Aus dem Lehrstuhl für Physiologische Chemie,
Lehrstuhl der Ludwig-Maximilians-Universität München

Dissertation

zum Erwerb des Doctor of Philosophy (Ph.D.)

an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Dissecting the nutrient-driven role of Creb3L transcription factor family to coordinate ER function

vorgelegt von:
Haris Ahmad Khan

aus:
Jhang

Jahr:
2023

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Erstes Gutachten von:	Prof. Dr. Andreas G. Ladurner
Zweites Gutachten von:	Prof. Dr. Peter B. Becker
Drittes Gutachten von:	Prof. Dr. Jürgen Bernhagen
Viertes Gutachten von:	Prof. Dr. Klaus Parhofer

Dekan:	**Prof. Dr. med. Thomas Gudermann**

Datum der Verteidigung:

21.09.2023

Acknowledgements

First and foremost, I would like to thank Dr. Carla Margulies for giving me the opportunity to work on my PhD project with her. It has been a great learning experience for me due to her immense guidance and support.

I am especially grateful to Prof. Dr. Andreas Ladurner who as department head created an atmosphere where great science can take place while also enabling each individual to develop personally.

I also wish to acknowledge Prof. Dr. Andreas Ladurner and Dr. Marcus Buschbeck for initiating the Marie-Curie funded "ChroMe" training network. Being part of such a network allowed for personal and professional growth that I am forever grateful for.

I wish to acknowledge and thank all the management, group leaders and fellow PhD candidates that were part of the "ChroMe" training network for the amazing journey we have had during the three years.

I would like to especially thank Dr. Catherine Postic and Dr. Ana Conesa for hosting me in their labs and enabling me to develop further professionally.

I would also like to thank Iva Guberovic, Salva Casani, Paula Ortega-Prieto and Magdalini Serefidou for always being there. You made this journey a beautiful one.

I also would like to thank all the members of the Ladurner department for discussions and for providing a lot of help and support throughout my time in the department.

I would like to thank the current and previous members of the Margulies group especially Ming Toh, Sonja Muehlberger, Mario Torralba-Saez, Rory Beresford, Hui-Lan Huang, Teressa Burrell and Sandra Esser for all their support.

I wish to thank Christine Werner and Dr. Anton Eberharter for all the management-related help and making me feel greatly welcomed in the department and in Germany.

I would like to thank my family, my parents and my friends here in Munich and back home for being a constant support and motivating me in good and bad times.

Lastly, I would like to thank my wife and my best friend, Syeda Sana Ahmad, without whom this journey would have been impossible. Thank you for always being there.

1. Table of Contents

2. Summary

2.1. Summary (in English)

Incoming nutrients in to an organism's body results in a range of physiological responses starting from digestion to nutrient storage. The endoplasmic reticulum (ER) plays a key role in regulating nutritional homeostasis via intercellular communication and promoting the distribution of metabolites. Little is known about how nutrients adapt ER function to impact physiology. In this thesis, I present the findings that the conserved CrebA/Creb3L transcription factors are key regulators of the ER secretory capacity in response to feeding, a response that likely mediates nutritional homeostasis of an organism.

First, using *Drosophila melanogaster* as a model organism, I identify changes in gene expression of the machinery responsible for the sorting of proteins to the ER and between the ER and Golgi upon fasting and refeeding. Interestingly, proteomic assays reveal that feeding impacts protein secretion into fly hemolymph, the equivalent of blood in flies. Second, I identify the *Drosophila* Creb3L-family transcription factor CrebA as a vital regulator of the ER secretory pathway genes upon nutrient intake. CrebA activity is quickly and dynamically switched on upon feeding, driving the expression of the ER protein sorting machinery by releasing the polymerase for transcription elongation. Critically, transient overexpression of CrebA suppresses feeding thus regulating animal behavior upon food ingestion possibly via changes in secretion of satiety hormones. Similarly, the mouse homologs Creb3L1 and Creb3L2 are also upregulated upon feeding and induce the transcriptional activation of ER protein sorting machinery genes.

In summary, this thesis reveals a conserved transcriptional switch, that prepares cell's secretory capacity upon nutritional influx. I propose that CrebA coordinates a negative feedback loop that inhibits feeding via regulating ER function and secretion. These findings are promising advances in our understanding of ER function regulation and how it controls the secretion of proteins upon metabolic stress. Further studies built on this understanding can be crucial towards our fight against ER- and secretory defects- related pathologies.

2.2. Zusammenfassung

Die Aufnahme von Nährstoffen in den Körper eines Organismus führt zu einer Reihe physiologischer Reaktionen, angefangen von der Verdauung bis zur Nährstoffspeicherung. Das endoplasmatische Retikulum (ER) spielt eine Schlüsselrolle bei der Regulierung der Ernährungshomöostase über die interzelluläre Kommunikation und die Förderung der Verteilung von Metaboliten. Es ist wenig darüber bekannt, wie Nährstoffe die ER-Funktion anpassen, um die Physiologie zu beeinflussen. In dieser Dissertation präsentiere ich die Ergebnisse, dass die konservierten CrebA/Creb3L-Transkriptionsfaktoren Schlüsselregulatoren der ER-Sekretionskapazität als Reaktion auf die Nahrungsaufnahme sind, eine Reaktion, die wahrscheinlich die Ernährungshomöostase eines Organismus vermittelt.

Zunächst identifiziere ich unter Verwendung von Drosophila melanogaster als Modellorganismus robuste Veränderungen in der Genexpression der Maschinerie, die für die Sortierung von Proteinen zum ER und zwischen dem ER und Golgi beim Fasten und Wiederernähren verantwortlich ist. Interessanterweise zeigen Proteomik-Assays, dass die Fütterung die Proteinsekretion in die Fliegen-Hämolymphe beeinflusst, das Äquivalent von Blut in Fliegen. Zweitens identifiziere ich den Transkriptionsfaktor CrebA der Drosophila-Creb3L-Familie als den wichtigsten metabolischen Regulator dieser Gene des ER-Sekretionswegs als Reaktion auf die Nährstoffaufnahme. Die CrebA-Aktivität wird bei der Fütterung schnell und vorübergehend eingeschaltet, wodurch die Expression der ER-Protein-Sortiermaschinerie vorangetrieben wird, indem die Polymerase für die Transkriptionselongation freigesetzt wird. Entscheidend ist, dass die vorübergehende Überexpression von CrebA die Nahrungsaufnahme unterdrückt und somit das Tierverhalten bei der Nahrungsaufnahme reguliert, möglicherweise über Änderungen in der Sekretion von Sättigungshormonen. Darüber hinaus werden die Maus-Homologe Creb3L1 und Creb3L2 nach der Fütterung ebenfalls hochreguliert und treiben die transkriptionelle Aktivierung von Genen der ER-Proteinsortierungsmaschinerie an.

Zusammenfassend zeigt diese Arbeit einen konservierten Transkriptionsschalter, der als Reaktion auf die Nahrungsaufnahme eingeschaltet wird. Ich schlage vor, dass CrebA eine negative Rückkopplungsschleife orchestriert, die das Sättigungsgefühl fördert, indem es die ER-Funktion und -Sekretion reguliert. Diese Ergebnisse sind vielversprechende Fortschritte in unserem Verständnis der ER-Funktionsregulation

und wie sie die Sekretion von Proteinen bei metabolischem Stress steuert. Weitere Studien, die auf diesem Verständnis aufbauen, können für unseren Kampf gegen ER- und sekretorische Defekte im Zusammenhang mit Pathologien von entscheidender Bedeutung sein.

3. Introduction

All life requires energy to perform tasks necessary for survival. Ingestion of food containing nutrients that provide this energy is thus crucial for a healthy life. Nutrients are macromolecules such as carbohydrates, amino acids, lipids, vitamins and metals that can be utilized to produce energy. Organisms have evolved multiple mechanisms to sense and absorb nutrients that serve as energy sources. Most organisms do this through the use of taste receptors at the oral cavity or along the digestive tract that are connected to the central nervous system. These mechanisms 1) allow the control of taste-dependent and taste-independent ingestion of nutrients based on the need of the organism and 2) ensure that appropriate digestive and hormonal responses are elicited following the ingestion of nutrients. Indeed, Jean Mayer proposed over six decades ago that animals can, through neuronal detection of blood glucose in the brain, sense blood glucose levels (Mayer, 1955). Later, action of specific neurons was also identified to promote food intake. However, in order to maintain long term metabolic homeostasis, cells must also closely regulate gene expression. Transcription factors, through the action of their DNA-binding domain, can bind to the regulatory regions of genes to turn them on or off. These factors thus lie at an important intersection where, upon nutritional intake, they can be regulated to influence gene expression of multiple genes required for a metabolic response. Metabolically active organs, such as the liver, house multiple such transcription factors to orchestrate metabolic pathways in response to organismal-level metabolic changes.

In this thesis, I present an evolutionarily conserved mechanism that dynamically switches the cellular program altering Endoplasmic Reticulum (ER) function in response to nutrients. I show that the mammalian Creb3L transcription factor family and its *Drosophila* counterpart, CrebA are directly regulated by nutrients which in turn regulate ER activity, impacting the secretion of hormones and other, behavior-relevant secreted signals. Using genome-wide approaches such as RNA sequencing (RNA-seq) and Chromatin Immunoprecipitation sequencing (ChIP-seq), I show that Creb3L transcription factors directly bind and regulate the transcription of the ER protein sorting machinery genes via the release of RNA polymerase from a paused to an elongating state. *Drosophila melanogaster* is a suitable organism for such studies because it is a complex organism with distinct organs, tissues and cell-types and depicts similar feeding behavior as in mammals. Taking advantage of the powerful

genetic tools in fruit flies, I show that the manipulation of CrebA levels alters fly feeding behavior. These findings provide a novel insight in to how nutrients connect with transcriptional regulation of ER function, secretion of hormones / metabolic regulators and feeding behavior, an important finding in the field of metabolism.

3.1. Nutrient Metabolism

Metabolism of nutrients involves a cascade of responses from smell to the uptake of metabolized energy products by individual cells. Organisms have evolved complex mechanisms to make these processes as efficient as possible. Our sensory systems play an important role in perceiving the outside world in the form of chemical (taste and olfaction) and physical (mechanical, sound, vision and temperature) stimuli. But what factors drive our responses to nutritional sensing? Our responses can be classified in to two broad categories: immediate and post-digestive system level responses and each of them are discussed in detail below.

3.1.1. Immediate responses to nutrition

All organisms are dependent on effectively identifying and ingesting nutrients. Olfaction and taste act as first sensory responders to determine whether food is initially accepted or rejected. This initial sensory input in response to foods is thus critical in determining our innate and acquired feeding responses. Olfaction is mediated by olfactory or smell receptors that bind to odor molecules. Activated olfactory receptors then trigger electric signals that are transmitted to the central nervous system (CNS) where the odor is perceived. Our brain's capability to associate odors to desirable or undesirable foods forms part of our acquired feeding behavior and influences nutrition.

The sensation of flavor acts as the second response to nutritional intake. This sensation is part of a complex integration of taste, aroma, texture and chemesthetic (oral and nasal irritation cues) from a food source (Tepper and IT., 2020). Thousands of taste receptor cells sitting on our tongues detect different types of taste and project taste receptors present in our tongues are capable of detecting different types of taste. These include sweet, bitter, sour, salty and umami. Umami and sweet are "good" signals that promote feeding of nutritive food whereas bitter and sour are "bad" tastes that alert the organism to toxins and discourage food consumption (Yarmolinsky et al., 2009). Overall, all these signals can be transmitted to the feeding centers of the brain where they can regulate hunger or satiation. In general, both olfaction and taste can

heavily modulate feeding behavior based on the initial sensory input from a food source.

3.1.2. Post-digestive system level responses to nutrition

Apart from the initial sensory responses to nutrition, metabolism of the food ingested and the appropriate response post-digestion plays a major role in maintaining metabolic homeostasis. Interestingly, we know that olfaction and taste responses to nutritive foods primes the metabolic organs responsible for the digestion of incoming nutrients. Perception of food through sight, smell and taste, known as the cephalic phase starts a myriad of physiological responses, including increased heart rate (LeBlanc and Cabanac, 1989) and secretion of saliva and digestive enzymes (Feldman and Richardson, 1986). These cephalic phase responses are a transient and fast way to prepare the organism to digest, absorb and metabolize nutrients (Power and Schulkin, 2008).

System level responses to nutrition include cellular changes in specialized metabolic tissues which are broadly conserved between flies and mammals and are discussed below.

3.1.3. Parallels between mammalian and fly metabolism

Many metabolic pathways are conserved from worms to mammals as it involves the breakdown of the same energy containing molecules. Target of rapamycin (TOR) is an example of one such pathway. Both flies and mammals contain a TOR protein that responds to a myriad of signal from nutrient availability to cellular stressors. In most cases of conserved pathways between flies and mammals, the fly system is simpler. It contains less genes that are able to perform the same functions as multiple mammalian orthologs.

Drosophila has contributed significantly to our understanding of various biological processes including metabolism. It offers many advantages including 1) a short life cycle 2) low maintenance costs 3) availability of powerful genetic tools and 3) a ~75% shared disease genome homology to humans. In terms of physiology, many of the fly metabolic organs are obvious analogs to vertebrate counterparts. I discuss major metabolic organs below (Figure 3.1).

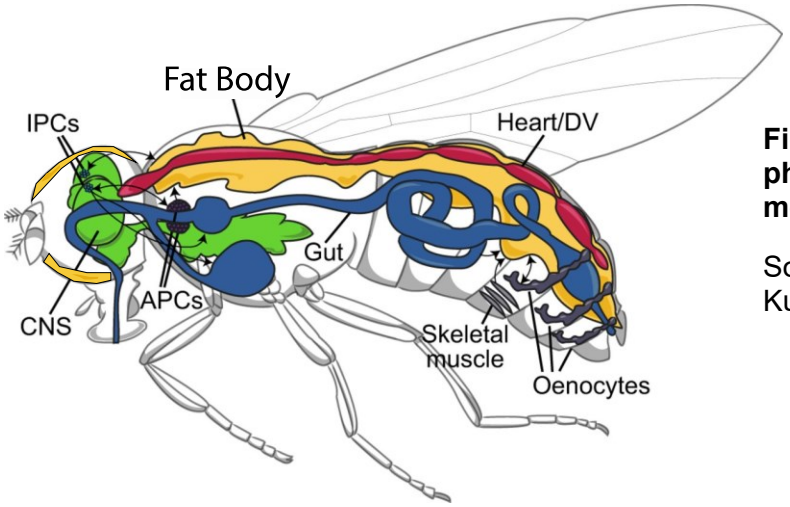

Figure 3.1 Metabolic physiology of Drosophila melanogaster.

Source: Musselman and Kuhnlein, 2018, JEB.

3.1.3.1. Gut

Drosophila gut is composed of three main parts, foregut, midgut and hindgut, each specializing for various functions in nutrient absorption. It absorbs macronutrients such as proteins, sugars and lipids. Food and water enter to the foregut and travel to the posterior parts where they are further metabolized in to intermediates. As in humans, *Drosophila* alimentary tract consists of multiple sub-regions that are specialized compartments for processing different nutrients. A region in the midgut, for example, specializes in dietary lipid absorption as indicated by regional accumulation of triacylglyceride (TAGs) (Buchon and Osman, 2015). Therefore, gut also houses protein complexes necessary for loading and transport of lipids. These lipoprotein complexes are made up of apolipophorins, orthologs of human apoA and apoB proteins that carry diacylglycerols and sterols from the gut to other tissues (Palm et al., 2012). Activation of a gastric lipase, *magro*, in flies via a nuclear receptor promotes midgut absorption and fat storage. Interestingly, a human anti-obesity drug orlistat, a gastric lipase inhibitor, is also able to reduce body fat in adult flies (Sieber and Thummel, 2009). Furthermore, a cross species expression of a human gut specific peptide, neurotensin, increases lipid accumulation in *Drosophila* midgut and fatbody (Li et al., 2016).

Similar to humans, Drosophila gut also hosts a complex microbiome. Diet affects the composition of microbiota species which in turn affects fly health and metabolism (Wong et al., 2016). Consumption of dietary sugars by the gut bacteria also affects overall lipid storage (Huang and Douglas, 2015).

Overall, the conservation of multiple pathways suggests the presence of a conserved gut system in *Drosophila* that has an important role in maintaining metabolic homeostasis.

3.1.3.2. Fat body

Fly fat body is the equivalent to mammalian liver and adipose tissues and acts as the central fat body storage hub. It accumulates fat stores in the form of large intracellular lipid droplets (LDs) during development and caloric load and executes lipolysis when energy is needed during starvation or egg production. Interestingly, adult fat body cells are allocated in various, unconnected adipose tissue throughout the body potentially allowing for more efficient storage and use of nutrients.

