ACT3 and ACT4 each lack the final β-strand and ACT2 contains two additional β-strands (Fig. 1a, Extended Data Fig. 1a).

The dimerization interface buries ~950 Å² of surface area at the intersection of α1 and α5, from ACT1 and ACT3, respectively (Fig. 1b). Four inward facing isoleucines (Ile28 and Ile202) form the hydrophobic core of the symmetrical interface, flanked on each side by tyrosine-histidine pairs (His25 and Tyr207) that form both pi-stacking and hydrogen-bond contacts with the opposing monomer (Fig. 1b). To understand the importance of dimerization for CASTOR1 function, we generated constitutively monomeric mutants of CASTOR1 (Y207S and
I202E, Fig. 1c). Interestingly, although dimerization is dispensable for arginine binding (Extended Data Fig. 2a), these mutants interacted weakly with GATOR2 and failed to inhibit mTORC1 signalling in cells (Fig. 1c, Extended Data Fig. 2b), indicating that CASTOR1 must be dimeric to robustly inhibit GATOR2 upon arginine starvation.

CASTOR1 binds arginine through a narrow pocket at the interface of ACT2 and ACT4, distal to the dimerization interface (Fig. 1a, 2a, b). The side-chain of arginine points towards a loop connecting β15 and β16 (β15-loop), where the backbone carbonyls of Thr300, Phe301, and Phe303 coordinate the guanidinium group of arginine (Fig. 2a). Immediately adjacent to the β15-loop, the anionic side-chain of Asp304 forms an additional stabilizing salt-bridge with the cationic arginine side-chain (Fig. 2a). On the opposite side of the pocket, the hydroxyl side-chain of Ser111 and backbone carbonyl of Val112 in the α3-loop anchor the free amine of arginine in place, while the free carboxyl points towards a water-filled cavity that separates it from ACT2 (Fig 2a, b). Consistent with a critical role for these contacts in arginine sensing by CASTOR1, mutation of either Ser111 or Asp304 (S111A, D304A) abolished the arginine binding capacity of CASTOR1 in vitro (Fig. 2c). Furthermore, when expressed in HEK-293T cells, these mutants bound constitutively to GATOR2 and strongly inhibited mTORC1 signalling even in the presence of arginine (Fig. 2D).

Together, these data explain the molecular determinants of specificity in the CASTOR1-arginine interaction. While Ser111 fixes the position of the free amine, the location of the β15-loop and Asp304 sets a strict length requirement for the bound ligand (Extended Data Fig. 3a). In addition, the positions of the three hydrogen-bond donating nitrogens in the guanidinium group facilitate contacts with both the carbonyl oxygens in the β15-loop and the side-chain of Asp304 (Fig. 2a). Finally, the gap behind the free carboxyl group of arginine suggests that CASTOR1 can tolerate ligands with modifications to that functional group (Fig. 2b). We tested these predictions by interrogating the ability of various arginine analogues to disrupt the CASTOR1-GATOR2 interaction in vitro (Fig. 2e, Extended Data Fig. 3b). Consistent with our
structural analysis, while the carboxy-modified arginine-methyl ester triggered full dissociation of CASTOR1 from GATOR2, compounds with alterations to the guanidium group, α-amine, or side-chain length had no effect.

Table 1. Data collection and refinement statistics (SAD)

<table>
<thead>
<tr>
<th></th>
<th>CASTOR1 + Arg Native</th>
<th>CASTOR1 + Arg SeMet</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organism</strong></td>
<td>H. sapiens</td>
<td>H. sapiens</td>
</tr>
<tr>
<td><strong>PDB ID</strong></td>
<td>5i2C</td>
<td></td>
</tr>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>P2₁</td>
<td>P2₁</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>91.39, 82.60, 96.67</td>
<td>91.76, 82.35, 96.71</td>
</tr>
<tr>
<td>a, b, g (°)</td>
<td>90, 116.23, 90</td>
<td>90, 116.04, 90</td>
</tr>
<tr>
<td><strong>Wavelength (Å)</strong></td>
<td>0.9792</td>
<td>0.9792</td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>86.7 – 1.80</td>
<td>86.89 – 2.20</td>
</tr>
<tr>
<td><strong>R_sym (%)</strong></td>
<td>7.2 (62.8)</td>
<td>10.4 (&gt;100)</td>
</tr>
<tr>
<td><strong>I / sI</strong></td>
<td>25.9 (1.2)</td>
<td>22.6 (1.4)</td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
<td>97.85 (87.1)</td>
<td>98.2 (98.1)</td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td>3 (2.5)</td>
<td>6.4 (5.9)</td>
</tr>
<tr>
<td><strong>Anomalous Completeness (%)</strong></td>
<td>96.8</td>
<td></td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>86.71 – 1.80</td>
<td></td>
</tr>
<tr>
<td><strong>No. reflections</strong></td>
<td>116,883</td>
<td></td>
</tr>
<tr>
<td>R_work / R_free</td>
<td>17.2% / 20.4%</td>
<td></td>
</tr>
<tr>
<td><strong>No. atoms</strong></td>
<td>9,872</td>
<td>9,912</td>
</tr>
<tr>
<td>Protein</td>
<td>9,012</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>796</td>
<td></td>
</tr>
<tr>
<td><strong>Average B-factors (Å)</strong></td>
<td>40.2</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>40.0</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>26.8</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>46.4</td>
<td></td>
</tr>
<tr>
<td><strong>R.m.s. deviations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>0.85</td>
<td></td>
</tr>
</tbody>
</table>

*Values in parentheses are for highest-resolution shell.
Extended Data Figure 1: Multiple sequence alignment of CASTOR1 homologues

a) Expanded Multiple Sequence Alignment of CASTOR1 homologues from various organisms. Positions are colored white to blue according to increasing sequence identity. Secondary structure features are labeled and colored by ACT domain as in 1A.
Figure 1: Architecture of human CASTOR1

(A) Two orthogonal views of the arginine-bound CASTOR1 homodimer shown as ribbon diagrams, with ACT-domains 1-4 colored in green, purple, wheat, and pink, respectively. The stick representation of the bound arginine is shown in yellow. Disordered regions not present in the crystal structure are omitted. (B) Close-up view of the CASTOR1 dimerization interface, with side chains of key residues represented in stick form. ACT domains are colored as in 1a. (C) Dimerization deficient CASTOR1 Y207S and I202E mutants display weaker interactions with endogenous GATOR2. HEK-293T cells transiently expressing FLAG-tagged CASTOR1 wild type (WT) and the indicated HA-tagged constructs were starved of arginine for 50 min and, where indicated, restimulated for 10 min. HA-immunoprecipitates were generated from cell lysates and analyzed by immunoblotting for the indicated proteins. Mios was used as a representative GATOR2 component.
Extended Data Figure 2: Dimerization deficient CASTOR1 mutants bind arginine but fail to inhibit mTORC1 in cells

(A) The dimerization deficient CASTOR1 Y207S and I202E mutants bind arginine *in vitro*. FLAG-immunoprecipitates prepared from HEK-293T cells transiently expressing indicated FLAG-tagged proteins were used in binding assays with [3H]arginine as described in the methods. Unlabeled arginine was included as a competitor where indicated. Values are Mean ± SD for 3 technical replicates from one representative experiment. 