Utilization of LDs is mediated by the human glucagon ortholog, adipokinetic hormone (dAkh), which circulates in the hemolymph. Similar to mammalian glucagon signaling, Akh is activated upon fasting to meet the energy demands of the organism. dAkh is released from a neuroendocrine organ, reminiscent of pancreatic alpha cells. It is released under conditions of nutrient deprivation and acts on the fat body to induce lipolysis and glycogen breakdown (Kim and Rulifson, 2004; Lee and Park, 2004). Binding of Akh to its G protein-coupled receptor (AkhR) (Bharucha et al., 2008; Gronke et al., 2007) on the surface of fat body cells triggers the activation of the canonical cyclic AMP (cAMP)/protein kinase A (PKA) signaling. Increased intracellular cAMP levels activate pro-lipolytic pathways via the canonical cAMP responsive element binding (CREB) transcription factor (Iijima et al., 2009). Additionally, cAMP activates PKA dependent phosphorylation of proteins that promote the access of TAG lipases to the LDs (Patel et al., 2006) or boosts the transcriptional mechanisms to upregulate the production of lipases such as brummer, the fly ortholog of human adipose triglyceride lipase (ATGL) (Gronke et al., 2005; Wang et al., 2011). Similar to ATGL mutants, bmm flies are obese and unable to process lipolytic pathways under starvation conditions (Gronke et al., 2007).

Akh signaling also controls fat body fat storage via intracellular calcium (iCa) signaling using the conserved store-operated calcium entry (SOCE) system. SOCE acts as an adiposity regulator as changes as genetic changes that reduce or increase iCa levels greatly affect fat storage causing flies to become obese or lean, respectively (Baumbach et al., 2014). Changes in levels of many of the components involved in maintaining iCa levels modulates fat body TAG content in adult flies (Pospisilik et al.,

2010; Subramanian et al., 2013). Interestingly, these findings in flies have been recapitulated in mammals showing a role for SOCE in fat storage control (Maus et al., 2017). Clearly, *Drosophila* continues to be an important model organism for research in metabolism.

Drosophila fat body is also the target of the antagonist insulin signaling pathways. *Drosophila* expresses eight insulin-like peptides (Dilps)(Liu et al., 2016). Two of these Dilps, Dilp2 and 5 are released by the insulin producing cells (IPCs) in the brain upon high nutritional intake (Geminard et al., 2009). Deletion of Dilps results in type I diabetes, elevated blood sugar levels and reduction in stored fat – phenotypic responses that model insulin depletion in mammals (Zhang et al., 2009).

Until the last decade, it was thought that flies lacked the human satiety hormone, leptin. Leptin was believed to be exclusive to vertebrates where it inhibits hunger and diminishes fat stores by acting on the hypothalamus (Klok et al., 2007). However, a JAK/STAT pathway ligand, Unpaired 2 (Upd2) was later identified to be released by the fly fat body and promote insulin secretion. Interestingly, the neural circuits Upd2 stimulates is very similar to that of leptin further highlighting a conserved role for Upd2 in satiation in flies (Rajan and Perrimon, 2012).

3.1.3.3. Brain and neurosecretory cells

Neurons and neurosecretory cells in flies are crucial for regulating nutrient metabolism. Although flies do not possess a specific hypothalamus region, multiple independent neuronal populations act as modulators of initiation/termination of a meal by interacting with other sensory systems (reviewed in (Itskov and Ribeiro, 2013)). Furthermore, genetic screens have identified neurons that control body fat levels by mediating food intake and/or internal metabolic rates (Al-Anzi et al., 2009).

At the molecular level different signaling hormones/neuropeptides work to modulate food intake. Apart from Upd2, flies also express another leptin-like cytokine, Upd1 that increases food intake in response to a high fat diet (Beshel et al., 2017). Upd1 functions by suppressing the expression of neuropeptide F (dNPF), a fly ortholog of mammalian orexigenic neuropeptide Y. Like mammals, dNPF prolongs feeding and reduces food aversion (Wu et al., 2003).

Flies also possess special neurosecretory cells. Two populations of these cell types play an important role in fat and carbohydrate storage. These constitute of adipokinetic hormone producing cells (APCs) and insulin producing cells (IPCs).

Unlike humans, fly glucagon and insulin is produced in these specialized cells in the brain. The functions of both hormones in maintaining energy homeostasis are conserved in flies as discussed earlier.

3.1.3.4. Heart

The fly heart is a long, multichambered tube that circulates fly hemolymph throughout the body via an open circulatory system. This circulation is important for the transport of metabolites to peripheral tissues. Interestingly, obese *Drosophila* exhibit cardiac steatosis and fibrosis analogous to humans (Birse et al., 2010; Diop et al., 2015). Similar to mammals, obesity is associated with the accumulation of fat stores in the heart and mutants of fatty acid transporter proteins result in higher chances of heart failure (Sujkowski et al., 2012).

Normal heart function is required for the circulation of lipoprotein particles that carry lipids to where energy is required. Interestingly, *Drosophila* cardiomyocytes can function non-autonomously to control circulating sterols and lipids by synthesizing new apolipoproteins, previously known to be only generated in the fat body (Lee et al., 2017). Under high-fat diet conditions, these form a significant portion of the total TAG circulating molecules impacting system-wide lipid metabolism. Taken together, fly heart muscle has a crucial role in maintaining overall metabolism and its similarity to mammalian heart functions allows it to be a powerful tool to study metabolic diseases.

Many of the conserved pathways described above rely on the proper function of the ER. Lipoprotein formation and transport, two key processes affecting organismal lipid metabolism, depend on the complexes integrated in to the ER. Furthermore, ER morphology also affects this process heavily. In the next section, I will describe the role of ER in metabolism.

3.2. ER in metabolism

A structurally complex and dynamic ER allows it to efficiently respond to fluctuations in environmental cues. Regulation of nutritional changes, in particular, are handled through the ER as the organelle houses the secretory machinery that adapt and secrete hormones to communicate the nutritional status of an organism. Therefore, ER function becomes particularly important in organs / tissues with high secretory demands such as the liver. In this section, we will take a deeper look in to some of ER’s major roles in regulating organismal metabolism.

3.2.1. ER has a crucial function in regulating cellular metabolism

ER is an important gateway for many intracellular metabolic activities. Some of the most important pathways such as the production of glucose from glucose-6-phosphate occurs in the ER (Hutton and O'Brien, 2009; Waddell and Burchell, 1991). ER is also responsible for processing and transporting cholesterol molecules (Chang et al., 2006) and is a pre-requisite for steroidogenesis, oxysterol and bile acid synthesis (Kennelly and Tontonoz, 2022). Furthermore, ER is also responsible for the formation and release of lipid droplets (Balla et al., 2019; Gillon et al., 2012; Koerner et al., 2019; Yao et al., 2013) and if disrupted can lead to metabolic disorders (Baiceanu et al., 2016). These include lipid accumulation and increased glucose levels that can cause an excess release of calcium from the ER which drives it stress and chronic inflammation (Zhang and Kaufman, 2008). Although ER function and its role in metabolism has been extensively studied, it is not clearly understood how nutrient uptake requires ER's function to initiate the downstream molecular pathways for its metabolism.

3.2.2. ER is required for the formation of lipid droplets and the transport of lipid particles

Lipid droplets (LDs) are neutral lipid storage organelles found in most eukaryotic cells primarily in the cytoplasm. Since most of the enzymes required for neutral lipid synthesis are localized in the ER (Buhman et al., 2001), de novo LDs are formed in the ER. Many proposed models exist on exactly how this process occurs. One such mechanism is that they are formed by a conventional synthesis of an oil lens which buds out in to the aqueous cytosol (Figure 3.2; (Walther and Farese, 2009)).

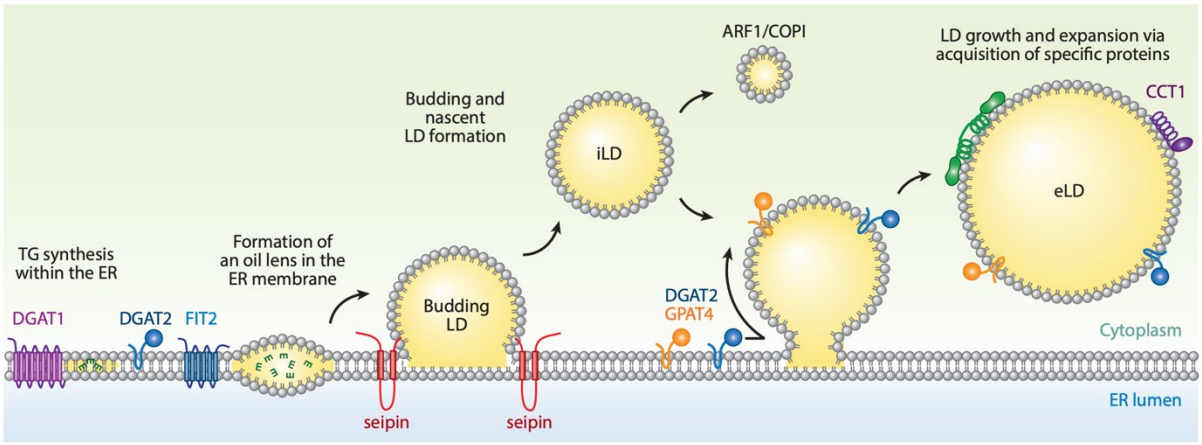

Figure 3.2 Formation of lipid droplets.

TG synthesis within the ER accumulates at the boundary causing the formation of an oil lens that buds away from the ER in to the aqueous cytosol. Source: (Walther and Farese, 2009)

This process includes the synthesis of basic lipid structures for example, triacylglycerols (TG) within the ER by the diacylglycerol acetyltransferases (DGATs; (Cases et al., 1998; Cases et al., 2001)). This is followed by the formation of an oil lens of neutral lipids within the ER bilayer driver by the accumulation of TG. ER structure plays a crucial role here in providing the right curved structures on the outer sheet that allow for an efficient budding process. Depletion of proteins that maintain ER shape and balance results in alterations of LD morphologies (Falk et al., 2014; Klemm et al., 2013). Lastly, when sufficient neutral lipids accumulate, LDs bud out in to the cytosol. A gradual increase in the angle with the ER pushed by the load of neutral lipids eventually leads to the fission of the bud from the ER (Thiam and Foret, 2016). Although this is considered to be a biophysical process, some proteins may be involved. However, this remains to be elucidated. Once LDs are formed, they may be targeted by proteins to become even larger structures and categorize them in to distinct populations (Wilfling et al., 2013). Clearly, the vast network of ER structures sitting in the cytoplasm provide a platform for the formation of LDs that then serve additional metabolic functions.

Since transporting lipids in aqueous milieu is difficult, organisms have developed mechanisms to transport lipids within hydrophobic protected shields called lipoprotein particles. Vertebrates express two different proteins ApoA and ApoB type proteins to transport very high density (HDL) and very low density lipoproteins (VLDL; Figure 3.3).

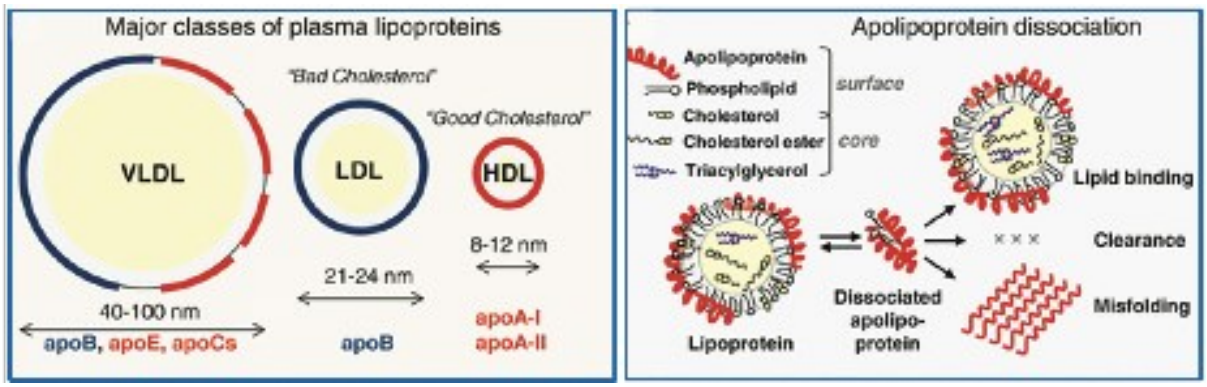

Figure 3.3 Major classes of plasma lipoproteins in humans.

Different apolipoproteins come together with a phospholipid layer to form distinct lipoprotein structures that carry lipids in the plasma. Source: (Das and Gursky, 2015).

Interference in the formation and transport of these particles has been shown to result in lack of circulating lipoproteins that are vital for cellular function and can result in metabolic disorders such as hepatosteatosis and hypobetalipoproteinemia (Burnett et al., 2003; Burnett et al., 2007; Minehira et al., 2008; Zhong et al., 2010).

3.2.3. ER houses major protein sorting machinery complexes

Secretion of proteins in to the intra- and extra-cellular space post production needs to happen quickly and dynamically upon a stimulus to ensure regulation of key pathways such as cell growth. The secretory protein sorting machinery sit at the ER that acts as a switch to meet the intra- and extracellular secreted protein demands (Feizi et al., 2017). More than 30% of all eukaryotic proteins use the ER protein sorting machinery (Figure 3.4) to secrete proteins in to the cytosolic and extracellular space (Gemmer and Forster, 2020). As the name suggests, these proteins are primarily housed within the ER and some in the Golgi. They are involved in the processing of nascent mRNAs starting with recognition of mRNAs targeted to the ER, translating them in to polypeptide chains, translocate in to the ER and processing them further to facilitate their folding and transport within vesicles between the ER and the Golgi (Aviram and Schuldiner, 2017; McCaughey and Stephens, 2019). A hydrophobic N-terminal signal peptide (SP) binds the soluble signal recognition particle (SRP) which allows the ribosome nascent chain to be recruited to the ER via the signal recognition particle receptor (SrpR). Sec61, a trimeric membrane protein complex binds the ribosome and enables the nascent unfolded peptide to enter through its channel. Other protein complexes sitting at the ER include the translocon associated protein (TRAP), oligosaccharyl transferase complex (OST) and the translocating chain associated membrane protein (TRAM). They serve important functions for the proper folding of the nascent polypeptide i.e., to support the recruitment of specific SPs (Nguyen et al., 2018), N-glycosylation (Kellenberger et al., 1997) and to determine which domains of the nascent chain are visible to the cytosol during translocational pausing (Hegde et al., 1998) respectively.

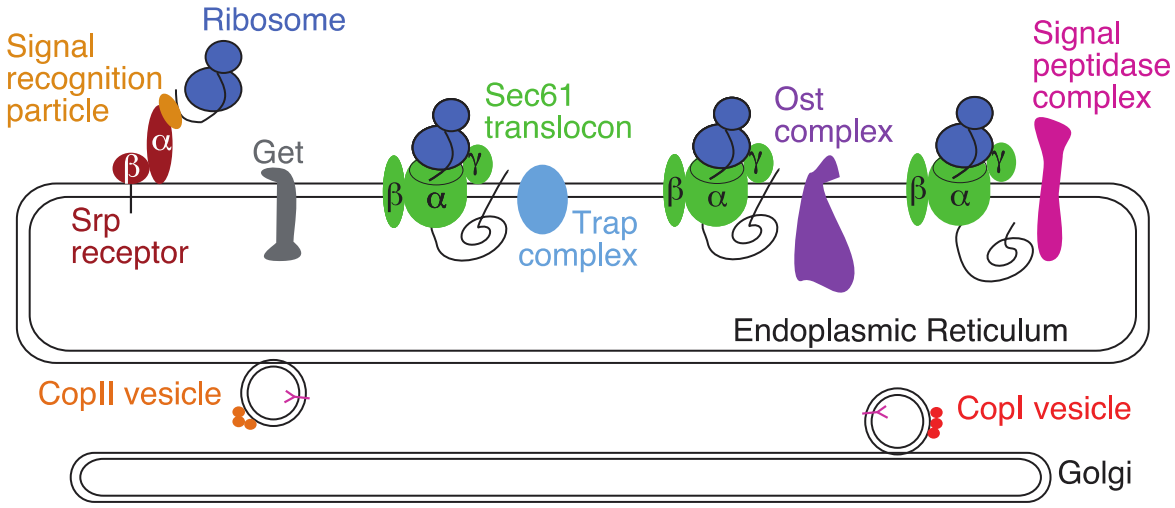

Figure 3.4 The Endoplasmic Reticulum houses the ER protein sorting machinery complexes.

ER protein sorting machinery genes are responsible for the processing of almost 30% of eukaryotic transcripts.

However, there are other components of the ER such as the TMCO1, the Get1 subunit of the Get1-Get2 complex that are thought to function independently of the Sec61 translocon in the insertion of tail-anchored proteins (Guna et al., 2018). The signal peptidase complex (SPC) cleaves off the SP from the proteins ready to be transported. Coat protein complex I and II (COPI and COPII) drive the formation of small vesicles and sort protein and lipid cargo to then transport these newly synthesized proteins from the ER to the golgi and vice versa (Arakel and Schwappach, 2018; Jensen and Schekman, 2011). Based on the functional relevance of these proteins, it is clear that the ER protein sorting machinery may stand at an important intersection of recognition of nutrient intake and the capacity of a cell to secrete and release relevant proteins required to adequately metabolize these nutrients.

The major components of the ER protein sorting machinery genes are highly conserved between flies and mammals. Mammals contain very similar complexes that recognize the SRP, a processing translocon and the downstream machinery to modify the nascent polypeptide chain. For a functional complex, mammals contain a similar ortholog that is needed. In some cases, however, mammals contain an additional copy of a gene which likely have additional tissue-specific role but this is not yet fully understood.

Many secreted hormones and other downstream signaling proteins are processed via the protein sorting machinery at the ER. Proper functioning of each of these protein processing molecular machines hence becomes vital for a healthy cell, tissue, organ and an organism.

3.3. Nutritional regulation of gene expression

The ability of an organism to detect changes in the environment and adapt to these changes is crucial for survival. Adaptation to nutritional changes need to be quick and transient in order for the body to adequately metabolize and/or store the incoming nutrients. Alterations to these evolutionary conserved metabolic readjustments are a major cause of metabolic syndromes. One-way organisms orchestrate a response to nutritional input is via the tight control of gene expression. Therefore, the intersection of nutrients, gene expression and metabolic disease is a growing area of study.

In response to environmental fluctuations, organisms need a tight control over gene expression (Figure 3.5). One-way organisms achieve this is through a class of proteins called the transcription factors. Transcription factors (TFs), belonging to the most studied class of proteins (Yusuf et al., 2012), are gene products that can regulate the expression of multiple target genes. TFs most commonly contain a sequence-specific DNA-binding domain that allows it to bind to a promoter region of a target gene (Latchman, 1997; Mitchell and Tjian, 1989; Ptashne and Gann, 1997). Some TFs can also bind to distal cis-regulatory elements and regulate transcription from distance (Spitz and Furlong, 2012). A trans-activating domain allows for the interaction with transcription co-factors or the RNA polymerase II to drive the transcription.

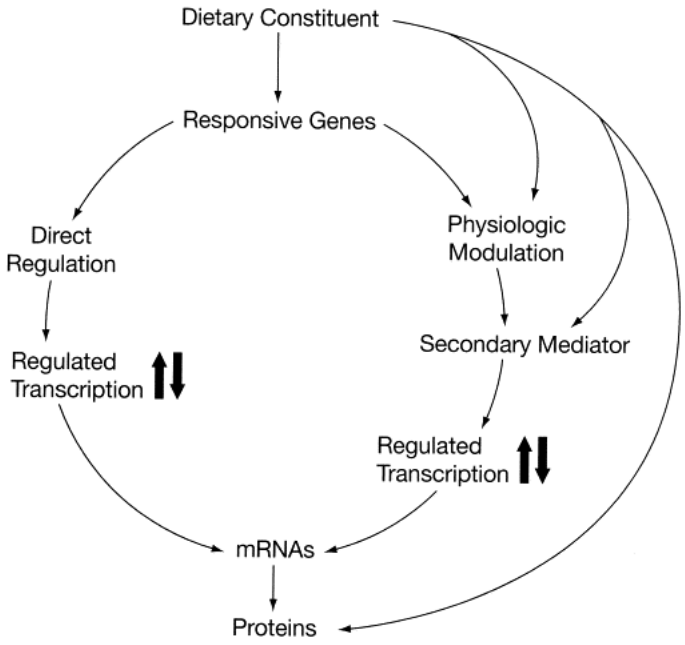

Figure 3.5 Dietary nutrients can regulate gene expression directly or indirectly.

Source: (Cousins, 1999)

Transcription factors are also the targets of many nutrient signals (Figure 3.6). Changes in nutrition regulate many of the transcription factors functioning in metabolic tissues to regulate the expression of genes that directly control metabolic pathways. This process occurs either via direct binding of a nutrient to the transcription factor or indirect activation of a signaling cascade. Sugars, fatty acids and amino acids all can regulate TFs.

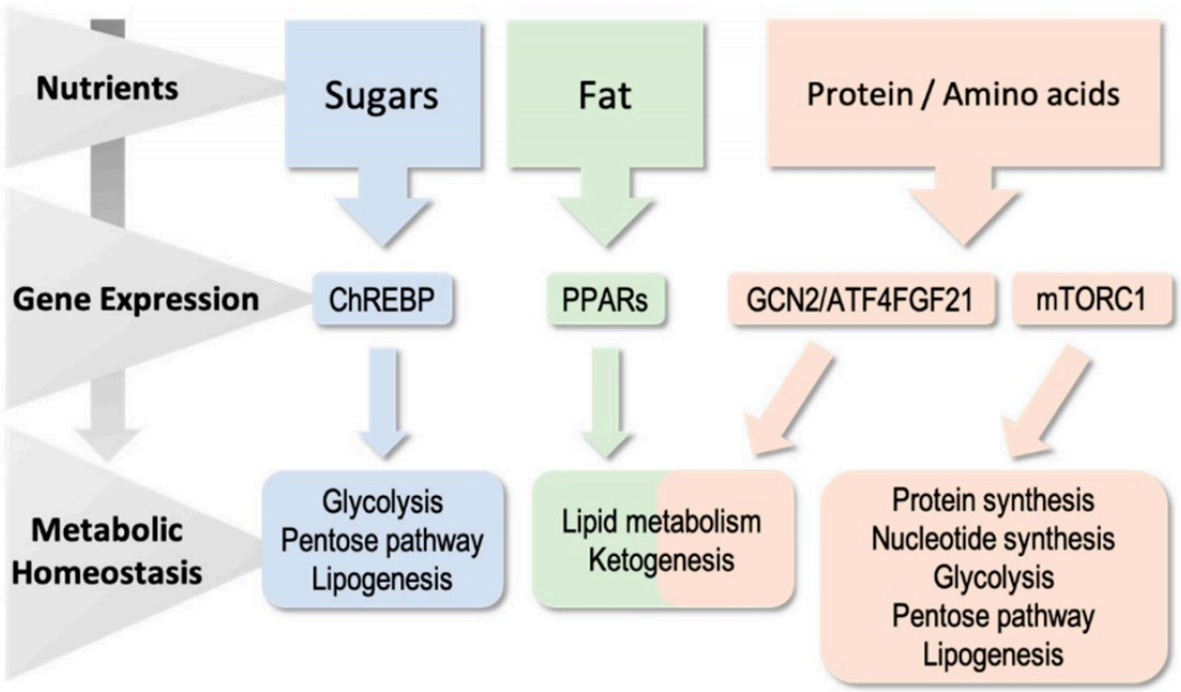

Figure 3.6 Nutrients such as sugars, amino acids and fatty acids initiate a metabolic response via the control of gene expression.

Many transcription factors have been identified to regulate gene expression upon specific nutrient macromolecules. ChREBP initiates glycolysis and lipogenesis pathways upon sugar recognition. Fat products can be directly recognized by PPAR proteins in order to initiate lipid metabolism. mTORC1 has been extensively studied as the master regulator of amino acid metabolism. Source: (Haro et al., 2019).

In this part, I discuss the current knowledge on the role transcription factors play in regulating metabolic homeostasis, focusing on their activation upon a nutrient signal and the downstream regulation of targets that optimizes cellular and organism-level metabolism.

3.3.1. Transcriptional responses to sugars

The quick and transient control of simple sugars such as glucose is mediated by the hormonal action of insulin and glucagon. The pancreas is the chief organ that senses glucose levels and induces the production and secretion of insulin and glucagon accordingly. The liver acts as a “buffer” where it is the main organ of action of the two hormones providing glucose under fasting conditions and storing it as glycogen when it is abundant.

Insulin is mainly produced in the β-cells of the pancreatic islets where its expression is tightly controlled at the transcription level. This is mainly controlled by the glucose-sensitive transcription factor pancreatic duodenum homeobox 1 (PDX1) (Ohlsson et al., 1993). PDX1 is a phosphorylation target of PI3K pathway which is activated under high glucose conditions (MacFarlane et al., 1994). Other transcription factors are also involved in the PDX1 mediated regulation of insulin production. The hepatocyte nuclear factor 3 β (HNF3β) also targets the promoter of PDX1 and positively regulates its transcription (Gerrish et al., 2004). Other members of the HNF transcription factor family are also expressed in the pancreatic β-cells, whose expression alterations results in maturity-onset diabetes of youth (MODY), a inheritable form of diabetes mellitus occurring in children and young adults (Winter and Silverstein, 2000).

Glucose signaling can also affect transcription via insulin-independent pathways. Entry of glucose in to the cells is mediated by glucose transporters (Glut) sitting at the plasma membrane (Oka et al., 1990). Interestingly, glucose can directly act as a ligand to bind and activate transcription factors. One such transcription factor called the carbohydrate responsive element binding protein (ChREBP) is a helix-loop-helix protein that binds a carbohydrate responsive element (ChoRE; (Yamashita et al., 2001)). The ChoRE sequence was found to be present upstream of key glycolytic and lipogenesis enzymes such as the liver pyruvate kinase (L-PK) and fatty acid synthase (FAS). ChREBP expression is high in metabolic tissues where upon glucose sensing it is phosphorylated to affect its cellular localization, stability and its transcriptional activity (Ortega-Prieto and Postic, 2019). ChREBP knockout mice are unable to survive on high sugar foods suggesting the protein is required to maintain the tight control of glucose levels.

Drosophila also possesses an intricate system of transcriptional responses that ensure mediation of sugar metabolism. Similar to mammalian ChREBP, flies contain a basic helix-loop-helix leucine zipper containing transcription factor, Mondo. Fly Mondo together with its binding partner *Mlx* regulates the gene expression of many carbohydrate and lipid regulating genes. Interestingly, larvae with reduced Mondo function are unable to survive on high sugar containing food and show increased blood sugar levels (Havula et al., 2013). Clearly, both human ChREBP and fly Mondo have

very similar functions and are necessary for initiating transcriptional cascades upon sugar intake that are necessary for survival.

3.3.2. Transcriptional responses to amino acids

Complex organisms synthesize 20 amino acids necessary for protein synthesis within a cell. Most of these amino acids are obtained in diet. Circulating levels of amino acids thus are a function of protein synthesis and protein breakdown. Therefore, how well an organism is able to metabolize ingested food to break it down to meet the amino acids requirement of all cells is crucial for cell survival.

The amino acid response (AAR) is the canonical pathway by which cells maintain adequate levels of amino acids upon a deficiency (Haro et al., 2019). Reduction of amino acid levels within a cell are marked by a deacetylation event of the corresponding tRNAs. These resulting uncharged tRNAs are able to bind and activate the general control nonderepressible 2 (GCN2) kinase, a direct sensor of amino acids (Garcia-Barrio et al., 2000). An activated GCN2 can apart from initiating the integrated stress response (ISR; (Harding et al., 2000)), also increase the translation of the activating transcription factor 4 (ATF4; (Kilberg et al., 2009; Vattem and Wek, 2004)). Elevated levels of ATF4 can up regulate the transcription of genes involved in amino acid import and metabolism (Harding et al., 2003). Further, ATF4 also initiates other transcriptional cascades reduce the overall use of amino acids by suppressing transcription (Lopez et al., 2007; Pan et al., 2007).

The mammalian target of rapamycin complex (mTOR) is a serine/threonine kinase and one of the most well-studied amino-acid sensor. In humans, mTOR complex is made up of additional protein complexes, TORC1 and TORC2. TORC1 is activated by amino acids via GTPase complexes sitting at the cell periphery and integrates nutrient signals and the energy status of cells (Kim and Guan, 2019). An important binding partner of TOR, called raptor is a crucial mediator of signal from GTPases to TOR. Raptor is a 150kDa protein that is necessary for the mTOR catalyzed phosphorylation of 4EBP1 resulting in increased translation initiation events and protein synthesis (Hara et al., 2002). One family of GTPases called Rag proteins have been identified to interact with mTORC1 via raptor in an amino-acid sensitive manner and the downstream activation of mTORC1 (Sancak et al., 2008). It is also known that TORC1 itself senses cytosolic and intralysosomal amino acids and activates downstream signaling to promote protein synthesis (Wyant et al., 2017). But

how does mTOR mediate the transcriptional signaling of genes involved in protein production? Initially it was thought that mTOR only functions by repressing the function of translational machinery such as the eukaryotic initiation factors (Brunn et al., 1997; Gingras et al., 1999). However, more recently it has been shown that mTOR can more specifically regulate transcriptional programs encoding amino acid transporters and metabolic enzymes via the ATF4 transcription factor (Park et al., 2017). TORC regulates the translation and stability of the ATF4 mRNA which in turn tunes the expression of many enzymes and transporters to balance translation supply and demand.

3.3.3. Transcriptional responses to fatty acids

An organism's inability to properly metabolize dietary fat is the leading cause of many metabolic diseases such as diabetes, fatty liver conditions and obesity. The perixosome proliferator-activated receptor (PPAR) were the first identified sensors of fatty acids (Lin et al., 1999; Xu et al., 1999). PPAR proteins are ligand-activated nuclear receptor transcription factors that can bind a PPAR responsive element (PPRE) via their DBD. PPAR transcription factor family comprise of three isotypes (α, Δ and γ) with different expression patterns and metabolic functions. Upon binding of fatty acids or its derivatives, PPAR heterodimerizes with the nuclear receptor retinoid X receptor (RXR) and regulate the expression of variety of processes including fatty acid storage and adipogenesis (Desvergne et al., 2006).

PPARα, for example is highly expressed in the adipose tissues and the liver. Its role is particularly highlighted during fasting when fatty acid load in the liver is high to be used as an energy source. It serves as the main driver of fatty acid oxidation by up regulating a broad set of genes involved in fatty acid transport, intracellular binding and oxidation of these fatty acids in cellular compartments such as the peroxisomes and the mitochondria (Desvergne and Wahli, 1999). PPARα null mice which survive under a normal diet and exhibit no abnormalities, are unable to sustain fasting (Lee et al., 1995).

PPARγ is enriched in the white and brown adipose tissues in mice. It is necessary for adipocyte differentiation as in the absence of PPARγ mice do not develop adipose tissue and develop insulin resistance and lipodystrophy (Jones et al., 2005; Ludtke et al., 2007). Interestingly, PPARγ exerts its effect on adipogenesis also

via regulating C/EBP transcriptional cascade, a key pathway required for adipocyte differentiation (Wu et al., 1999). Because of PPARγ role in insulin sensitivity, it is a target of anti-diabetic drugs that improve the glycemic control in type-2 diabetes (Lehmann et al., 1995).

PPARΔ is the third member of the PPAR family and is more ubiquitously expressed. It has a major function in modulating cellular energy consumption by switching energy production from glycolysis to fatty acid oxidation (Fan et al., 2017) or increasing the processes of fatty acid uptake and oxidation (Holst et al., 2003). Due to its crucial role in many cell types, ligand agonists have been developed and subsequent treatment with these has shown improvements in lipid metabolism and insulin sensitivity (Sznaidman et al., 2003).

Intriguingly, it has been proposed that flies do not contain PPAR orthologs. How does adipogenesis and the regulation of key enzymes involved in the process regulated in lower organisms? An answer to this might lie in understanding the roles of PPAR transcriptional coactivators. Mitochondria is a key powerhouse with critical functions in metabolizing nutrients and adapting cellular physiology tightly linking it with growth-promoting pathways and nutrient availability. PPARγ coactivator 1 (PGC-1α and PGC-1β) are PPARγ transcriptional coactivators and are potent inducers of mitochondrial biogenesis (Kamei et al., 2003; Puigserver and Spiegelman, 2003). Deletion of these proteins in heart and adipose tissues leads to severe respiration defects (Uldry et al., 2006). Flies contain a single PGC-1 homolog, *spargel*, which is required for expression of mitochondrial proteins in fly fat body and thus demonstrates a critical role for *spargel* in the coordination of mitochondria with cellular metabolism (Gershman et al., 2007; Tiefenbock et al., 2010). Having a single fly PGC-1 homolog together with the conservation of mammalian adipose tissue in fly fat body tissue, primes *Drosophila* to be an ideal system to study the many important roles of PPAR and its coactivator proteins.

Furthermore, there is additional evidence of a PPAR homolog existing in flies. A BLAST search reveals ecdysone inducing protein (E75) sharing ~ 60% homology to the DBD of PPARα and ~ 30% homology to the ligand binding domain (LBD). E75 is a nuclear receptor and a member of the ligand-regulated transcription factor family and can regulate feeding behavior and fat deposition (Schwedes and Carney, 2012). Additionally, ultraspiracle (USP), a fly homolog of RXR, is known to form heterodimers

with ecdysone receptor (EcR) for DNA binding (Thomas et al., 1993). Taken together, these findings suggest that fly E75 protein could act as a PPAR homolog in *Drosophila* and may perform similar functions as the PPAR proteins.

In contrast, a transcription factor family that is fully conserved from yeast to humans are the sterol regulatory element binding proteins (SREBP). SREBPs play a key role in coupling lipid metabolism with nutrition, energy stress, and pathological processes (Shimano and Sato, 2017). In vertebrates, SREBPs are activated upon a deprivation of cholesterol upon which they translocate from the ER to the Golgi Apparatus, proteolytically cleaved and targeted to the nucleus to induce the transcription of genes involved in sterol and fatty acid biosynthesis (Brown and Goldstein, 1997). Interestingly, *Drosophila* encodes a single dSREBP protein whose nuclear presence enhances transcription of fatty acid biosynthesis genes but not cholesterol or isoprenoid biosynthesis (Seegmiller et al., 2002). Further evolutionary studies need to done to understand if mammalian SREBPs acquired additional roles like regulating cholesterol homeostasis and lipid regulation in liver in response to insulin (Matsuda et al., 2001). Interestingly, entangled SREPB proteins reside in the ER and only(Hirose and Ohkuma, 2007) upon a stimulus are activated by binding to a chaperone that allows the loading of SREBPs on to the COPII vesicle for translocation to the Golgi (Espenshade et al., 2002). Clearly, the ER plays an important role for housing important nutrient responders that then regulate transcriptional cascades to alter metabolism.