(B) Dimerization deficient CASTOR1 Y207S and I202E mutants fail to inhibit mTORC1. HEK-293T cells transiently expressing FLAG-S6K1 and HA-tagged CASTOR1 WT, Y207S, or I202E were starved of arginine for 50 min and, where indicated, restimulated for 10 min. FLAG-immunoprecipitates were prepared from lysates and analyzed as in 1C. Phospho-S6K1 was used as an indicator of mTORC1 activity.
Figure 2: The arginine-binding pocket of CASTOR1

(A) Close-up view of the arginine-binding pocket in CASTOR1, together with its F_o-F_c electron density map calculated and contoured at 4a from an omit map lacking arginine. The stick model of the bound arginine is shown in yellow. Hydrogen bonds or salt-bridges are shown as black dashed lines. Residues 269-273 of the β14-loop are omitted for clarity. ACT domains are colored as in 1a. (B) Steric view of the arginine-binding pocket, depicting the surface representation of CASTOR1 and stick model of arginine (yellow). The β14-loop (residues 269-280) is omitted for clarity. (C) CASTOR1 S111A and D304A mutants do not bind arginine in vitro. FLAG-immunoprecipitates prepared from HEK-293T cells transiently expressing indicated FLAG-tagged proteins were used in binding assays with [3H]Arginine as described in the methods. Unlabeled arginine was included as a competitor where indicated. Values are Mean ± SD for 3 technical replicates from one representative experiment. (D) The CASTOR1 S111A and D304A mutants constitutively bind GATOR2 and inhibit mTORC1 signaling in cells. HEK-293T cells transiently expressing FLAG-S6K1 and the indicated HA-tagged constructs were starved of arginine for 50 min and, where indicated, restimulated for 10 min. Both FLAG- and HA-immunoprecipitates were prepared from lysates and analyzed as in 1c. (E) Effects of various arginine analogues on the CASTOR1-GATOR2 interaction in vitro. HEK-293T cells transiently expressing HA-CASTOR1 WT were starved of arginine for 50 min. HA-immunoprecipitates were prepared from cell lysates then incubated with 400 μM of the indicated compounds for 20 min and analyzed as in 1c.
Extended Data Figure 3

(A) Comparison of the arginine-bound pocket of human CASTOR1 with a model of the pocket with lysine in place of arginine. Arginine and lysine stick representations are shown in yellow and orange, respectively. The distances in the lysine-bound model, 3.8 Å and 5.0 Å, are beyond the range of standard hydrogen-bonds and salt-bridges, respectively. ACT domains are labeled as in 1A.

(B) Chemical structures of arginine analogues used in Fig. 2E. Differences relative to L-Arginine are highlighted in orange boxes.

Extended Data Figure 3: Model of lysine-binding in CASTOR1 (A) Comparison of the arginine-bound pocket of human CASTOR1 with a model of the pocket with lysine in place of arginine. Arginine and lysine stick representations are shown in yellow and orange, respectively. The distances in the lysine-bound model, 3.8 Å and 5.0 Å, are beyond the range of standard hydrogen-bonds and salt-bridges, respectively. ACT domains are labeled as in 1A. (B) Chemical structures of arginine analogues used in Fig. 2E. Differences relative to L-Arginine are highlighted in oranges boxes.
In addition to the main pocket contacts described above, a highly conserved, glycine-rich loop connecting β14 and α7 in ACT4 ("β14-loop", residues 269-280) wraps over the arginine-pocket, fully burying the bound ligand (Fig. 2a, 3a, Extended Data Fig. 1a). The β14-loop forms several hydrogen-bonds with arginine through the backbone amides of Gly279 and Ile280, as well as the backbone oxygens of Gly274 and Glu277 (Fig. 2a, 3a). The ordered conformation of the β14-loop also places it just along the ACT2-ACT4 interface, enabling it to form several intramolecular contacts with residues in ACT2 (Fig. 3a). Cys278 hydrogen bonds with the backbones of Val110 and Ser111 in the α3-loop, while Asp276 forms a salt-bridge with Arg126. In addition, Glu277 extends in the opposite direction to form another salt-bridge with His175. (Fig. 3a). Thus, the β14-loop facilitates the formation of numerous inter-ACT-domain contacts in the presence of arginine. Indeed, the arginine and β14-loop contribute ~40% of the total buried surface area in the ACT2-ACT4 interface of the arginine-bound structure (390 Å² out of 980 Å²).

The glycine-rich β14-loop is predicted to have a high propensity for disorder, and our structure suggests that these inter-ACT-domain contacts could stabilize it in an ordered conformation over the bound arginine. Indeed, mutation of key residues in both the β14-loop (D276A, E277A, C278A) and the adjacent ACT domains (R126A, H175A) significantly reduced the arginine binding capacity of CASTOR1 (Fig. 3b, c), indicating that the inter-ACT-domain contacts formed by the β14-loop are required for arginine sensing by CASTOR1. In addition, we found that the N-terminal (ACT1 and ACT2) and C-terminal (ACT3 and ACT4) halves of CASTOR1 associate in both an arginine- and β14-loop-dependent manner when expressed as separate polypeptides in HEK-293T cells (Fig. 3d), indicating that arginine likely induces a conformational change in CASTOR1 by stabilizing the ACT2-ACT4 interaction.
Figure 3: Arginine facilitates the intramolecular association of ACT2 and ACT4 domains of CASTOR1

(A) Top-down view of the arginine- and β14-loop-mediated contacts between ACT2 and ACT4. Hydrogen bonds and salt-bridges are shown as black dashed lines. Arginine is represented as a stick model (yellow). (B) CASTOR1 D276A, R126A, E277A, H175A, and C278A mutants display reduced arginine-binding capacity in vitro. Binding assays were performed and immunoprecipitates analyzed as in 2c. Values are Mean ± SD for 3 technical replicates from one representative experiment. (C) The CASTOR1 D276A, R126A, E277A, H175A, and C278A mutants constitutively bind GATOR2 in cells. HEK-293T cells transiently expressing the indicated HA-tagged constructs were starved of arginine for 50 min and, where indicated, restimulated for 10 min. HA-immunoprecipitates prepared and analyzed as in 1c. (D) CASTOR1 ACT1-2 (1-169) and CASTOR1 ACT3-4 (169-329) associate in an arginine- and β14-loop dependent manner. HEK-293T cells transiently expressing the indicated HA-tagged constructs were starved of arginine for 60 min and, where indicated, restimulated for 60 min. HA-immunoprecipitates were prepared and analyzed as in 1c.

In addition to CASTOR1, human cells express a related protein, CASTOR2, which shares 63% sequence identity but does not bind arginine. Although the regions of CASTOR1 directly involved in arginine binding are well conserved (Extended Data Fig. 1a), we identified residues along the ACT2-ACT4 interface (His108 to Val110) that differ between CASTOR1 and CASTOR2 (Extended Data Fig. 4a). Switching these residues of CASTOR1 with those in...
CASTOR2 abrogated arginine binding in vitro and converted CASTOR1 to a nearly constitutive GATOR2-interactor in cells, resembling CASTOR2 (Extended Data Fig. 4b-d). Interestingly, these residues immediately precede Ser111 and hydrogen bond with Cys278 in the β14-loop (Fig. 3a, Extended Data Fig. 4a), suggesting that their identity may be critical for the proper positioning of the α3-loop to enable arginine binding and/or the association of ACT2 and ACT4. The corresponding mutation in CASTOR2 (108QNI-110HHV) however was not sufficient to confer arginine binding, suggesting that additional differences also contribute (Extended Data Fig. 4d).