3.4. Role of RNA Polymerase II in nutrient driven transcription

Transcription of DNA in to precursor mRNA requires the action of a multi-protein complex called the RNA Polymerase II (Pol II). RPB1 is the largest subunit of the enzyme that contains the catalytic site for RNA synthesis and a regulatory C-terminal domain. Other subunits of the enzyme such as RPB2 and RPB3 are required for complex assembly (Kolodziej et al., 1990). In total, RNA Pol II is comprised of 12 subunits that come together to perform vital functions from promoter DNA binding, forming transcriptional complexes and transcribing DNA in to RNA. RNA Pol II requires other general transcription factors and co-activators to perform its function.

3.4.1. RNA Pol II in transcription initiation

The first complex that the RNA Pol II forms is called the pre-initiation complex (PIC) together with 7 other factors TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH and mediator which come together at the promoter regions of protein coding genes. Different compositions of PIC have been observed at other classes of transcribed genes (Sadowski et al., 1993) or in different cell types (Deato and Tjian, 2007). Binding of PIC initiates the “opening” of chromatin that then allows for transcription to initiate. A key molecular subunit of the TFIID complex is the TATA-binding protein (TBP) that binds DNA, bends it and sets up a 3D DNA architecture for active promoters (Patel et al., 2018). Pol II binding is also assisted by the mediator complex and is an important protein via which TFs communicate their activation signals to Pol II (Schier and Taatjes, 2020). Bent DNA structures are then stabilized by TFIIA, TFIIB and TFIIF complexes (Imbalzano et al., 1994). TFIIH consists of a kinase module that allows it to phosphorylate serine residues on the CTD domain of RPB1 subunit of Pol II and regulate the early elongation steps of its release (Buratowski, 2009). This early elongation is not sufficient to produce mRNA and is a way for the system to be prepared for induction signals (developmental or other stimuli) that releases the inhibition of Pol II from a paused state (Muse et al., 2007).

3.4.2. RNA Pol II in transcription elongation

While some factors such as NELF bind Pol II and negatively regulate transcription elongation, there are other proteins that positively promote Pol II release and transcription elongation. A major player that is required for this event is the positive transcription elongation factor b (P-TEFb), a kinase / cyclin pair that phosphorylates the Serine 2 residues on the RPB1 subunit of Pol 2 (Peterlin and Price, 2006). P-TEFb also phosphorylates a subunit of NELF, resulting in its release from the complex (Vos et al., 2018). These events promote the association of CTD with other RNA processing and chromatin modifying complexes that in turn set the platform for transcription initiation (Ebmeier et al., 2017; Kizer et al., 2005). Histone modification critically histone H3 acetylation is imperative for continuous and efficient transcription. Another class of proteins that are critical and are part of the transcription elongation complex are ATP-dependent chromatin modifying enzymes. The SWI2/SNF family of ATPases are one of the most well characterized chromatin remodelers that slide nucleosomes

during transcription to adopt positions that favor other Pol II complexes to transcribe DNA (Becker, 2002).

It is clear that transcription initiation, elongation and termination are complex processes that require the input from multiple other protein complexes. This dependency on other factors results in multiple of entries for how transcription can be regulated and potentially targeted in human health. Similarly, nutritional cues also affect transcription. In the next section, we will take a look at how nutritional cues affect transcription specifically by regulating Pol II mediated transcription.

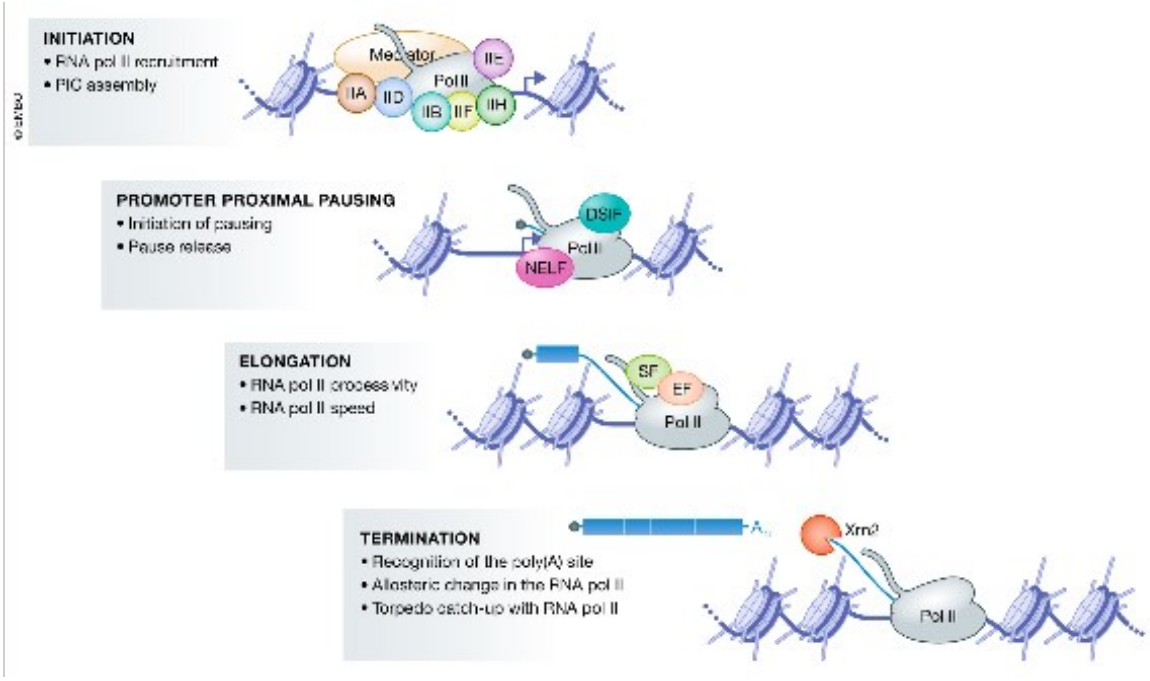

Figure 3.7 Key steps in the RNA Pol II transcription cycle

RNA Pol II transcription cycle starts with the assembly of the Pre-Initiation Complex (PIC) at the promoter regions. Initially, Pol II can be paused at the start by the inhibition of Pol II release. Once the paused Pol II is released, it transcribes DNA in to RNA before getting terminated by signals such as the poly(A) site. Source: (Muniz et al., 2021)

3.4.3. RNA Pol II regulation in response to nutritional cues

Metazoans have developed mechanisms to adapt to the influx of nutrients in order to gain energy from nutritional sources and get rid of unwanted materials. Transcriptional regulation is a key pathway that allows the organism to meet the demand for proteins to metabolize nutrients. As discussed in the previous section (see section 3.3), many metabolic signals can directly interact with and activate transcription factors to impact the expression of genes they regulate. Transcription factors achieve this by binding DNA recognition sites to recruit or block transcriptional machinery to the promoters of target genes (Castellanos et al., 2020). Although there have been efforts to understand

how transcription factors are regulated with nutritional cues, little is known how transcription factors in turn regulate the control of Pol II recruitment and release over target genes. Some studies have looked at effect of different diets and fasting on RNA Pol II activity. Bound RNA Pol II levels were decreased in mice liver upon fasting (Warnick and Lazarus, 1982). It is also known that transcription factors can regulate RNA Pol II activity. For instance, c-Myc, a crucial transcription factor that plays a role in cell cycle progression functions by interacting with P-TEFb to contribute to Pol II pause release over its target genes. Clearly, Pol II and the associated transcriptional machinery are regulated via their interactions with other proteins such as transcription factors. But how is this process mediated in response to nutritional cues? A better and clear understanding of how Pol II is regulated in response to nutritional cues by transcription factors will vastly improve how we approach and tackle nutritional disorders.

3.5. CrebA/Creb3L transcription factors as nutrient responders

Regulating secretory capacity of cells is an important pathway that requires to be regulated in response to environmental cues. Creb3 family of transcription factors are an interesting set of proteins that play an important role in regulating the secretory capacity of cells. These proteins were first identified as a single homolog in *Drosophila* named the cyclic-AMP responsive element binding protein A (CrebA) where it was shown to bind cyclic-AMP responsive DNA elements (CRE) (Abel et al., 1992; Smolik et al., 1992). Interestingly, Creb3/Luman a member of the mammalian family along with CrebA was pulled down with the host cell factor (HCF) protein that is involved in activating viral gene transcription (Abel et al., 1992; Smolik et al., 1992). Later, other members of the mammalian family were identified to be very similar proteins with roles in ER stress response (Asada et al., 2011; DenBoer et al., 2005) and metabolism (Chin et al., 2005; Kim et al., 2017; Lee et al., 2011; Lee et al., 2010; Zhang et al., 2012). Here, we take a deeper look in to the known Creb3 family of proteins, their similarities and differences as a better understanding of their expression patters and cross-talk will assist us in understanding what roles this family of proteins play in healthy and disease metabolism.

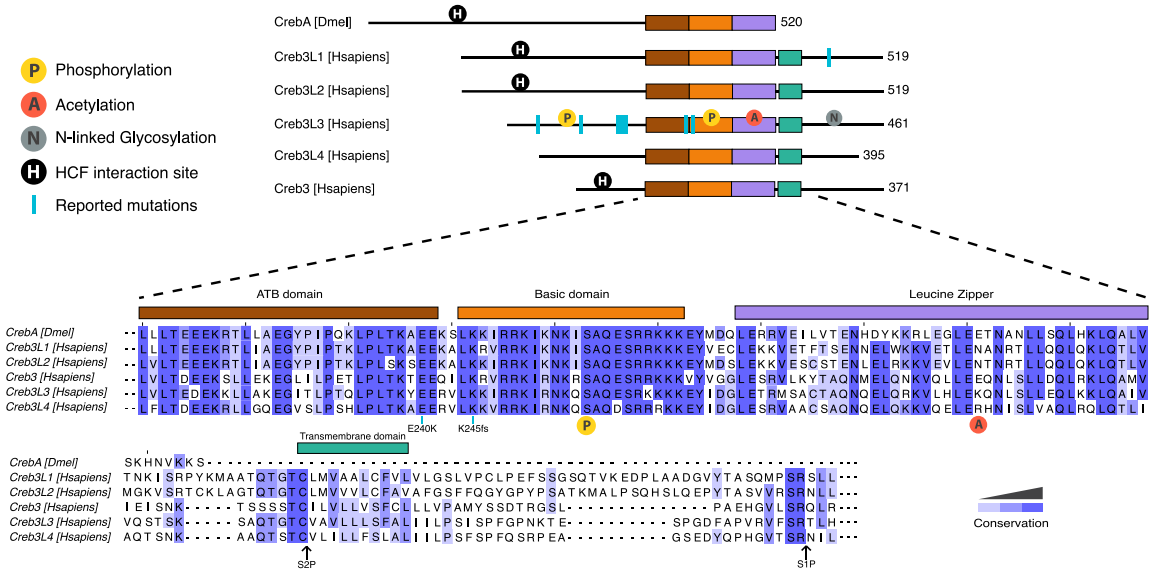

Figure 3.8 Protein organization Creb3 family transcription factors in humans and *Drosophila.*

Adjacent to bZip (ATM) (shown in brown), basic (orange), and leucine zipper (purple) domains are shared across all family members. Human members contain a transmembrane domain for anchoring in to the ER. An HCF interaction site is present in CrebA, Creb3L1, Creb3L2 and Creb3 proteins. Sequence alignment was performed using Clustal Omega.

3.5.1. Creb3 proteins share evolutionarily conserved domains

Mammals have five Creb3 family members, including Creb3/Luman, Creb3L1/OASIS (old astrocyte specifically induced substance), Creb3L2/BBF2H7, Creb3L3/CrebH and Creb3L4/AlbZIP/Atce1/Tisp40/Creb4. Creb3 transcription factors and the conserved domains within them have been traced back to ancient origins with presence of conserved domains in sponges (Barbosa et al., 2013). Being part of the b-Zip family of transcription factors, the leucine zipper serves as a dimerization domain. The leucine zipper allows Creb3 proteins to act as homodimers and/or heterodimers (Cui et al., 2016; Zhang et al., 2006) (Figure 3.8). Additionally, all Creb3 proteins possess a very similar basic domain. Basic domains allow for sequence-specific DNA binding. Presence of a nearly similar basic domain in distinct Creb3 proteins suggest they likely bind a very similar DNA motif. Numerous *in-vitro* and *-in-vivo* studies have shown that the basic domain is able to bind CRE and B-box motifs (Abel et al., 1992; Chin et al., 2005; DenBoer et al., 2005; Kondo et al., 2005; Lu et al., 1997; Nagamori et al., 2006; Omori et al., 2002). What type of gene promoters contain these sequences? Interestingly, all Creb3 proteins can drive the expression of an identical set of genes in *Drosophila* embryos (Barbosa et al., 2013; Fox et al., 2010). Similarly, knock-out

studies of Creb3 homologs in different organisms have revealed similar phenotypes suggesting these proteins a high level of functional and structural conversation (Hino et al., 2014; Ishikawa et al., 2017; Melville et al., 2011).

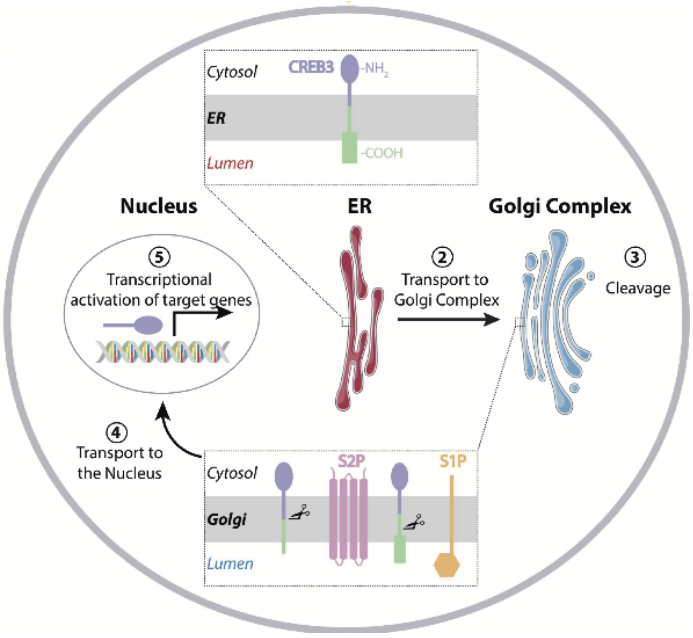

Figure 3.9 Role of Creb3 transcription factors transmembrane domain.

Creb3 transcription factors are anchored at the ER to be transported and cleaved off at the Golgi Complex and active part translocated in to the nucleus for its function. Source: (Sampieri et al., 2019)

Creb3 transcription factors also contain a transmembrane (TM) domain (marked green in Figure 3.8) that allows for anchoring at the ER and transport to the Golgi complex before a sequential cleavage by S1P and S2P proteases (Sakai et al., 1998). Interestingly, the d-CrebA does not contain the transmembrane domain suggesting that the mechanism might have evolved in higher organisms for tighter regulation of these pathways. It is also important to note that other transcription factors such as ATF6 and SREBP (Lemberg, 2011)(see section 3.3.3) having a role in metabolic regulation contain a similar TM domain. Once cleaved the N-terminal form of Creb3 is translocated in to the nucleus where it may function as homo- and hetero-dimers (Vinson et al., 2006). Therefore, this intramembrane proteolysis is an important precursor before transcriptional activation of Creb3 proteins.

Other domains present in Creb3 proteins include the adjacent to bZip (ATB, marked in brown in Figure 3.8). The ATB domain is required for transcriptional activity but it is not clearly understood what role it plays in transcriptional activation (Barbosa et al., 2013). Additionally, a HCF binding linear motif is present in all Creb3 proteins except Creb3L3 and Creb3L4 (Lu et al., 1998; Misra et al., 2005) **(see (Khan and Margulies, 2019)**.

3.5.2. Creb3 transcription factors target the metabolic and ER protein sorting machinery genes

Many genome-wide studies in multiple organisms have demonstrated the role of Creb3 transcription factor family in regulating the evolutionary conserved ER protein sorting machinery genes (Khan and Margulies, 2019). First evidence came from Deborah Andrew's lab in HeLa cells from overexpression studies. Overexpressing Creb3L1 in HeLa cells, which are biologically not specialized for secretion, was sufficient to increase expression of the ER protein sorting machinery genes (Fox et al., 2010). Similar phenotype was observed in human pancreatic beta-cells (Vellanki et al., 2010). One could argue that overexpression studies being gain of function experiments, this can result in indirect effects. However, different Creb3 proteins were shown to regulate the same set of proteins. d-CrebA target genes in the fly embryo are comparable to its mammalian homologs (Fox et al., 2010). Furthermore, a Creb3L2 knockout study in Medaka fish demonstrates a reduce in expression of genes involved in COPII vesicle formation (Ishikawa et al., 2017). Similarly, CrebL4 regulates genes responsible for protein sorting, maturation and degradation in human prostate cells (Ben Aicha et al., 2007).

In addition to transcription profiling studies, multiple studies have demonstrated the ability of Creb3 proteins to bind the promoters of ER protein sorting machinery genes directly. Using lacZ reporter assays in *Drosophila* embryos, Fox et al., demonstrated that d-CrebA can bind and activate transcription of multiple ER protein sorting machinery genes including the signal peptide receptor α, sec61ß, spase25, p24.1 and Copζ (Fox et al., 2010). Interestingly, ectopically expressed Creb3L1 or Creb3L2 in fly embryos was sufficient to induce the expression of the same components of the secretory pathway.