Extended Data Figure 4

(A) Multiple sequence alignment of human CASTOR1 and CASTOR2, highlighting differences in amino acid sequence that are in close proximity to arginine binding residues in CASTOR1. (B) The CASTOR1 HHV108-110QNI mutant constitutively binds GATOR2 in cells. HEK-293T cells transiently expressing HA-metap2 or the indicated HA-tagged CASTOR1 constructs were starved of arginine for 50 min and, where indicated, restimulated for 10 min. HA-immunoprecipitates prepared and analyzed as in 1C. (C) The CASTOR1 HHV108-110QNI
mutant displays reduced arginine-binding capacity in vitro. Binding assays were performed with the indicated CASTOR1 or CASTOR2 constructs and immunoprecipitates analyzed as in 2C. Values are Mean ± SD for 3 technical replicates from one representative experiment. (D) Comparison of the CASTOR1 HHV108-110QNI mutant and WT CASTOR2. HEK-293T cells transiently expressing HA-metap2 or the indicated HA-tagged CASTOR1 or CASTOR2 constructs were starved of arginine for 50 min and, where indicated, restimulated for 10 min. HA-immunoprecipitates prepared and analyzed as in 1C.

To understand how arginine induces dissociation of CASTOR1 from GATOR2, we identified five highly conserved sites in CASTOR1 required for the interaction with GATOR2 (Y118, Q119, D121, E261, D292, Fig. 4a, Extended Data Fig. 1a). Importantly, these mutants still bind arginine in vitro and homodimerize when expressed in cells (Extended Data Fig. 5a, b). Interestingly, these residues cluster along the surface of the ACT2-ACT4 interface, adjacent to but on the opposite face of the protein as the arginine-binding pocket (Fig. 4b, c). Glu261 and Asp292 are closely linked to the β14-loop, separated only by β14 and helix α7, respectively (Fig. 4c). Furthermore, the critically important residue Asp121 is buried in the ACT2-ACT4 interface, providing one explanation for why the arginine-bound conformation is incompatible with the CASTOR1-GATOR2 interaction (Fig. 4c).

Together, these results suggest a model where arginine binding orders the glycine-rich β14-loop to enable the intramolecular association of ACT2 and ACT4 (Fig. 3a-d). This association of these domains would alter the positions and exposure of the residues required for GATOR2 binding, which also lie along the ACT2-ACT4 interface (Fig. 4a-c), thereby triggering the dissociation of CASTOR1 from GATOR2 and the subsequent activation of mTORC1 (Fig. 4e).

The observations that CASTOR1 both inhibits mTORC1 signalling and interacts with GATOR2 in an arginine-sensitive manner suggests that CASTOR1 may regulate mTORC1 by inhibiting GATOR2, a mechanism analogous to that of the recently identified leucine sensor Sestrin2 (Chantranupong et al., 2014; Parmigiani et al., 2014; Saxton et al., 2016; Wolfson et al., 2016). Using our GATOR2 binding-deficient mutants, we were able to test this hypothesis.
Indeed, in contrast to wild-type CASTOR1, the YQ118-119AA and D121A mutants both failed to inhibit mTORC1 signalling in cells (Fig. 4d). Moreover, due to their ability to dimerize with endogenous CASTOR1, these mutants also functioned as dominant negatives, rendering mTORC1 fully resistant to arginine starvation (Fig. 4d). Thus, the CASTOR1-GATOR2 interaction is required to signal arginine deprivation to mTORC1.

Although defined by their common topology, ACT domains are highly diverse in sequence and form a wide range of structural assemblies (Chipman and Shaanan, 2001; Grant, 2006). Comparison of our structure with other ACT domain-containing proteins in the Protein Data Bank (PDB) revealed that CASTOR1 shares substantial structural homology with the allosteric regulatory domains of bacterial Aspartate Kinases (AKs), including those found in E.

Figure 4: The GATOR2 binding site of CASTOR1 is at the ACT2-ACT4 interface and is required for signalling arginine deprivation to mTORC1

(A) The CASTOR1 D292A, E261A, D121A, and YQ118-119AA mutants are deficient in GATOR2 binding. HA-immunoprecipitates prepared from HEK293T-cells transiently expressing the indicated HA-tagged constructs were analyzed as in 1c. (B) Solvent-exposed surface view of the CASTOR1 homodimer highlighting the GATOR2-binding sites (red). (C) Cross-sectional view of the ACT2-ACT4 interface showing the positions of the critical GATOR2-binding residues relative to the bound arginine (yellow) and the β14-loop. (D) The GATOR2-binding-deficient
YQ118-119AA and D121A mutants of CASTOR1 fail to inhibit the mTORC1 pathway and render cells insensitive to arginine starvation. FLAG-immunoprecipitates were prepared and analyzed as in 1d. (E) Model of how arginine binding triggers dissociation of CASTOR1 from GATOR2. Upon arginine starvation CASTOR1 binds and inhibits GATOR2 via residues in ACT2 and ACT4, which are partially dissociated in the arginine-free conformation.

Extended Data Figure 5

Extended Data Figure 5: GATOR2-binding deficient CASTOR1 mutants still bind arginine and homodimerize
(A) The CASTOR1 YQ118-119AA, D121A, E261A and D292A mutants bind arginine in vitro. FLAG-immunoprecipitates prepared from HEK-293T cells transiently expressing indicated FLAG-tagged proteins were used in binding assays with [3H]Arginine as described in the methods. Unlabeled arginine was included as a competitor where indicated. Values are Mean ± SD for 3 technical replicates from one representative experiment. (B) The CASTOR1 YQ118-119AA, D121A, E261A and D292A mutants dimerize in cells. HA-immunoprecipitates prepared from HEK293T-cells transiently expressing CASTOR1-FLAG and HA-metap2 or the indicated HA-tagged CASTOR1 constructs were analyzed as in 1C.

coli (AKeco) and cyanobacteria (AKsyn) (Fig. 5a, Extended Data Fig. 6a) (Kotaka et al., 2006; Robin et al., 2010). AKs catalyze the first step of a metabolic pathway that synthesizes several amino acids, including lysine, and display allosteric feedback inhibition when downstream products bind to the regulatory domains (Dumas et al., 2012). Interestingly, AKeco binds lysine through pockets that bear striking resemblance to the arginine-binding pocket of CASTOR1 (Fig. 5b) (Kotaka et al., 2006). Furthermore, AKeco residues Arg305, Glu346, and Val347, which correspond to the positions of the critical GATOR2-binding residues Glu261, Tyr118, and
Gln119, respectively, directly participate in the lysine-dependent inhibition of the kinase domain in AKeco (Fig. Extended Data Fig. 6b)(Kotaka et al., 2006). Thus, the overall structure, mode of amino-acid binding, and likely allosteric mechanism of CASTOR1 all resemble those found in the regulatory domain of prokaryotic AKs.