Chromatin Immunoprecipitation (ChIP) followed by sequencing is a commonly used technique to assess direct binding of transcription factors to its target genes across the whole genome. Interestingly, ChIP-seq experiments with Creb3L2 pull downs showed that it directly binds to promoters of Sec23α and Sec24δ (Kondo et al., 2007). Clearly, all this evidence suggests that Creb3 proteins have a conserved role in regulating the expression of ER protein sorting machinery genes. This is further evident from multiple experiments done across different cell-types and organisms. However, further genome-wide experiments are needed to better understand their role

in response to environmental stimuli specifically nutrition. Feeding / fasting are strong candidates for external stimuli that would require quick adaptability to increase / decrease the secretion of proteins in to the intra- and extra-cellular space. Creb3L3, for example, regulates genes involved in gluconeogenesis (Chin et al., 2005; Kim et al., 2017; Lee et al., 2010), triacylglycerol synthesis and fatty acid elongation (Zhang et al., 2012), lipid storage and transport such as the fat-specific protein 27 (Fsp27) (Xu et al., 2015) and the secreted lipoproteins ApoC2 and ApoA4 (Dandekar et al., 2016; Xu et al., 2014; Zhang et al., 2012). Do other members of the Creb3 protein family have a similar role in metabolic regulation? A better understanding of their role in metabolic regulation in response to different nutritional stimuli might aid us in answering key questions such as how an organism efficiently and dynamically maintains metabolic homeostasis in response to nutrition in healthy and unhealthy organisms.

3.5.3. Nutrient and stress signals regulate Creb3 family transcript levels

In order to understand the role of Creb3 proteins in metabolic regulation, it is important to study mechanisms which activate these proteins. One pathway that induces Creb3L3 expression is ER stress (Jang et al., 2011; Kondo et al., 2005; Kondo et al., 2007; Shin et al., 2012; Vellanki et al., 2010). Interestingly, activated inflammatory pathways also trigger Creb3L3 expression. Cytokines such as TNFα, bacterial challenges and hepatitis C virus (HCV) can induce Creb3L3 gene transcription (Dandekar et al., 2016; Jang et al., 2015; Jang et al., 2011; Shin et al., 2012; Song et al., 2017; Troha et al., 2018).

Are Creb3 proteins also activated by changes in nutritional cues? Again, evidence comes from the most studied family member, Creb3L3 which is regulated by metabolic signals in the liver and small intestine. Interestingly, this regulation does not seem to be straightforward. Creb3L3 transcript levels do not fluctuate just with feeding on vs feeding off cues. Both fasting and a high-fat diet induces Creb3L3 transcript levels in mice liver (Danno et al., 2010; Lee et al., 2010; Vecchi et al., 2014; Xu et al., 2014; Zhang et al., 2012). A similar phenotype was observed in zebrafish intestine where high fat diet induced Creb3L3 transcript levels (Zeituni et al., 2016). Little is known about what transcription factors are able to bind promoters of Creb3 genes and regulate their transcription in response to nutrients. Reporter assays done in mouse hepatocytes suggest that expression of PPARalpha in this system increased Creb3L3

promoter activity. Further, a PPARalpha was confirmed to bind a PPRE DNA motif located in the promoter of Creb3L3 (Danno et al., 2010). PPARalpha's major role in fatty acid metabolism (see section 3.3.3) and the nutrient-state dependent evidence of Creb3L3 transcript regulation suggests Creb3L3 is involved in nutritional regulation. Creb3L3 transcript levels were also shown to be dependent on the liver-specific transcription factor, hepatocyte nuclear factor alpha (HNF4alpha) (Luebke-Wheeler et al., 2008). HNF4alpha is a key hepatic transcription factor that is required for lipid homeostasis (Hayhurst et al., 2001). Further studies might reveal additional factors that are required for the regulation of Creb3 transcript levels.

Creb3 transcriptional control evidence highlighted above coupled with the presence of the transmembrane domain suggests a multi-faceted control might exist that allows for a quick and dynamic response to environmental stimuli specifically nutrition. It would be interesting to further explore how these two methods of Creb3 control work in tandem to ensure proper function in different cell types and in response to different nutritional cues.

4. Identification of feeding dependent gene expression and secretion changes in *Drosophila melanogaster*

4.1. Summary

Understanding gene expression and proteome changes in response to dietary cues is crucial to dissect the molecular mechanisms responsible for maintaining metabolic homeostasis. In this chapter, we present findings from time-course experiments targeted to understand transcriptional reprogramming and subsequent circulating proteome changes upon feeding in *Drosophila melanogaster*.

Using next generation sequencing, we sequenced mRNA collected from flies that were fed, fasted or refed after fasting along the time-course of 2,4,6 and 24 hours. Analyzing this dataset, I identified 1588 transcripts that changed in pairwise comparisons between these conditions. I find up-regulation of transcripts involved in lipid and carbohydrate metabolism upon fasting. Interestingly, among the genes up-regulated upon refeeding, I identified a group of genes belonging to the ER protein sorting machinery genes that were co-regulated together at 6 hours of refeeding. These group of genes are responsible for protein synthesis, folding and post-translational modification of proteins which enter the ER. This machinery is responsible for the translation of one third of all proteins (Chen et al., 2005; Choi et al., 2010). Important components of the ER protein sorting machinery such as members of the Sec61 translocon were regulated by fasting and refeeding. However, not all components of associated with ER function such as the guided entry of the tail-anchored proteins (GET) pathway were affected by these nutritional cues.

In order to determine the consequences of these changes in the ER protein sorting machinery genes upon the secretome, I performed liquid-chromatography mass spectrometry (LC-MS) analysis on the circulating hemolymph collected from flies under the same conditions. Interestingly, I identify 405 time- and feeding-dependent secreted proteins that are predicted to have a signal peptide (enriched with 66% of total proteins detected). Several secreted proteins including fit – a satiety hormone, dawdle – sugar signaling TGF-B ligand and members of Torso signaling are up-regulated upon feeding. I further validated these results using western blots. Overall, these results provide the first evidence for the response of a crucial ER protein sorting pathway upon refeeding that is specifically co-regulated to increase secretion of important extracellular hormones and ligands in the fly hemolymph.

4.2. Introduction

Animal feeding behavior is regulated by pre and postprandial processes. Nutrient sensing mechanisms allow the organism to mediate multiple downstream processes such as digestion, nutrient absorption and cessation of feeding. In order to understand how these processes are regulated, it is first important to understand feeding behaviors. Flies naturally feed in a circadian manner consisting of meals eaten once in the morning and once in the evening (Ro et al., 2014). Long-term fasting has been previously used as a way to induce feeding in flies synchronously (Ro et al., 2015) (Qi et al., 2015, Methodology). However, how quickly flies satiate themselves once given access to food *ad libitum* after fasting is not clearly understood. Furthermore, most of these studies have focused on feeding sugar-only diets whereas flies naturally feed on sugar and yeast substrates. Previous work in our lab therefore focused on understanding how long and how much flies ate during the transition from a state of hunger to satiation (Figure 4.1A). Flies were fed a mixture of 20% sucrose and 5% yeast extract to mimic their natural diet and consumption over several hours was measured using the CAFÉ assay (Figure 4.1B). Intuitively, total food consumption and rate of consumption depended on the amount of time fasted (Figure 4.1C). Flies were satiated within an hour of being given food (Figure 4.1D). These experiments gave the impetus to study transcriptional changes upon quick refeeding in flies.

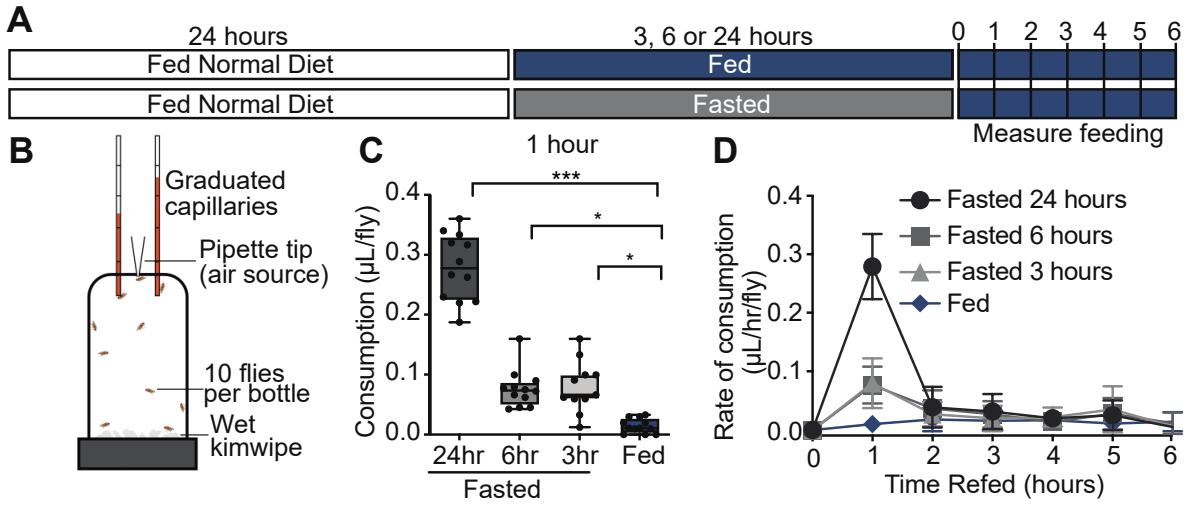

Figure 4.1 Flies satiate themselves quickly upon access to food. (Experiment by Teresa Burrell).

(A) The feeding regiment followed for measuring fly food consumption using (B) the CAFÉ assay as previously described (Ja et al., 2007). (C) Feeding post 1 hour from female flies that were continuously fed or fasted. Data (n=12, n=11 for 3h fasting condition) plotted as box-and-whiskers plot (min to max), analyzed with Kruskal-Wallis and post-hoc Dunn's multiple

comparisons test (***p<0.0001; *p<0.05). (D) Feeding rate of flies that were fasted for either 3,6 or 24 hours compared with continuously fed flies. Data plotted as the mean ±SD (n=12, except the 3 hr fasting condition n=11).

As discussed earlier, machinery integrated in to the ER sorts almost 30% of all eukaryotic proteins (**see section 1.2.4**). Most of the hormones and neuropeptides are also processed via the ER protein sorting machinery and are released in to the endocrine system where they regulate distant target organs. Insulin, for example, is one such hormone released from the pancreatic islet cells in to the blood affects feeding behavior and body energy stores via its action in the central nervous system (CNS) (Gray et al., 2014). In flies, a recently discovered satiety hormone called female-specific independent of transformer (FIT), is released in to the hemolymph by the fatbody cells and mediates feeding behavior via its action in the brain (Sun et al., 2017). Importantly, both insulin and FIT contain a SP signal that allows its processing via the ER protein sorting machinery.

In this chapter, I analyze transcriptional changes upon a fasting – feeding timecourse experiment focusing on quick changes in transcripts upon a nutritional stimulus. I identify that ER protein sorting machinery genes are co-regulated at 6-hours of refeeding. Furthermore, using a mass spectrometry approach, I analyse the hemolymph proteome along the same timecourse. I report that SP containing proteins are enriched in this dataset and change with nutritional cues. Overall, these results provide a detailed insight in to fasting – feeding dependent transcriptional and proteomic changes in the fly head and the circulating hemolymph.

4.3. Methods

RNA extraction for sequencing and feeding behavior experiments were done by the lab technician Teresa Burrell.

4.3.1. Capillary feeding assay (CAFÉ)

A modified capillary feeding assay (CAFÉ) was used to measure food intake and rate of consumption over time described previously (Ja et al., 2007). CAFÉ bottles were prepared by placing a wet Kimwipe on the bottom and two graduated capillaries containing 5ul of colored food. Food was prepared with 20% sucrose (Roth, Cat #4661.3), 5% yeast extract (Serva, Cat# 24540.3) and 5% (v/v) red dye (McCormick). 10 three to five days old flies were blown in to a single CAFÉ bottle with a hole on the side using an air mouth pipette. Amount consumed was measured at desired time-points by eye. The amount of time spent on filling the bottles with flies was taken in to account when measuring the food consumption to get more accurate readings. Food consumption per fly (µL) was measured as food uptake divided by the total number of flies in a bottle at the end of the experiment. Rate of consumption (µL/hr) was calculated as food consumed in a single hour divided by the total number of flies in a single bottle.

4.3.2. RNA extraction

For each time-point / condition, two biological replicates were collected. Each biological replicate comprised of thirty female fly heads. Biological replicates were collected on different days. Fly heads were homogenized in 200 µL Trizol reagent (Life Technologies, Cat#15596018) with a motorized grinder and left at RT for 5 minutes. Grinder ends were washed with 100µl Trizol in order to get all biological material off from it. The lysates were extracted once with 100 µL chloroform using a vortex and spun down at 12000g, 4 degrees Celsius for 10 minutes. The upper aqueous phase was taken and RNA was precipitated with isopropanol and 1µl of 200mg/ml glycogen (Thermofischer, Cat#10814010) using dry ice. Pellets were washed twice with 75% ethanol, air-dried and resuspended in 30µl RNAse free dH20.

4.3.3. cDNA preparation

15µg of RNA was taken and treated with 0.5µL TURBO DNA-free DNAse 1 (Ambion, M1907) at 37ºC for 30 minutes. The reaction was inactivated using 6µL of DNAse

Inactivation Reagent as described in manufacturer's instructions. cDNA amounts were measured afterwards using a nanodrop and checked for quality. For cDNA synthesis, 1µg of DNAse-free RNA was reverse transcribed using random primers (Thermo, 48190011) and 1µL SuperscriptIII (Thermo, 18080085) according to manufacturer's instructions.

4.3.4. RT-qPCR

cDNAs were quantified by qPCR using FAST or POWRUP SYBR Green Master Mix (Life Technologies) and a 7500 Fast or QauntStudio™ 3 Real-Time PCR systems. cDNA was amplified using gene specific primers and normalized based on a standard curve generated against genomic DNA extracted from wild-type flies. Standards were created separately for each primer to make sure all calculated measurements fall within the standard curve. Absolute values were plotted normalized to H2AZ.

4.3.5. Transcriptomics

RNA was extracted as described above. A sequencing library was prepared using Illumina polyA-mRNA library preparation methods with paired-end option. Both library preparation and sequencing were performed by the EMBL Genomics Core Facility. Fifty base pair reads were obtained from an Illumina HiSeq 2000 sequencer.

4.3.6. Informatic processing of the transcriptomics

Raw sequencing outputs were demultiplexed using barcodes used for library preparation. Reads were then checked for quality using fastqc and aligned to the reference genome (version dm6) using STAR (version 2.6.0). Uniquely mapped reads were counted per genes using STAR-quantMode GeneCounts using the annotation dm6.13 and –outSAMtype BAM SortedByCoordinate to get sorted BAM files. Transcript per million (TPMs) were calculated using rsem-calculate-expression and used as gene expression levels for plotting and observations. Coverage vectors were calculated from BAM files (generated by STAR) using tstools (https://github.com/musikutiv/tsTools/) and normalized over total coverage. Read counts obtained from STAR were normalized using trimmed means of M values using the NOISeq package (version 2.2) in R. Exploratory analysis was done by log2 transforming counts followed by batch correction using ComBat from the sva package in R. PCA plots were generated after centering the log2 transformed counts by the

mean and scaling by the standard deviation. Differential expression analysis was performed using maSigPro (version 3.6) in R. Minimum read counts per million (CPM) cutoff of 10 was used to get rid of low or zero abundance genes. Adjusted significance level (padj) cutoff of 0.05 and a R-squared value cutoff of 0.6 was used for variable selection in the stepwise regression models. Hierarchal clustering based on Euclidean distance was performed within maSigPro to identify 3 clusters in Figure 4.4B. Significant genes called were used to plot the heatmap in Figure 4.5A. GO term analysis was done using a Fischer's Exact Test and the database from biomaRt (version 2.38) in R.

4.3.7. Hemolymph extraction

Hemolymph was extracted as previously described (http://musselmanlab.com/wp-content/uploads/2018/09/adult-hemolymph-isolation-and-sugar-assays.pdf). A total of 3 biological replicates were used for each timepoint / condition with 1 biological replicate comprising of hemolymph from 20 flies collected on a different day. 20 flies were immobilized on ice to avoid carbon dioxide exposure. Flies were pricked one time in the thorax with a tungsten needle (Musselman et al., 2013) and placed in 0.5mL eppendorf tube with three 0.25um holes at the bottom. The smaller Eppendorf was placed in a bigger 1.5mL Eppendorf tube. This apparatus was centrifuged for 5min, 5000rpm at 4 degrees Celsius to collect 1µL of hemolymph. Hemolymph was immediately frozen with liquid nitrogen and stored at -80c before further processing.

4.3.8. Mass spectrometry

Peptide preparation: Peptides from hemolymph samples were prepared using the Preomics iST sample preparation kit (Cat# P.O.00027) as per manufacturer's instructions. For IP experiments samples were trypsin digested and peptides were desalted using Cº18 Stagetips.