These similarities suggest that CASTOR1 shares an evolutionary origin with prokaryotic AKs. Interestingly, AKs are found throughout bacteria, archaea, and many eukaryotic lineages, but were lost prior to the emergence of metazoa, while CASTOR1 homologues are only present in metazoa (Fig. 5c). Thus, in order to acquire arginine sensitivity in early multicellular animals, the mTORC1 pathway may have taken advantage of this more ancient, lysine-sensitive regulatory mechanism (Fig. 5d). This exploitation of a pre-existing allosteric module is analogous to the models proposed for the evolution of hormone-receptor signaling (Bridgham et al., 2006) and yeast MAP kinases (Coyle et al., 2013), and may function to enable the more rapid incorporation of novel signalling responses into existing pathways (Peisajovich et al., 2010).

Conclusion

Our results provide a mechanistic model of arginine sensing by CASTOR1, where arginine binding at the interface of ACT2 and ACT4, accompanied by the ordering of a glycine-rich loop, transmits an allosteric signal to release CASTOR1 from GATOR2 and enable downstream activation of mTORC1. In addition, our data using arginine analogues suggests that our structure may be useful for predicting compounds that can modulate arginine sensing by CASTOR1 in vivo. As deregulation of mTORC1 is common in human diseases, including cancer, the identification of novel pharmacological regulators of mTORC1 activity is of particular interest (Saxton and Sabatini, 2017; Shaw and Cantley, 2006; Zoncu et al., 2011b).
Despite these insights, several critical questions remain unanswered. Understanding how CASTOR1 inhibits GATOR2 to regulate mTORC1 activity awaits the elucidation of the molecular function of GATOR2. In addition, how mTORC1 integrates the cytosolic arginine signal from CASTOR1 with the activity of the putative lysosomal arginine sensor SLC38A9 is unclear.

Finally, our structure reveals that CASTOR1 is structurally homologous to the allosteric regulatory domain of prokaryotic Aspartate Kinases, suggesting a common evolutionary origin.
Interestingly, while homologs of GATOR1/2 and mTOR are found in many single celled eukaryotes (Saxton and Sabatini, 2017), CASTOR1 is metazoan-specific, suggesting that the coupling of cell growth with the availability of free arginine may have been specifically selected for in the evolution of early multicellular animals.

**Extended Data Figure 6**

(A) Ribbon diagram views of human CASTOR1 (this paper), AKeco (PDB ID: 2JOx) and AKsyn (PDB ID: 3L76), highlighting the different modes of dimerization. AKs can dimerize through an interlocked-ACT domain conformation (as in AKeco) or through their kinase domains (AKsyn), both of which are distinct from the side-by-side ACT-domain dimerization in CASTOR1. (B) View of AKeco depicting positions of residues R305, E346, and V347, which correspond to the positions of GATOR2-interacting residues of CASTOR1.

**Chapter 4 - Mechanism of arginine sensing by CASTOR1**
Chapter 4 - Mechanism of arginine sensing by CASTOR1
Acknowledgements

We thank all members of the Sabatini and Schwartz laboratories for helpful insights. This work is based on research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. This work has been supported by grants from NIH (R01CA103866 and AI47389) and the U.S. Department of Defense (W81XWH-07- 0448) to D.M.S. Fellowship support was provided by NIH to L.C. (F31 CA180271). D.M.S. is an investigator of the Howard Hughes Medical Institute. Coordinates and structure factors for the x-ray crystal structure of CASTOR1 have been deposited in the Protein Data Bank (PDB) with accession code 5I2C. D.M.S. is a founder, a member of the Scientific Advisory Board, a paid consultant, and a shareholder of Navitor Pharmaceuticals, which is targeting for therapeutic benefit the amino acid sensing pathway upstream of mTORC1.

Materials and Methods

Materials

Reagents were obtained from the following sources: HRP-labeled anti-rabbit secondary antibody from Santa Cruz Biotechnology; antibodies to phospho-T389 S6K1, S6K1, Mios and the FLAG epitope from Cell Signaling Technology; antibodies to the HA epitope from Bethyl laboratories; antibody to raptor from Millipore. All antibodies used have been published previously (Chantranupong et al., 2016; Saxton et al., 2016). FLAG-M2 affinity gel and amino acids from Sigma Aldrich; RPMI without leucine, arginine, or lysine from Pierce; DMEM from SAFC Biosciences; XtremeGene9 and Complete Protease Cocktail from Roche. Inactivated
Fetal Calf Serum (IFS) from Invitrogen; [³H]-labeled arginine from American Radiolabeled Chemicals.

**Protein production and purification**

Full-length, codon-optimized human CASTOR1 was N-terminally fused with a human rhinovirus 3C protease–cleavable His$_{10}$-Arg$_8$-ScSUMO tag and cloned into a PET-Duet-1 bacterial expression vector. This vector was transformed into *Escherichia coli* LOBSTR (DE3) cells (Kerafast)(Andersen et al., 2013). Cells were grown at 37 °C to 0.6 OD, then protein production was induced with 0.2 mM IPTG at 18 °C for 12–14 h. Cells were collected by centrifugation at 6,000g, resuspended in lysis buffer (50 mM potassium phosphate, pH 8.0, 500 mM NaCl, 30 mM imidazole, 3 mM β-mercaptoethanol (βME) and 1 mM PMSF) and lysed with a cell disruptor (Constant Systems). The lysate was cleared by centrifugation at 10,000g for 20 min. The soluble fraction was incubated with Ni-Sepharose 6 Fast Flow beads (GE Healthcare) for 30 min on ice. After washing of the beads with lysis buffer, the protein was eluted in 250 mM imidazole, pH 8.0, 150 mM NaCl and 3 mM βME. The Ni eluate was diluted 1:1 with 10 mM potassium phosphate, pH 8.0, 0.1 mM EDTA and 1 mM dithiothreitol (DTT), and was subjected to cation-exchange chromatography on a 5 ml SP sepharose fast flow column (GE Healthcare) with a linear NaCl gradient. The eluted CASTOR1 was then incubated with 3C protease and dialyzed overnight at 4 °C into 10 mM potassium phosphate, pH 8.0, 150 mM NaCl, 0.1 mM EDTA and 1 mM DTT, followed by a second cation-exchange chromatography run on an SP sepharose fast flow column (GE Healthcare) with a linear NaCl gradient. The protein was further purified via size-exclusion chromatography on a Superdex S200 16/60 column (GE Healthcare) equilibrated in running buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1 mM EDTA and 1 mM DTT). Selenomethionine (SeMet)-derivatized CASTOR1 was prepared as described previously (Brohawn et al., 2008) and purified as the native version, except that the reducing-agent concentration (βME and DTT) was 5 mM in all buffers.
Crystallization

Purified CASTOR1 was concentrated to 6 mg/ml and incubated in 2 mM arginine for >1 hour prior to setting crystal trays. Crystals were grown at 18 °C by hanging-drop vapor diffusion with 1 µl of protein at 6 mg/ml mixed with an equal volume of reservoir solution containing 0.1 M sodium acetate pH 5.0, 0.25 M ammonium acetate, and 22.5% PEG 3350. Selenomethionine-derivatized CASTOR1 crystallized in 0.1 M BIS-TRIS pH 5.6, 0.25 M ammonium acetate, and 22.5% PEG3350. Crystals were cryoprotected in mother liquor supplemented with 20% (v/v) ethylene glycol.

Data collection and structure determination

Data collection was performed at the Advanced Photon Source end station 24-IDC at Argonne National Lab, at 100 K. All data-processing steps were carried out with programs provided through SBgrid (Morin et al., 2013). Data reduction was performed with HKL2000 (Otwinowski and Minor, 1997). A complete native dataset was collected to 1.8 Å (at wavelength 0.9792 Å) and a complete SeMet dataset, at the selenium peak wavelength (0.9792 Å), was collected to 2.2 Å. The phase problem was solved using single-wavelength anomalous dispersion (SAD) and selenium positions were determined in HYSS, run as part of the PHENIX AutoSol program (Adams et al., 2010), for the SeMet dataset (space group P2₁, 4 molecules per asymmetric unit). An interpretable 2.2 Å experimental electron density map was obtained, and manual model building was carried out in Coot (Emsley et al., 2010). Subsequent refinement was carried out with the superior 1.8 Å native data set using phenix.refine to a final R_work/R_free of 17.2%/20.4%. Ramachandran statistics in the final model are 99% favored, 1% allowed, and 0% outlier.
Structural analysis

Protein-protein and protein-ligand interfaces were analyzed using PDBePISA (Krissinel and Henrick, 2007). NCBI’s Vector Alignment Search Tool (VAST) (Gibrat et al., 1996) was used to identify structurally related proteins in the PDB. The multiple sequence alignment (MSA) was generated in Jalview (Waterhouse et al., 2009) with the T-Coffee alignment algorithm (Notredame et al., 2000). Sequences of CASTOR1 homologs were obtained via NCBI BLAST searches (Altschul et al., 1990). All structure figures were made in PyMol (Schrodinger, 2010).

Cell lysis and immunoprecipitation

Cells were rinsed one time with ice-cold PBS and immediately lysed with Triton lysis buffer (1% Triton, 10 mM β-glycerol phosphate, 10 mM pyrophosphate, 40 mM Heps pH 7.4, 2.5 mM MgCl2 and 1 tablet of EDTA-free protease inhibitor [Roche] (per 25 ml buffer). The cell lysates were cleared by centrifugation at 13,000 rpm at 4°C in a microcentrifuge for 10 minutes. For anti-HA-immunoprecipitations, the magnetic anti-HA beads (Pierce) were washed 3 times with lysis buffer. 30 μl of a 50/50 slurry of the affinity gel was then added to clarified cell lysates and incubated with rotation for 1 hour at 4°C. Following immunoprecipitation, the beads were washed 4 times with lysis buffer containing 500 mM NaCl. Immunoprecipitated proteins were denatured by the addition of 50 μl of sample buffer and boiling for 5 minutes as described (Kim et al., 2002), resolved by 8%-16% SDS-PAGE, and analyzed by immunoblotting.

For co-transfection experiments in HEK-293T cells, 2.5 million cells were plated in 10 cm culture dishes. Twenty-four hours later, cells were transfected using the polyethylenimine method (Boussif et al., 1995) with the pRK5-based cDNA expression plasmids indicated in the following amounts: 50 ng CASTOR1-HA (wild-type or mutant), 50 ng CASTOR1-FLAG, 1 μg HA-metap2, or 2 ng S6K. For in vitro dissociation experiments, 50 ng of wild-type CASTOR1-HA was transfected into HEK-293T cells. The total amount of plasmid DNA in each transfection was
normalized to 5 μg with empty pRK5. 36-48 hours after transfection, cells were lysed as described above.

For experiments that required amino acid starvation or restimulation, cells were treated as previously described (Tsun et al., 2013). Briefly, cells were incubated in arginine free RPMI for 50 minutes and then re-stimulated with 500 μM arginine for 10 minutes.

Arginine binding assay

5 million HEK-293T cells were plated in a 15 cm plate four days prior to the experiment. Twenty-four hours after plating, the cells were transfected via the polyethylenimine method with the pRK5-based cDNA expression plasmids indicated in the figures in the following amounts: 15 μg FLAG-Rap2A, 500 ng FLAG-CASTOR1 (wild-type or mutant). The total amount of plasmid DNA in each transfection was normalized to 15 μg total DNA with empty-PRK5. Forty-eight hours after transfection cells were lysed as previously described. If multiple samples of the same type were represented in the experiment, the cell lysates were combined, mixed, and evenly distributed amongst the relevant tubes.

Anti-FLAG beads were blocked by rotating in 1 μg/μl bovine serum albumin (BSA) for 20 minutes at 4 °C, then washed twice in lysis buffer and resuspended in an equal volume of lysis buffer. 30 μl of bead slurry was added to each of the clarified cell lysates and incubated as previously described. Post-IP, the beads were washed as previously and incubated for one hour on ice in cytosolic buffer (0.1% Triton, 40 mM HEPES pH 7.4, 10 mM NaCl, 150 mM KCl, 2.5 mM MgCl2) with the appropriate amount of [3H]-labeled arginine and cold argine. At the end of one hour, the beads were aspirated dry and rapidly washed three times with cytosolic buffer. The beads were aspirated dry again and re-suspended in 85 μl of cytosolic buffer. Each sample was mixed well and three 10 μl aliquots were separately quantified using a TriCarb scintillation counter (PerkinElmer). This process was repeated in pairs for each sample, to ensure similar incubation and wash times for all samples analyzed across different experiments.
**In vitro CASTOR1-GATOR2 dissociation assay with arginine analogues**

HEK-293T were transfected with HA-CASTOR1 constructs as described above. 48 hours after transfection, cells were starved for all amino acids for 50 minutes, lysed and subjected to anti-FLAG immunoprecipitation as described previously. The CASTOR1-GATOR2 complexes immobilized on the HA beads were washed twice in lysis buffer with 500 mM NaCl, then incubated for 20 minutes in 1 mL of cytosolic buffer with 400 μM of the indicated compound. The amount of GATOR2 and CASTOR1 that remained bound was assayed by SDS-PAGE and immunoblotting as described previously.

**Cell lines and tissue culture**

HEK-293T cells were cultured in DMEM 10% IFS supplemented with 2 mM glutamine. All cell lines were maintained at 37°C and 5% CO2.

**Statistical analysis**

For the arginine-binding assays, two-tailed t tests were used for comparison between two groups. All comparisons were two-sided, and P values of less than 0.005 were considered to indicate statistical significance. The data meet the assumptions of the test and the variance is similar between groups that are being statistically compared.

**References**


Chapter 4 - Mechanism of arginine sensing by CASTOR1


Chapter 4 - Mechanism of arginine sensing by CASTOR1
component of the lysosomal amino acid sensing machinery that controls mTORC1. Nature 519, 477-481.