LC-MS/MS: Samples were evaporated to dryness, resuspended in 15 µl of 0.1% formic acid solution and injected in an Ultimate 3000 RSLCnano system (Thermo), either separated in a 25-cm Aurora column (Ionopticks) with a 50-min gradient from 6 to 43% of 80% acetonitrile in 0.1% formic acid (hemolymph samples) or separated in a 15-cm analytical column (75µm ID with ReproSil-Pur C18-AQ 2.4 µm from Dr. Maisch) with a 50-min gradient from 5 to 60% acetonitrile in 0.1% formic acid (IP samples). The effluent from the HPLC was directly electrosprayed into a Qexactive

HF (Thermo) operated in data dependent mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 375–1600) were acquired with resolution R=60,000 at m/z 400 (AGC target of $3x10^6$). The 10 most intense peptide ions with charge states between 2 and 5 were sequentially isolated to a target value of $1x10^5$, and fragmented at 27% normalized collision energy. Typical mass spectrometric conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 250ºC; ion selection threshold, 33.000 counts.

4.3.9. LC-MS/MS data analysis

MaxQuant: Raw MS data files were processed using MaxQuant (version 1.6.3.4) with the Andromeda search engine with FDR < 0.01 at protein and peptide level. The default settings were used with the following modifications: variable modification methionine (M), acetylation (protein N-term) and the fixed modification carbamidomethyl (C) were selected, only peptides with a minimal length of seven amino acids were considered. Peptide identification was done using the drosophila melanogaster DB from Uniprot (uniprot_3AUP000000803_Drosophila_melanogaster_20180723.fasta).

Perseus: Contaminants and reverse hits were removed using Perseus. Label-Free Quantification (LFQ) intensity values generated by MaxQuant (Cox et al., 2014) were used. From a total of 12 samples representing 1 feeding condition (for example, Fed x 4 timepoints x 3 replicates), proteins containing equal to or greater than 8 out of 12 non-zero values were selected. Additionally, this selection criteria had to be met in at least 2 out of the 3 feeding conditions (Fed, Fasted and Refed) to ensure accurate representation. This ensured that proteins with missing values due to fasting condition (a biological and not a technical reason) were included in the analysis. There were no proteins observed having signal only in the fasting condition. Missing values were imputed based on a Gaussian distribution and values fit to the distribution with down shift of 1.8 and width of 0.5 within Perseus.

R: Imputed and normalized LFQ values were imported in R. Fasta sequences were downloaded for all detected proteins from Uniprot and signal peptide prediction was done using SignalP 4.0 with default parameters. PCA analysis was done with log2 transformed LFQ intensities centered by mean and scaled by standard deviation.

Visualization was performed using R base functions and graphics. Differential expression analysis was performed using maSigPro with default parameters.

4.3.10. Western Blotting

Proteins were separated using 6% SDS-PAGE gels at constant voltage of 200 V for 45 minutes. Transfer of proteins to a nitro-cellulose membrane was carried out at constant 100 V for 1 hour after which the membranes were blocked in 5% milk powder prepared in TBS-T solution for 1 hour. Primary antibody incubation was done over-night and membrane washed 3 times for 10 minutes in TBS-T. Secondary antibody incubation was done for 1 hour after which the membrane was washed again for 3 times with TBS-T for 10 minutes. Proteins were visualized by adding a 1:1 HRP substrate solution (Merck) with a developer. All primary and secondary antibodies were diluted with 5% milk containing TBS-T. 1µl of hemolymph was used for each sample. Apoltp, and Cv-d antibodies (gifts from the labs of Susan Eaton and Bruno Lemaitre) were used at 1:2000.

4.4. Results

4.4.1. Refeeding after overnight fasting results in transient changes in the transcriptome

In order to better understand mechanisms that regulate appetite, satiety and metabolism, we asked how nutritional intake after a period of fasting affects transcriptional expression. Since we were interested in understanding early transcriptional changes upon nutritional intake, RNA was extracted from very early timepoints after refeeding. Timepoints of 2, 4 and 6 hours were collected from wild-type flies that were continuously fed, continuously fasted or fasted for 16 hours and then refed (Figure 4.2A). An additional timepoint of 24 hours was also collected to understand expression dynamics post initial nutritional fluctuations. In order to compensate for circadian changes, all timepoints were collected between the different conditions were collected at the same time. A total of two biological replicates (collected on different days) were collected with each biological replicate comprising of 30 female fly heads. In order to determine the quality of sequenced reads obtained, I first analyzed the reads using fastqc (Figure 4.2B). Overall, all sample runs showed high quality reads with a single base pair quality of above the threshold of 28. Reads were then mapped to a reference genome, checked for mapping quality and quantified per gene.

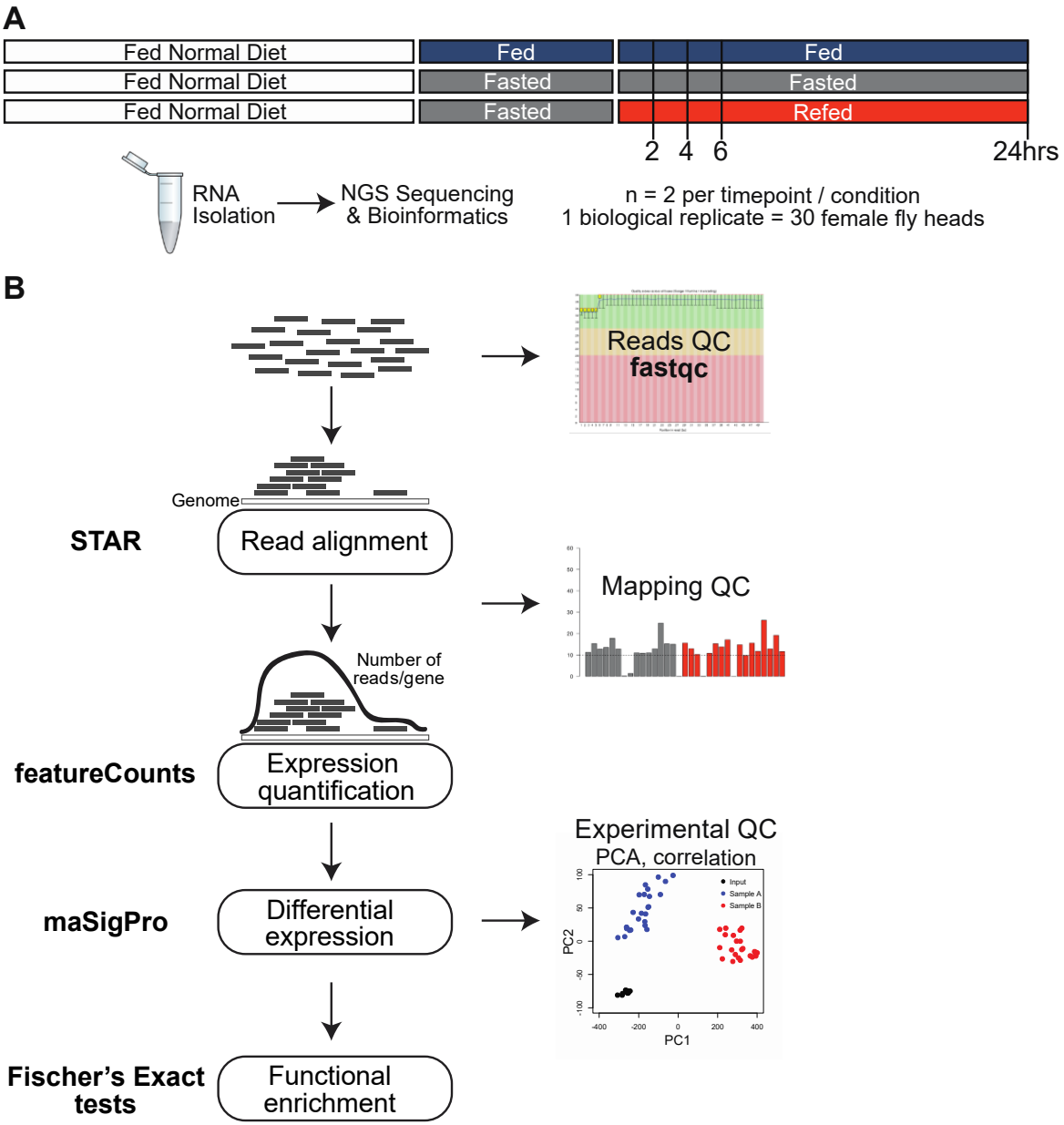

Figure 4.2 Experimental pipeline for the refeeding timecourse.

(A) The experimental regiment for used for the transcriptomics experiment. (B) The bioinformatic analysis pipline of the transcriptomics dataset. QC = Quality Check.

Next, I analyzed the normalized read counts per gene globally to assess 1) experimental reproducibility within the biological replicates and 2) overall changes in response to nutritional cues. I performed principal component analysis (PCA) on the most varying set of genes. Interestingly, the first two components (PC1 & PC2) explained 70% of variability within these genes (Figure 4.3). PC1 separated fed conditions from the fasted while the PC2 resolved variability occurring due to the feeding. As expected, refed samples showed the starkest differences. Already at 2

hours, my analysis showed that refeeding results in the samples being clustered even further away from fasted 2-hour samples. Interestingly, over time refed (24-hours) samples converged to resemble fed (24-hours) samples suggesting that the transcriptome activity is returning to resemble the continuously fed state. In addition, replicates clustered together depicting reproducibility between samples treated the same way.

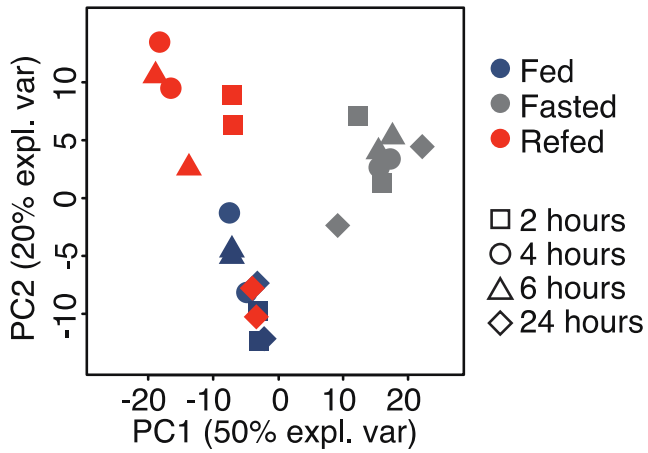

Figure 4.3 Principal component analysis of the most varying set of genes within the transcriptomics dataset.

PC1 explains 50% of variability within these genes and separates the fed, fasted and refed conditions. PC2 captures variability based on time as shown by Refed 24 hours replicates closest to Fed 24 hours state suggesting transcript levels returning to a normal fed state.

Once I had validated the quality of the data, I searched for differences between the nutritional conditions along the refeeding timecourse. I used a regression-based approach implemented in the maSigPro package in R (Conesa and Nueda, 2021) to find significant gene differences between the experimental groups. In total 1588 transcripts significantly changed with pairwise comparisons between the 3 different conditions (Figure 4.4A). Intuitively, most changes occurred between the fed and fasted conditions. Only a few transcripts (29) showed differences between the fed and the refed states belonging to the “Slow” cluster 2. Interestingly, these transcripts were enriched in mitochondrial proteins including the major components of the ATP synthase complex. Next, I used unbiased clustering on the 1588 transcripts to identify genes that may be coregulated upon feeding or fasting. A 621 gene cluster was seen to be upregulated with fasting which I termed as the “fasting genes”. Intuitively, this cluster was enriched in transcripts encoding for proteins involved in lipid and carbohydrate metabolism (Figure 4.4B, Cluster 3). Clearly, these genes play a role in making sure the organismal metabolic needs are met by for example, metabolizing nutrient reserves. Approximately half of differentially expressed transcripts are up-regulated with feeding. A cluster within these set of feeding up-regulated genes were a group of genes that were quickly and dynamically up-regulated within the 2-6 hours

of feeding (Figure 4.4B, Cluster 1). Additionally, another cluster containing 205 transcripts was identified which was consistently dependent on the feeding state. Low expression levels during fasting were returned to higher expression levels seen in the fed state over continuous feeding after 24-hours (Figure 4.4B, Cluster 2). Functional enrichment analysis did not reveal any group of genes that are specifically enriched in this cluster.

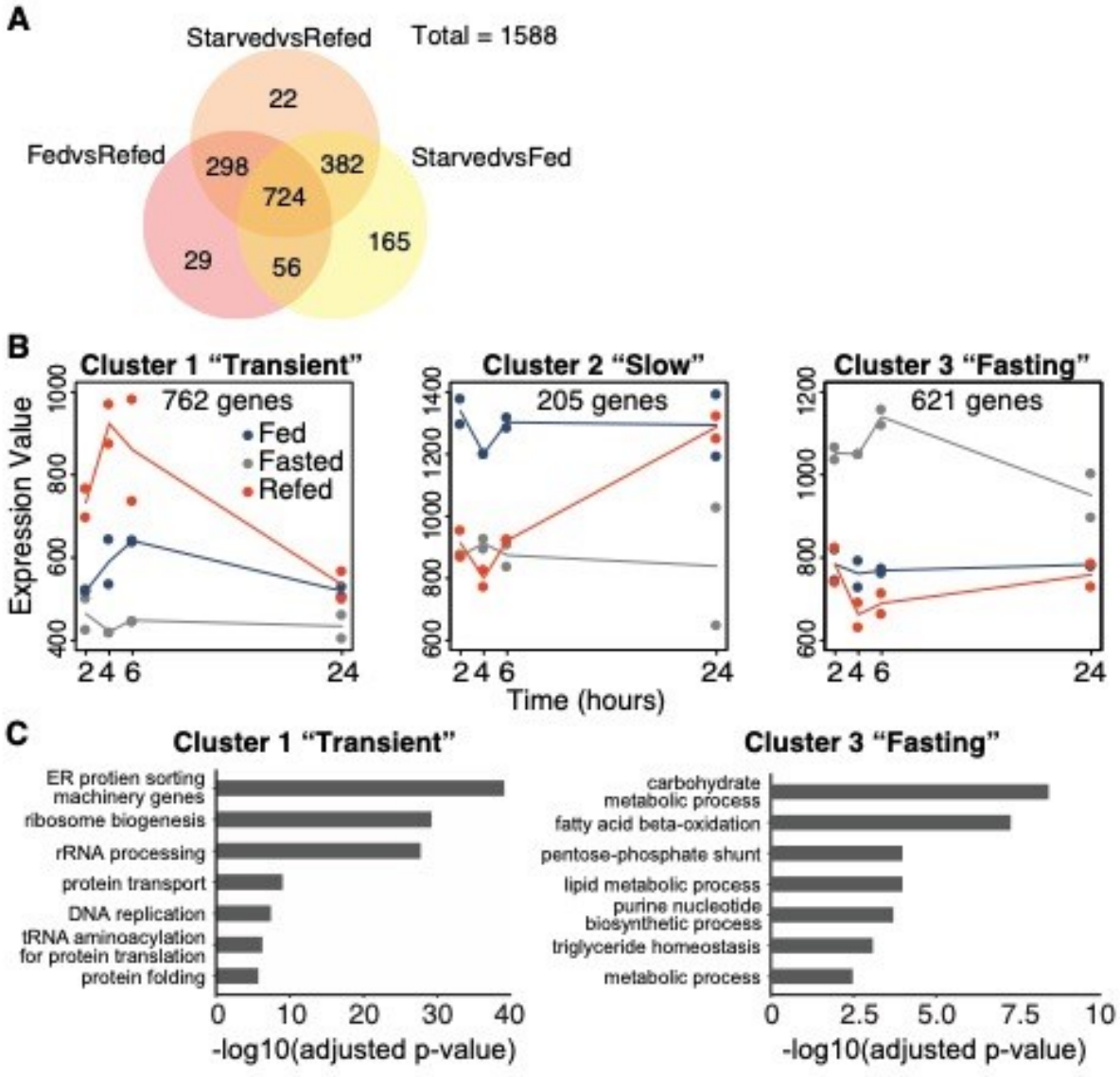

Figure 4.4 Refeeding induces changes in the gene expression landscape.

(A) Regression model based pairwise comparisons reveal changes between the three conditions. (B) Using maSigPro the 1588 differentially expressed transcripts separated into 3 main clusters. (C) GO term enrichment of the "transient" cluster 1 and "fasting" cluster 3 using a Fischer's Exact test.

4.4.2. ER protein sorting machinery genes are upregulated upon refeeding in a coordinated fashion

Next, I performed a GO term analysis on different clusters of genes that were coordinated together. Visualization of the data as heatmap revealed a clear cluster up-regulated with refeeding at the 6-hour timepoint (Figure 4.5A). Interestingly, functional enrichment analysis of this transiently upregulated transcripts contained transcripts annotated to a set of genes referred to as the “ER protein sorting machinery genes”. These set of genes are responsible for protein targeting to the ER to translate, modify and secrete one third of all eukaryotic proteins (Figure 4.4C). Specifically, these included the Sec61 translocon including Sec62 and Sec63, the Signal recognition particle receptor (Srpr), protiens incorporated into complexes of the Signal recognition particle (Srp), protein modifying complexes such as the Trap complex, Signal peptidase complex, Ost complex, and transcripts encoding for transportation machinery such as the CopII vesicle proteins, such as Sec23 and Sec13, and CopI coat proteins (including $\alpha, \beta', \gamma, \varepsilon,$ and δ Cop), in addition to several p24 proteins (Figure 4.5A). Most of these were up regulated together at the 6-hour timepoint. Collectively, this data demonstrated that nutrition induces a coordinated control of the ER protein sorting machinery genes.