Chapter 5

Summary and future directions

Parts of this chapter were first published as:

Section I. Summary

The results presented in this thesis have revealed the structural and biochemical mechanisms underlying cytosolic leucine and arginine sensing by the mTORC1 pathway. Using a combination of approaches including X-Ray crystallography, biochemistry, and cell biology, we have successfully elucidated how the leucine sensor Sestrin2 and the arginine sensor CASTOR1 recognize their respective ligands, and also established the likely molecular mechanisms through which ligand binding allosterically disrupts the interaction between these sensors and GATOR2, a key activator of mTORC1. Furthermore, our structures of Sestrin2 and CASTOR1 have also provided substantial insights into the evolutionary origins of amino acid sensing in mammals, showing that both Sestrin2 and CASTOR1 likely evolved from ancient bacterial metabolic enzymes. Together, these findings represent an important advance in our understanding of how mammals coordinate nutrient availability with metabolic processes, and will hopefully enable the improved therapeutic targeting of the mTORC1 pathway in the clinic.

In this chapter, I outline the major future directions and remaining questions on the topics of Sestrin2, CASTOR1, and nutrient sensing by the mTORC1 pathway more generally.
Section II. Leucine sensing by Sestrin2

A. Structure of apo- or GATOR2-bound Sestrin2

As discussed in Chapter 2, in our efforts to structurally characterize the leucine sensor Sestrin2, we were only able to obtain high quality crystals in the presence of saturating leucine concentrations (Saxton et al., 2016c). Indeed as discussed in chapter 3, even the lower quality crystals of Sestrin2 that we obtained without adding exogenous leucine contained leucine-bound Sestrin2 (Saxton et al., 2016b). Thus, the conformation of Sestrin2 in the absence of leucine, and the exact conformational change that leucine induces in Sestrin2 is unclear, and exactly how leucine triggers the dissociation of Sestrin2 from GATOR2 remains to be seen.

Our structure of leucine-bound Sestrin2 provides two plausible models to explain how leucine binding by Sestrin2 could influence the interaction with GATOR2, primarily due to our identification of a highly conserved GATOR2 binding site on the surface of Sestrin2 in close proximity to the leucine-binding pocket (Saxton et al., 2016c). The first model proposes that the "lid," a loop that helps bury the leucine in the leucine-binding pocket, must open in the absence of leucine, triggering a conformational change that alters the position of the GATOR2-binding residues, which are connected to the lid by a short alpha helix. The second model proposes that the release of leucine alters the conformation of the mostly disordered "linker" region of Sestrin2, potentially ordering it in such a way that it facilitates GATOR2 binding in the absence of leucine.

While both of these models are plausible and supported by our structural data, the definitive answer will require a structure of either apo-Sestrin2 or GATOR2-bound Sestrin2. We have made several attempts to obtain one or both of these structures by attempting crystallization of a) leucine-binding deficient Sestrin2 mutants, b) the Sestrin2 paralog with reduced affinity for leucine (Sestrin3), c) Sestrin2 in complex with the minimal region of GATOR2 that is sufficient to bind Sestrin2, and d) Sestrin2 in complex with single-domain camellid antibodies (nanobodies). Despite these efforts however, the structure of Sestrin2
without leucine remains elusive, and an important missing component of our understanding of the mechanism of leucine sensing by the mTORC1 pathway.

B. Role of Sestrin2 phosphorylation

Our early experiments with Sestrin2 showed a striking change in the gel mobility of Sestrin2 depending on the concentration of leucine in the cell culture media (Chantranupong et al., 2014; Wolfson et al., 2016). Subsequent unpublished data from the Sabatini lab has revealed that this is due to multi-site phosphorylation of Sestrin2 that occurs in the absence of leucine. More specifically, we found that this phosphorylation occurs only following GATOR2 binding and is actually independent of leucine levels per se, as Sestrin2 phosphorylation can be induced by GATOR2 overexpression even in the presence of leucine, and no phosphorylation is observed on the GATOR2-binding deficient mutant of Sestrin2 even under leucine starvation conditions (Saxton et al., 2016c; Wolfson et al., 2016). Our unpublished data along with those published by the Kimball group suggests that the phosphorylation of Sestrin2 enhances its affinity for GATOR2, thereby raising the threshold concentration of leucine required to reactivate mTORC1 signaling once it has been inhibited by leucine starvation (Kimball et al., 2016). This study identified three phosphorylation sites on Sestrin2, all located in the disordered linker region of our structure. One possibility therefore is that phosphorylation induces the ordering of this linker region, which can then participate in GATOR2-binding.

Going forward, it will be useful to understand the significance of Sestrin2 phosphorylation and its role in leucine sensing. Specifically, it will be informative to address the utility of raising the threshold for leucine sensing once the pathway is in the starved state. In addition, the identity of the kinase responsible for this event and how it recognize and phosphorylate only the GATOR2-bound population of Sestrin2 is also unknown. Addressing these questions will be important for gaining a more complete picture of how leucine sensing by Sestrin2 is altered depending on the existing nutritional state of the cell.
C. Kinetics of the Sestrin2-leucine interaction

In a variety of human cell lines, mTORC1 signaling is inhibited within minutes of switching into media lacking leucine, suggesting that the off-rate of leucine from Sestrin2 is also on the order of minutes. This is consistent with unpublished data generated in the Sabatini lab using radioactively labeled leucine showing that the half-life of the Sestrin2-leucine interaction is approximately 30 minutes at room temperature. However, as discussed in Chapter 3, Sestrin2 purified from bacteria will crystallize bound to leucine even without the addition of exogenous leucine, suggesting a non-trivial portion of Sestrin2 remains bound to leucine over the 24-48 hours between bacterial cell lysis and crystal tray setup (Saxton et al., 2016b). How this observation can be compatible with the established off rate of leucine from Sestrin2 remain unclear.

One possibility is that Sestrin2 in mammalian cells (or purified from mammalian cells) contains a modification or binding partner that enhances the off-rate of leucine, which is absent from Sestrin2 purified from E. coli. Such a “leucine exchange factor” would provide a way to regulate the response time of mTORC1 signaling following changes in leucine levels. Alternatively, the phosphorylation of Sestrin2 described above could also influence the kinetics of the interaction with leucine. Ultimately, understanding whether and how the kinetics of the leucine-Sestrin2 interaction is regulated will be important to address going forward.

D. Sestrin paralogs

In addition to Sestrin2, human cells also express two additional Sestrin genes, known as Sestrin1 and Sestrin3, which also bind to GATOR2 and inhibit mTORC1 signaling when over-expressed in cells (Chantranupong et al., 2014). Interestingly, despite sharing a high degree of sequence similarity, and complete conservation of residues in the leucine-binding pocket, the three Sestrin paralogs exhibit different affinities for leucine. Specifically, Sestrin1 binds leucine with a slightly higher affinity relative to Sestrin2, while leucine binding by Sestrin3 is
undetectable (Wolfson et al., 2016). The fact that these differences in affinity could not be predicted from our structure of leucine-bound Sestrin2 suggests that the structures of Sestrin1 and Sestrin3 would be insightful and contribute to our understanding of the molecular basis for leucine sensing by the mTORC1 pathway. Furthermore, the physiological role of Sestrin3, which functions as a constitutive negative regulator of mTORC1 signaling remains to be seen.

E. Evolutionary origins of Sestrin2

Another intriguing question surrounding Sestrin2 is that of how it evolved and became incorporated into the mTORC1 pathway. As discussed in Chapter 2, Sestrin2 shares a striking degree of structural homology with the bacterial enzyme AhpD, which functions as a peroxiredoxin reductase to protect bacterial cells from oxidative stress (Budanov et al., 2004; Saxton et al., 2016c). The high level of structural similarity, including in the region that forms the leucine-binding pocket of Sestrin2, strongly suggests a shared evolutionary origin between these two proteins. However, AhpD does not bind amino acids and the mTOR pathway does not exist in bacteria, raising the question of how AhpD became a leucine sensor upstream of mTORC1. One plausible model is that, as a stress responsive gene in primitive unicellular eukaryotes, an AhpD-like protein acquired the ability to bind GATOR2 to inhibit the mTORC1 pathway under conditions of cellular stress. This ancestral Sestrin/AhpD protein could then later have acquired leucine-sensing capabilities through subtle changes in it's the region that corresponds to the leucine-binding pocket (Chapter 2). Direct support for this hypothesis awaits the discovery of intermediate forms of AhpD/Sestrin that can suppress mTORC1 signaling but are not responsive to leucine. Interestingly, S. cerevisiae, which have a functioning TORC1 pathway and GATOR complex homologs, lack an identifiable Sestrin homolog (Saxton and Sabatini, 2017). Indeed, Sestrin homologs have thus far only been observed in metazoans, though a systematic search for potential “missing links” in the evolutionary chain from AhpD to Sestrin in unicellular eukaryotes remains to be done.
F. Role of Sestrin2 in mammalian physiology

As discussed in Chapter 1, the mTORC1 pathway plays a critical role in a number of tissues and physiological processes, including muscle growth, glucose homeostasis, immune function, and brain function (Saxton and Sabatini, 2017). As a key regulator of mTORC1 signaling in cells, Sestrin2 likely also plays important roles in regulating these various processes in response to changing dietary leucine levels. Consistent with this, deletion of all three Sestrin paralogs in mice leads to neo-natal lethality due to an inability to survive the neonatal fasting period, the same phenotype observed in RagA\textsuperscript{GTP} mice, in which mTORC1 is also hyperactivated in insensitive to nutrient starvation (Efeyan et al., 2013; Peng et al., 2014). However, the exact role of Sestrin2 in specific tissues, such as the muscle or liver where the importance mTORC1 signaling in response to nutrients is well established, remains to be seen (Anthony et al., 2000; Bentzinger et al., 2008; Sengupta et al., 2010). These studies will likely require tissue specific knockouts of Sestrin genes. In addition, it will likely be informative to look at the physiological effects of expressing Sestrin2 mutants that have reduced or abolished leucine-binding capacities in various tissues. For example by generating tissue specific knock-in mice expressing Sestrin2 E451A or W444L, described in Chapter 2 (Saxton et al., 2016c). These experiments are currently underway in the Sabatini Lab and will hopefully yield insights into the physiological role of Sestrin2 and of leucine sensing more generally.

Section III. Arginine sensing by CASTOR1

A. Structure of GATOR2-bound CASTOR1

Although our structure of arginine-bound CASTOR1 provides a likely mechanism for how arginine binding triggers the dissociation of CASTOR1 from GATOR2 (Saxton et al., 2016a), a complete picture of the arginine sensing mechanism will ultimately require the structure of CASTOR1 without arginine. However, as the true “apo”-form of CASTOR1 likely does not ever exist in cells, the structure of CASTOR1 bound to GATOR2, either by x-ray crystallography or
by cryo-EM, will likely be even more insightful and enable us to understand both the exact conformational change induced by arginine as well as how CASTOR1 regulates GATOR2 function. Our preliminary unpublished data suggests that CASTOR1 binds primarily to the Mios subunit of GATOR2, although both WDR24 and Seh1L are also required for maximal binding.

**B. Role of CASTOR2 and the CASTOR1/2 heterodimer**

In addition to CASTOR1, humans also express the closely related CASTOR2 (Chantranupong et al., 2016). Interestingly however, unlike CASTOR1, CASTOR2 does not bind arginine and appears to act as a constitutive inhibitor of mTORC1 signaling, in an analogous manner to Sestrin3. The presence of these two amino acid sensor paralogs that lack the ability to bind amino acids is therefore an emerging trend, although their functional role remains mysterious. Unlike the Sestrins however, CASTOR1 and 2 are able to both homo- and heterodimerize, indicating more complex sensor arrangements in which only one subunit of a CASTOR1-2 heterodimer can bind arginine. In our efforts to structurally characterize CASTOR1/2 we did successfully purify and crystallize the CASTOR1/2 heterodimer, however we were unable to solve this structure due to the severe twinning and pseudosymmetry observed in the X-ray diffraction data (R.A.S, T.U.S, D.M.S, unpublished data). However, although the functional significance of this heterodimeric species is unclear, our biochemical data suggest that it would be less sensitive to arginine and therefore raise the threshold arginine concentration required to dissociate CASTOR from GATOR2 and activate mTORC1 (Chantranupong et al., 2016; Saxton et al., 2016a). Such a setup would enable a fine-tuning of arginine sensitivity through the modulation of the relative expression levels of CASTOR1 and CASTOR2. Whether this heterodimeric complex is functionally relevant in cells or *in vivo* however remains to be seen.

**C. Integration of cytosolic and lysosomal arginine signals from CASTOR1 and SLC38A9**

Chapter 5 – Summary and future directions
In addition to CASTOR1, mammalian cells are also able to sense arginine levels through the lysosomal amino acid transporter SLC38A9 (Rebsamen et al., 2015; Wang et al., 2015). The current model suggests that arginine inside the lysosome can bind to SLC38A9, modulating the interaction between the N-terminus of SLC38A9 and the Rag-Ragulator complex to activate mTORC1. In addition, more recent data suggests that SLC38A9 is responsible for exporting lysosomal amino acids to the cytosol in an arginine-gated manner, thus making it both a key activator of mTORC1 signaling and a regulator of lysosomal amino acid content (Wyant et al., 2017).

These results raise the question of why mammalian cells need two independent arginine sensors upstream of mTORC1, and in which context one is more important than the other. Mechanistically, CASTOR1 is required to inhibit mTORC1 under arginine starvation, whereas SLC38A9 is specifically required for the reactivation of mTORC1 after starvation followed by arginine restimulation (Chantranupong et al., 2016; Wang et al., 2015). Consistent with this, CASTOR1 knock-down using RNAi in the background of SLC38A9 knock out yields a phenotype in which mTORC1 signaling is largely resistant to arginine starvation, but also reduced upon restimulation with arginine (Chantranupong et al., 2016). In addition, while CASTOR1 detects only cytosolic arginine levels, SLC38A9 responds only to arginine within the lysosomal lumen. Together, these results indicate that CASTOR1 and SLC38A9 serve distinct and complementary, rather than redundant, functions in the mTORC1 pathway. However, it is unclear under what conditions arginine levels in the cytosol and lysosome would be differentially impacted, and therefore what benefit is provided by sensing arginine in both compartments.