As discussed earlier (see section 3.4.3), one-third of eukaryotic proteins are secreted via the ER protein sorting machinery genes. Next, I asked if other protein secretion pathways are also coordinated in a similar way in our dataset. Interestingly, other such known pathways like the mammalian TMCO1-dependent pathway (Shurtleff et al., 2018) and the guided entry of the tail-anchored proteins (GET) pathway (Schuldiner et al., 2008), do not appear to be nutritionally regulated. These transcripts include CG1598, which encodes a protein which shares 68% amino acid identity with human Get3 and 48% identity with the *S. cerevisiae* ortholog, as well as CG10470, which encodes a protein which shares 75% identity with human TMCO1. The GET pathway was identified as an ER protein sorting pathway that was independent of the Sec61 translocon (Yabal et al., 2003). However, the TMCO1 dependent pathway may require Sec61 (McGilvray et al., 2020). In addition, other ER “stress” pathways such as the ER-associated protein degradation (ERAD) pathway (Lopata et al., 2020) nor the Unfolded Protein Response (UPR) pathway (Shen et al., 2001; Yoshida et al., 2001), change upon refeeding in our RNA-seq time course. Specifically, neither Der1, Der2, Edem1, Edem2, nor Ter94 (the homologue of yeast

Cdc48 or mammalian VCP) are nutritionally regulated. In summary, refeeding specifically induces expression of the ER protein sorting machinery while other components of the ER remain unaffected by nutrition. This suggests a role for the ER protein sorting pathway in regulating response to refeeding.

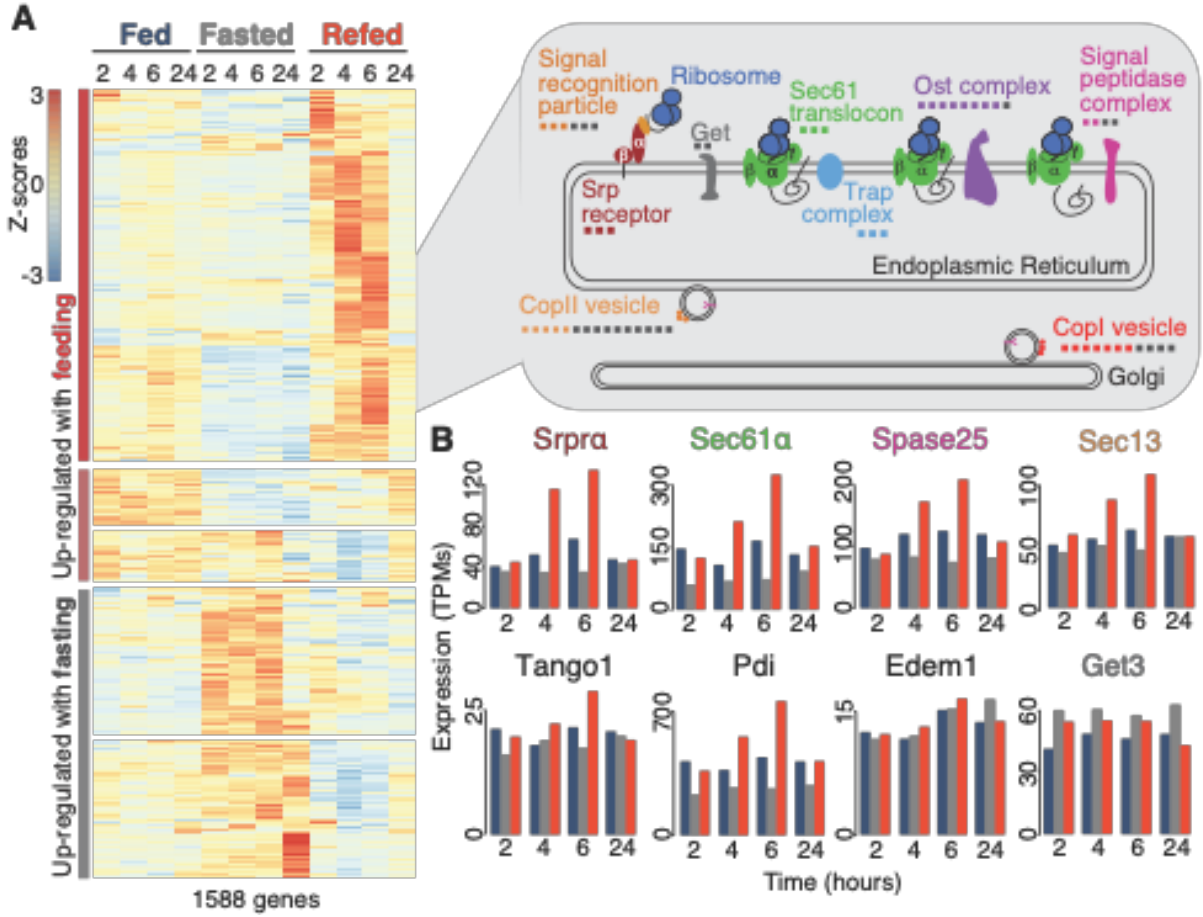

Figure 4.5 Refeeding specifically coordinates the expression of some ER sorting machinery genes.

(A) Representative heatmap of the 1588 “nutritionally regulated” transcripts identified from differential expression analysis from all conditions compared pairwise with maSigPro. 58 ER protein sorting machinery genes (out of 120) are enriched in a cluster upregulated 6 hours after refeeding. The gray inset illustrates a selection of the ER protein sorting machinery genes enriched in this cluster. Colored squares represent the number of genes differentially expressed in each complex out of total transcripts (dark grey). (B) Barplots from the transcriptomics data showing TPM expression levels for Signal particle receptor (Srprα), Sec61 translocon (Sec61α), Signal peptidase (Spase25), CopI vesicle (Sec13), Transport and Golgi organization 1 (Tango1), Protein disulfide isomerase (Pdi), ER degradation enhancer, mannosidase alpha-like 1 (Edem1) and Get3 (n=2).

4.4.3. Components required for lipoprotein particles are induced by refeeding

Lipoprotein particles play a key role in transporting lipids to where they are required in the body (see section 3.2.2). Clearly, they become even more important in response to an increase in nutritional influx. Next, I asked if pathways involved in regulating lipoprotein particles are affected in our timecourse analysis specially at early feeding timepoints. Interestingly, I identified transcripts such as the Tango1 (Transport and Golgi organization 1) upregulated upon refeeding (Cluster 1, Figure 4.4B; Figure 4.5). Tango1 forms part of the specialized CopII vesicles for transporting large molecules including preVLDL and pre-chylomicron lipoprotein particles (Dreyer et al., 2019; Rios-Barrera et al., 2017; Saito et al., 2009). Other transcripts included the microsomal triglyceride transfer protein (Mtp), an essential protein that transfers neutral lipids between vesicles and acts as a chaperone for the synthesis of many apolipoprotein B (apoB) (Hussain et al., 2003). Intriguingly, the interaction partner of Mtp, Protein disulphide isomerase (Pdi) that associates non-covalently with Mtp to form a fully functional lipid transfer complex, is also upregulated in the same cluster upon refeeding (Figure 4.5)(Wetterau et al., 1990). We know that fully functional Mtp-Pdi complex is required for metabolic homeostasis as its lack of function can result in fatty liver and intestinal lipid malabsorption (Berriot-Varoqueaux et al., 2000). Regulation of these transcripts in our dataset further suggests that these genes play a key role in maintaining metabolic homeostasis.

4.4.4. Hemolymph proteome is regulated by changes in nutrition

Our transcriptomics data clearly revealed the coordinated regulation of the ER protein sorting machinery genes upon refeeding. Importantly, since all subunits of the Sec61 translocon, an integral part of the protein translocation in to the ER process, are upregulated upon refeeding, I hypothesized that this response alters the ability of cells to secreted protein.

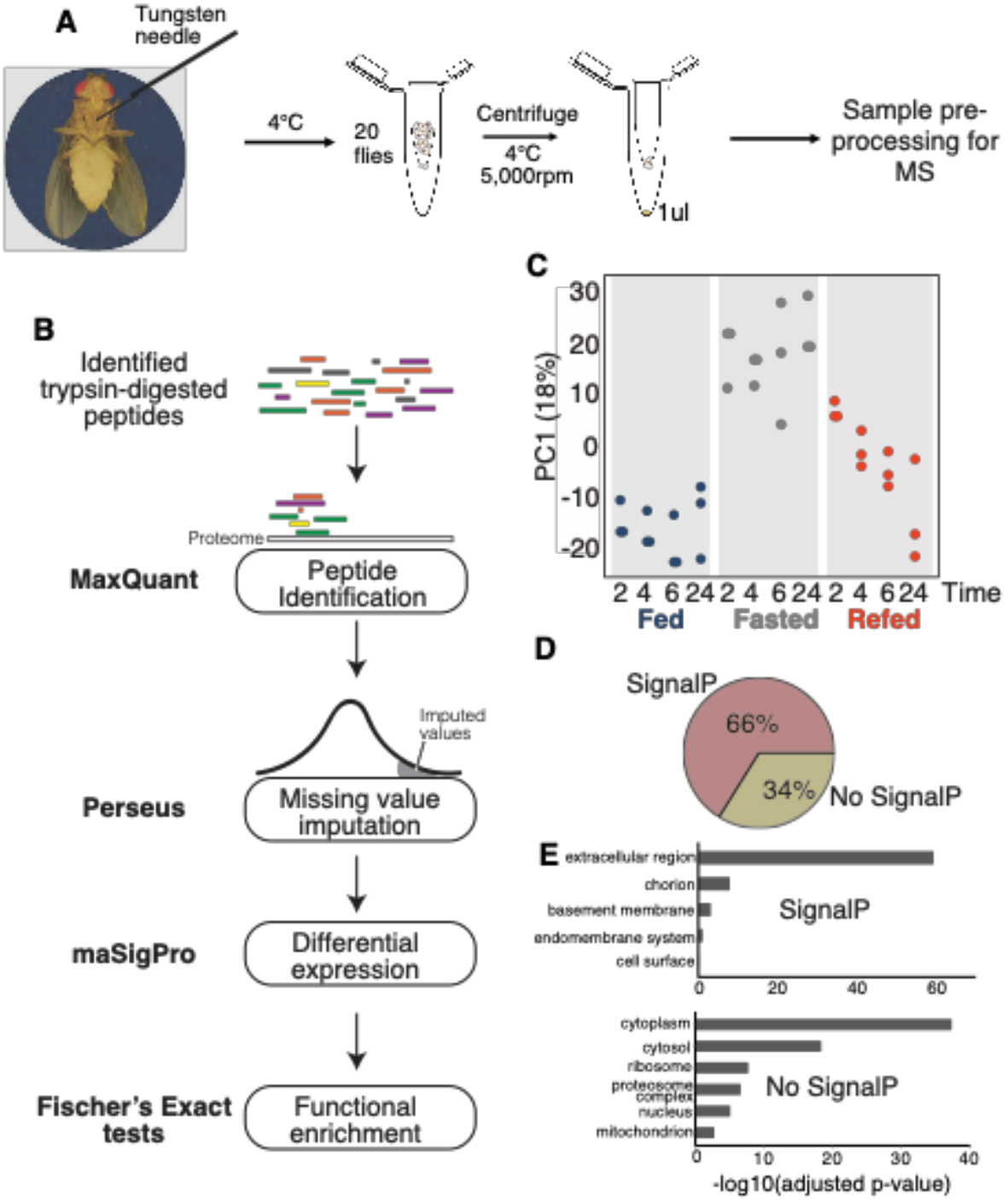

Figure 4.6 Analyzing fly hemolymph using a mass spectrometry approach.

(A) A scheme illustrates the protocol for isolating fly hemolymph (B) and the MS analysis pipeline. n = 3 with each n comprising of 20 flies collected on a different day (C) The first component from the principal component analysis plotted versus condition and time of proteins detected in the hemolymph. (D) Percentage of proteins predicted to have a signal peptide sequence plotted as a pie chart. Prediction was performed using SignalP-5.0 (http://www.cbs.dtu.dk/services/SignalP/). (E) Cellular component-based GO terms for proteins with or without a predicted signal peptide sequence evaluated using a Fischer's Exact test.

In order to test this hypothesis, I performed an orthogonal liquid-chromatography based mass spectrometry (LC-MS) analysis on the hemolymph of fed, fasted and refed flies. Hemolymph was collected from flies that were subjected to same feeding / time conditions as in the transcriptomics experiment (Figure 4.6A). I used a total of 3 biological replicates for each timepoint / condition with 1 biological replicate comprising of hemolymph from 20 flies collected on a different day. From 20 female flies that were used for a single collection, 1µl of hemolymph was collected which was used for trypsin-based peptide digestion followed by mass spectrometry runs (Figure 4.6B). In total, 1878 proteins were identified and used for downstream analysis.

First, I performed principal component analysis to assess the reproducibility and quality of data (Figure 4.2C). The first principal component captured both feeding state and timepoint dimensions. Continuously fed and fasted samples were clearly separated in this component I observed that early refed proteome resembled the fasting state with the secretome converging to resemble continuously fed samples over time (24-hours). Although I observed a similar pattern with the transcriptomics data, it was interesting to note that where the mRNA levels were highly regulated at refed 2-hours, the refed 2-hours proteome resembled the fasting state and took longer to return to continuously fed levels.

Next, I asked if the identified proteins are enriched in proteins localized to the ER protein sorting pathway. Using a prediction-based algorithm (SignalP), I searched for the number of proteins out of the 1878 proteins detected which contain a signal peptide. Interestingly, 66% of hemolymph detected proteins carry a signal peptide (Figure 4.6D). To further complement this this, I used the bioMart database to identify the cellular localization of the proteins containing a signal peptide. Interestingly, these were enriched to be extracellularly-localized while the ones not containing a signal peptide were enriched for locations inside a cell (Figure 4.6E). Since our method of hemolymph extraction also captured cells in the hemolymph, referred to as hemocytes, I suspect most of the proteins annotated to be cellular are a result of these. As we will see later, this was further confirmed by presence of immunity-related proteins in the dataset as hemocytes have a role in innate immunity and wound healing (Honti et al., 2014; Lemaitre and Hoffmann, 2007). Regardless, this data was enriched in proteins localized to the ER.

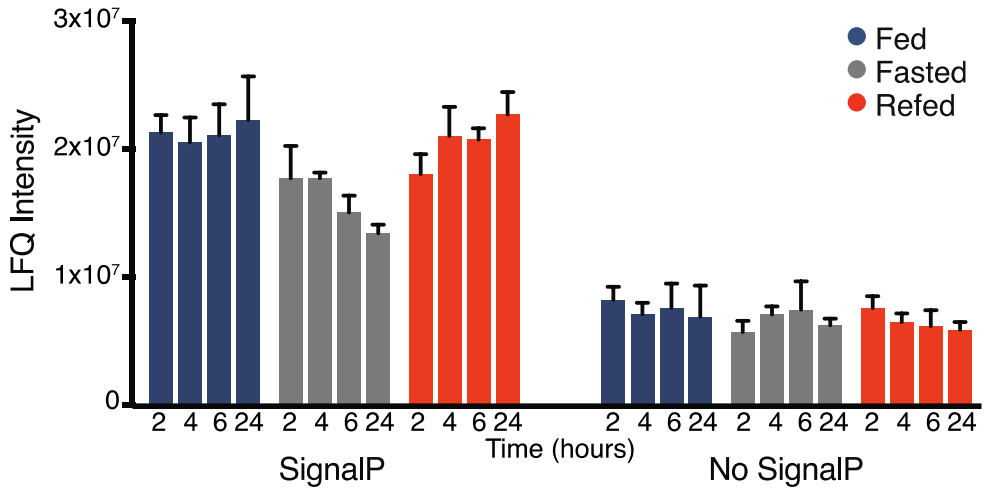

Figure 4.7 SignalP containing proteins change with fasting / feeding.

Proteins predicted to have a signal peptide change with fasting / feeding and have higher MS signals detected that proteins without a signal peptide. Average of all signalP vs non-signalP containing proteins plotted for biological replicates n=3 ± SEM.

In order to further understand the changing patterns of identified proteins, I asked if signal peptide containing proteins identified in Figure 4.6D differ in their hemolymph abundance than proteins not predicted to have a signal peptide. Interestingly, signal peptide containing proteins were significantly higher in abundance compared to no signalp proteins (Figure 4.7). Further, SignalP containing proteins oscillated with feeding and fasting while the No SignalP showed no oscillation. The decrease in abundance at fasted 24 hours for SignalP proteins was seen to be recovered at refed 2 hours and continued to increase to match levels of fed 24 hours. This pattern suggests that nutritional status affects the secretion of SignalP containing proteins possibly via its effect on the core protein sorting machinery genes.