D. Role of CASTOR1 in mammalian physiology

Likely the most important future direction relating to CASTOR1 is to understand its role in mammalian physiology. Specifically, it will be important to learn in what tissues and cell types arginine sensing by CASTOR1 is important, and whether defective arginine sensing by
CASTOR1 leads to metabolic defects at the organismal level. Such studies will hopefully also enable the pharmacological targeting of CASTOR1, which may be useful in a number of disease contexts. Gene expression data suggests that CASTOR1 is lowly expressed in most tissues in humans, while CASTOR2 is particularly abundant in muscles. The tissue where the regulation of mTORC1 by arginine is known to be particularly important is the muscle, where arginine promotes muscle growth and hypertrophy (Yao et al., 2008). However, the critical role of mTORC1 in the liver and brain suggest that studying CASTOR1 function in these tissues, through the use of tissue-specific knockout mice, will also be extremely informative.

Section IV. Integration of amino acid signals by the GATOR2 complex

A. Molecular function of the GATOR2 complex

The molecular characterization of Sestrin2 and CASTOR1 described in this thesis indicate that the GATOR2 complex plays a central role in integrating the signals from cytosolic leucine and arginine and transmitting their presence to the mTORC1 pathway. Specifically, both Sestrin2 and CASTOR1 bind to and inhibit GATOR2 function in the absence of their respective amino acids, and dissociate from GATOR2 upon ligand binding. Therefore, understanding the molecular function of GATOR2 and how these sensors regulate it likely holds the key for full understanding how the amino acid signals are transmitted to mTORC1.

Mechanistically, we have found that Sestrin2 and CASTOR1 bind to GATOR2 non-competitively, and therefore likely inhibit it through distinct mechanisms. In addition, starvation of either leucine or arginine alone, or over expression of Sestrin2 or CASTOR1 alone, is sufficient to inhibit mTORC1 signaling, suggesting that binding to either site is sufficient to inhibit GATOR2 function. Notably, neither Sestrin2 nor CASTOR1 disrupt the interaction between GATOR2 and GATOR1, suggesting that they may act by blocking the catalytic activity of GATOR2 rather than its interaction with a substrate or binding partner (Chantranupong et al., 2016; Chantranupong et al., 2014). Sequence analysis has revealed that three components of
GATOR2 contain RING domains, which in all known cases comprise the catalytic domain of E3 ubiquitin ligases, catalyzing the transfer of ubiquitin to substrate proteins. Thus, we have hypothesized that GATOR2 acts as a ubiquitin ligase whose ubiquitin transfer activity is blocked by both CASTOR1 and Sestrin2. However, the substrate targeted by GATOR2 has been elusive, making it impossible to test this hypothesis. Genetic and proteomic evidence point to GATOR1 as the most likely substrate (Bar-Peled et al., 2013), however no ubiquitin modifications have been detected on GATOR1 to date. Identifying the true substrate of GATOR2 and the mechanism through which Sestrin2 and CASTOR1 regulate this ubiquitylation event will therefore be a central focus for the field going forward.

**B. Structural characterization of the GATOR2 complex**

Another approach that will be essential for understanding the function of GATOR2 and its regulation by the amino acid sensors is the structural characterization of the complex, ideally in the presence and absence Sestrin2 and CASTOR1. Given the large size of this complex (~400-500 kD), this will likely require the use of cryo-Electron Microscopy (cryo-EM). In addition, crystallization of Sestrin2 and CASTOR1 in complex with the minimal domains of GATOR2 necessary for their interaction would likely also yield significant insight both into the conformational changes induced in these sensors upon ligand binding and the mechanism through which they regulate GATOR2 activity. We have successfully mapped the regions of GATOR2 where Sestrin2 and CASTOR1 bind to high resolution (R.A.S, T.U.S, D.M.S, unpublished data), and are working towards cocrystallization of these complexes.

**C. Identification and characterization of additional nutrient sensors upstream of GATOR2**

In addition to leucine and arginine, several additional nutrients are thought to regulate mTORC1 signaling as well, including glucose, lysine, and methionine. Indeed, recent work from the Sabatini lab has identified SAMTOR, an S-adenosyl methionine sensor for the mTORC1
pathway (Gu et al., 2017), which is required to inhibit mTORC1 signaling under methionine starvation. Interestingly, unlike Sestrin2 and CASTOR1, SAMTOR does not bind to GATOR2 but rather to the GATOR1-KICSTOR complex, suggesting an alternative mechanism of regulation. An important future direction will be the structural characterization of SAMTOR, both bound to SAM and bound to GATOR1-KICSTOR, in order to characterize both its mechanism of sensing and regulation of GATOR2/GATOR1 function.

Along with methionine, glucose has also been shown to regulate mTORC1 activity upstream of the Rag GTPases (Efeyan et al., 2013; Kalender et al., 2010), although the glucose sensor remains elusive. In addition, preliminary work implicates the essential amino acid lysine in regulating mTORC1 activation upstream of the Rags, though the mechanism of sensing is unclear. Given the sensing mechanisms for Sestrin2, CASTOR1 and SAMTOR, it is tempting to hypothesize that the sensors for glucose, lysine, and other unknown nutrients may also act by modulating GATOR2 function.

Section V. Conclusion

It is now clear that the mTOR pathway plays a central role in sensing environmental conditions and regulating nearly all aspects of metabolism at both the cellular and organismal level. In just the last several years, many new insights into both mTOR function and regulation have been elucidated, and extensive genetic and pharmacological studies in mice have enhanced our understanding of how mTOR dysfunction contributes to disease.

While many inputs to mTORC1 have now been identified and their mechanisms of sensing characterized, an integrated understanding of the relative importance of these signals and the contexts in which they are important remains largely unclear. For example, it remains mysterious which inputs to the TSC complex are dominant and how this depends on the physiological context. Similar questions exist for nutrient sensing by the Rag GTPases, specifically regarding the purpose of sensing of both lysosomal and cytosolic amino acids, as
well as the tissues in which the nutrient sensors such as Sestrin2, CASTOR1, and SLC38A9 are most important. Such insights will likely require both deeper biochemical studies of these complexes \textit{in vitro} as well as improved mouse models that enable more nuanced perturbation and monitoring of these sensors \textit{in vivo}.

The major focus of mTOR research going forward however will be to address whether these molecular insights can improve the therapeutic targeting of mTOR in the clinic. Although rapalogs and catalytic mTOR inhibitors have been successful in the context of immunosuppression and a small subset of cancer types, clear limitations have arisen which limit their utility. Specifically, given the critical functions of mTOR in most human tissues, complete catalytic inhibition causes severe dose-limiting toxicities, while rapalogs also suffer from the drawbacks associated with lack of tissue specificity and unwanted disruption of mTORC2. Future work should focus on the development of mTOR-targeting therapeutics outside of these two modalities, such as truly mTORC1-specific inhibitors for use in diabetes, neurodegeneration, and life-span extension, or tissue-specific mTORC1 agonists for use in muscle wasting diseases and immunotherapy. Such approaches will likely require going beyond targeting mTOR directly to instead developing compounds that modulate tissue-specific receptors or signaling molecules upstream of mTOR, such as the amino acid sensors Sestrin2 and CASTOR1, which, as demonstrated in this Thesis, contain small-molecule binding pockets and specifically regulate mTORC1. Ultimately, such insights may enable the rational targeting of mTOR signaling to unlock the full therapeutic potential of this remarkable pathway.

References


Peng, M., Yin, N., and Li, M.O. (2014). Sestrins function as guanine nucleotide dissociation inhibitors for Rag GTPases to control mTORC1 signaling. Cell 159, 122-133.