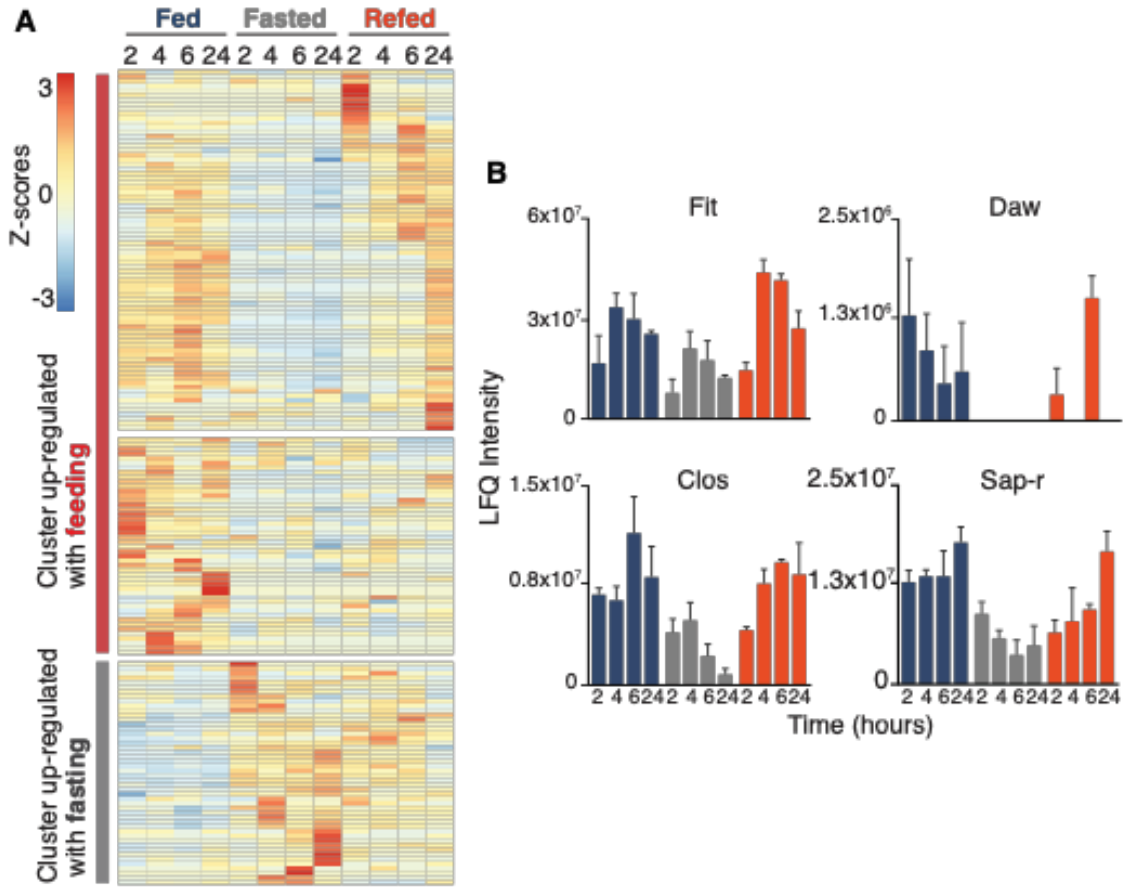

Figure 4.8 Hemolymph proteome changes with refeeding.

(A) Representative heatmap of 405 differentially expressed proteins predicted to have a signal peptide expressed in the hemolymph. Data are z-scores plotted for n=3 biological replicates. (B) Protein expression along the refeeding timecourse of Fit, Daw, Clos and Sap-r proteins plotted as LFQ intensities for n=3 ±SEM.

Next, I used maSigPro to identify any changes that occur in the pairwise comparisons of the three different nutritional conditions. 405 proteins out of a total of 1878 proteins identified changed with nutrition and were predicted to be secreted through the ER protein sorting machinery genes. Unbiased clustering on these 405 proteins identified a group of proteins that are upregulated with fasting and another larger group that are upregulated with feeding. Within the feeding-upregulated cluster, I identified a small set of proteins that were transiently upregulated at refed 2-hours. Interestingly, the abundances of this cluster were even higher than the fed state and were only detected at 2 hours after refeeding. When searched within this cluster, I identified a set of proteins involved in Torso signaling including Fs(1)N, Fs(1)M3 and Closca (Clos) (Ventura et al., 2010) (Figure 4.8B). Further exploration in to this

coordinated regulation of proteins involved in Torso signaling may identify a new role for these proteins in the future.

The hemolymph MS data also identified changes in feeding-dependent hormones. Interestingly, the sugar signaling TGF-β ligand, Dawdle (Daw), was transiently increased upon refeeding (Figure 4.8B). Daw is a circulating ligand produced by the *Drosophila* fat-body upon sugar consumption and activates TGF-β signaling to repress the activity of digestive enzymes (Chng et al., 2014). Another hemolymph circulating hormone, Female-specific independent of transformer (Fit), was also identified to be transiently secreted within 4 hours of refeeding (Figure 4.8B). Fit is a satiety hormone that forms part of the feeding-cessation pathways in flies (Sun et al., 2017). Activation upon initial refeeding after a fasting period may suggest that this forms part of the feeding behavior control mechanism. Other observations included the identification of Saposin-related (Sap-r), a protein involved in sphingolipid metabolism (Sellin et al., 2017), highly abundant at 24 hours after refeeding.

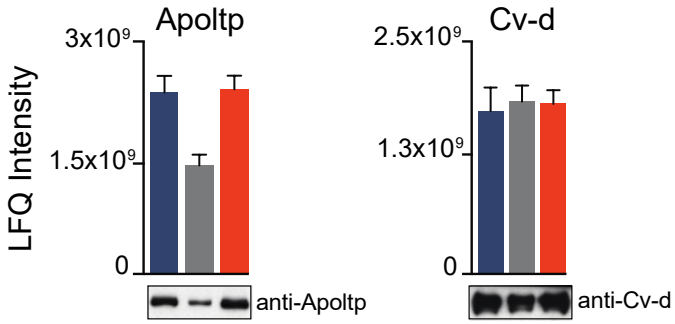

Figure 4.9 Confirmation of MS data with westerns.

LFQ intensities plotted as barplots showing levels of Apoltp and Cv-d from the MS hemolymph data together with complementary western blots.

Human ApoB proteins form the membranes of lipoprotein particles, required for lipid transport. I identified the *Drosophila* ApoB ortholog, Apoltp (Palm et al., 2012) also regulated upon feeding in the MS dataset. In order to confirm these changes, I acquired already developed antibodies against Apoltp and Crossveinless d (Cv-d, control) and performed western blot experiments using similarly acquired hemolymph samples. Secreted apoltp levels were seen to decrease upon fasting as in the MS data at 24-hours while Cv-d levels remain unchanged (Figure 4.9). Together, this data revealed many hormones, signaling peptide and other secreted proteins that responded to feeding / fasting cues and are likely involved in metabolic regulation. Clearly, some of the changes may be a direct result of changes previously observed in the expression of the ER protein sorting machinery genes.

4.5. Discussion

In this chapter, I used genome-wide analytical approaches to dissect gene expression and protein-level changes in female *Drosophila melanogaster*. By performing an extensive time-course analysis on fasting – feeding states, I identify the ER protein sorting machinery genes as being co-regulated with nutrition which in turn affects secretion of important metabolic regulators in the hemolymph of the fly.

Application of omics approaches has facilitated the understanding of molecular changes upon nutrition. Here, I used an RNA-seq approach to understand transcriptional dynamics upon feeding that prepares the organismal for the necessary metabolic responses. I performed unsupervised clustering of the most varying genes between the feeding and fasting conditions and identified three broad gene expression patterns. Importantly, I identify of the ER protein sorting machinery genes as being co-regulated upon refeeding at 6-hours. Clearly, ER plays an important role in organizing intracellular metabolism and is crucial for lipoprotein synthesis and secretion. This analysis provides the first evidence for the coordinated change in ER protein sorting machinery genes in response to nutrients which can be crucial for the ER to adapt and support its metabolic roles. Interestingly, gene expression changes in response to feeding seem to be specific only the conventional secretory machinery genes. Other ER associated pathways such as the GET pathway and the UPR pathway genes are not affected. In the future, it would be interesting to study how ER morphology is directly affected by feeding and if it plays a role in incorporation of the increased bulk of protein sorting complexes in to the ER.

I report a consistent up-regulation of ER protein sorting machinery genes at the 6-hour time point. Protein level changes however, when checked for Sec61alpha did not fully corroborate the mRNA changes. Sec61alpha protein was found to be higher in the fed states when compared to the fasted levels but only after 9 hours of feeding compared to 4 hours at the mRNA level. Interestingly, it did not show any differences between the fed and refed condition. This could be due to the requirement for the cells to revive the mRNA load of these genes after an overnight fast to then be able to adapt more quickly to changes in nutrition. Further studies will be needed to understand the translational dynamics of Sec61alpha and other ER protein sorting machinery genes.

The endocrine system in complex multi-cellular organisms provides crucial organ to organ communication. Fly hemolymph contains a myriad of signaling mediators

apart from metabolites. However, little is known how these signaling mediators are regulated upon nutrition in adult flies. Here, I used a liquid-chromatography based mass spectrometry to identify changes in the circulating hemolymph upon a similar refeeding timecourse experiment. It revealed changes in 405 proteins predicted to be secreted through the ER protein sorting machinery genes. These included transient changes in feeding dependent hormones such as Dawdle and Fit. Furthermore, I picked out more slower changes in important lipid carrier proteins which were confirmed with western blots. However, flies also express other known peptides that regulate feeding behavior. These include sNPF, fly ortholog of human sNPY neuropeptide (Lee et al., 2004), Hugin (Melcher and Pankratz, 2005), CCHa2 (Ren et al., 2015), dilps and adipokinetic hormone (Kim and Rulifson, 2004; Lee and Park, 2004; Gáliková et al., 2015). Inability of our MS approach to identify these peptides could be a result of trypsin digestion sites within these already small peptides and / or generally low abundances within the hemolymph. Targeted approaches may be needed to identify these peptides in the future.

In summary, these experiments provide a first detailed analysis of transcriptional and proteomic changes in flies upon changes in nutritional input. In doing so, I uncover coordinated changes in key pathways that allow the organism to adapt to changes in the metabolic environment.

5. CrebA, a conserved and essential transcriptional regulator of ER protein sorting machinery in response to nutrition

5.1. Summary

To identify the mechanism regulating the expression of ER protein sorting machinery genes, I examined more immediate changes in the refeeding-timecourse transcriptome. This analysis identified an obvious candidate which regulates the ER protein sorting machinery during development, the transcription factor cyclic-AMP Responsive Element Binding A (CrebA).

I show that CrebA up regulated transiently within 2 hours of refeeding. *Drosophila* CrebA, an essential and an evolutionarily conserved transcription factor, has been shown to regulate the ER protein sorting machinery genes during development (Fox et al., 2010). However, it is not known if CrebA acts as an upstream regulator of these genes in response to environmental changes. Further analysis confirmed this quick and transient upregulation in CrebA mRNA translated to protein levels and subsequent binding of the transcription factor to its target regions. Using ChIP-seq, I confirmed that nutritionally responsive CrebA binding sites were highly enriched in genes annotated to the ER protein sorting machinery genes suggesting CrebA as the upstream regulator of these genes. In addition, analysis of Pol2-ChIP suggests that CrebA mediates transcriptional regulation by promoting Pol2 release from its paused state to elongate over the gene body as confirmed by time-dependent changes in Pol2 elongation over the gene body.

Gene expression control is key to maintaining nutritional homeostasis in response to feeding (Haro et al., 2019). But how these gene expression changes occur in a relatively quick and transient manner is still not understood. I hypothesized that CrebA might act as a responder to feeding and in turn regulate the gene expression changes necessary for nutrient metabolism. To further dissect this mechanistic and physiological role of CrebA as a quick and transient responder to feeding, I used fly genetic tools to overexpress CrebA exogenously and mimic the refed state CrebA protein levels and check for feeding behavior changes. I found that even “hungry” flies overexpressing CrebA fed significantly less that controls which was not due to any disability to access food. Furthermore, I show that most of the CrebA expression in fly head is coming from the fat body, the major fly secretory organ hinting towards the

possibility that the core secretory machinery genes are a major player in downstream action of CrebA.

Overall, these data reveal a novel role for feeding dependent regulation of CrebA which, through its downstream action on the ER protein sorting machinery genes, orchestrates the metabolic state of the organism in response to food intake.

5.2. Introduction

Normal ER function is crucial to maintaining many of the metabolic processes necessary for survival. There is a tight interaction between the ER function and homeostasis and the nucleus via transcriptional regulation. A number of transcription factors that are physically associated with the ER. Presence of these transcription factors at the ER allows for a way to monitor ER state. Interestingly, there are a number of such transcription factors which have implications in metabolism. The Sterol regulatory element-binding proteins (SREBPs) respond to cholesterol levels in the ER to maintain cholesterol homeostasis (Shimano and Sato, 2017) (Moslehi and Hamidi-Zad, 2018). The X-box-binding protein 1 (Xbp1) is another such transcription factor that responds to ER stress from the influx of nascent polypeptides that form part of the unfolded protein response (UPR)(Fox and Andrew, 2015). But how do cells maintain ER homeostasis in response to changing nutritional cues? Our understanding of the impact of nutrition on ER and how ER manages the influx of nutrients remains limited.

CrebA and its mammalian orthologs, Creb3L transcription factors, lie at an interesting intersection of molecular functions in ER and response to nutrients. On one hand, CrebA has been shown to directly bind the promoters of the core secretory machinery genes (Fox et al., 2010) which lie at the ER and on the other hand some evidence suggests this transcription factor family might be regulated with nutritional cues (see section 3.5). Therefore, it would be interesting to take a deeper look in to what nutritional cues activate these transcription factors upon which they regulate ER homeostasis.

Clearly, a key downstream effector of transcriptional factor function is its ability to recruit the necessary machinery and initiate transcription of downstream targets. In eukaryotic cells, RNA Polymerase II (Pol2) is the molecular machine that transcribes DNA in to precursor messenger RNAs. Numerous studies trying to understand how transcription factors mediate their downstream function by coordinating with RNA Pol2 has provided crucial insights in to this long-standing mystery of gene regulation (Fuda et al., 2009). Genome-wide analysis of RNA Pol2 and transcription factor distribution across genomes has vastly helped in understanding these questions.

The transcription cycle is a multistep process with many rate-limiting steps. It begins with Pol2 being able to access to the promoter of a gene to be transcribed. This requires the site being cleared of nucleosomes so that the Pol2 is able to access the

region. Transcription factors play a role in both removing these physical barriers for Pol2 access and / or physically recruiting Pol2 to a promoter of a gene. For example, transcriptional activators at genes recruit chromatin remodeling enzymes to slide nucleosomes out of the promoter region. SWI/SNF complex is one such DNA-dependent ATPase that alters chromatin structure in this way (Peterson and Workman, 2000). In other cases, the promoter region is free of nucleosomes but requires Pol2 to be recruited to the location for transcription. The pioneer transcription factor PHA-4/FoxA responsible for pharynx organ identification during embryogenesis in *C. elegans* has been shown to perform both functions. In early development, PHA-4 binds promoters and recruits Pol2 to accumulate at transcription start sites while at a later development stage, it promotes chromatin opening (Hsu et al., 2015).

Another rate-limiting step in transcription by Pol2 is the release of accumulated and paused Pol2 at promoters. Studies in the 1980s of specific heat-shock genes in *Drosophila* showed that upon activation, paused Pol2 is released across the gene body in to productive elongation (Core and Lis, 2008). Later, the transcription factor P-TEFb was identified as one of the key mediators of this elongation release of Pol2 (Ni et al., 2008). Therefore, in order to understand gene expression changes in response to nutrients, it is critical to identify how these upstream mechanisms are regulated.

In this chapter, I describe findings on the *Drosophila* transcription factor, CrebA, which regulates a cell's secretory capacity in response to nutritional influx. I complement the results from the genome-wide gene expression data with Pol2 ChIP-seq data to report that CrebA acts as a key mediator of upregulating the expression of the ER protein sorting machinery genes in response to feeding. Interestingly, the findings described in this chapter suggest that CrebA likely promotes the release of paused Pol2 and not its recruitment in response to nutritional uptake. Collectively with the gene expression studies, these data suggest that CrebA mediated Pol2 release and transcription of its targets results in a function specific control of the ER protein sorting machinery genes. I also test weather CrebA levels directly alter fly feeding behavior. Taken together, this data reveals that CrebA plays an integral role in mediating metabolic homeostasis by altering secretion and feeding behavior.

5.3. Methods

5.3.1. Fly strains

Flies were raised on standard sugar-yeast-agar medium. For the majority of experiments, the 2202U w^{1118}, a Canton-S derivative strain was used as wild-type strain (Boynton and Tully, 1992). All transgenic flies used were backcrossed 6 generations to the 2202U wild-type background with the exception of the Sec61α RNAi. Other strains used included the Hsp70-Gal4 driver strain (Brand and Perrimon, 1993), nSyb-Gal4 (Bloomington stock center, 458), ppl-Gal4 (Bloomington stock center, 58768), TO-Gal4 (Dauwalder et al., 2002), UAS-CrebA (Rose et al., 1997), UAS-CrebAshRNAi (Vienna stock center, 330349) and UAS-Sec61αRNAi (Vienna stock center, 109660). Mated females were used for all experiments as virgins have been shown to be different in many biological aspects. 3-5 days old flies were frozen in liquid nitrogen and stored at -80ºC until used for further experiments. All experiments were done at 25ºC, 60% humidity on a 12h light:12h dark cycle.

5.3.2. Fly climbing assay

Fly climbing performance was measured using a rapid iterative negative geotaxis (RING) assay described previously (Gargano et al., 2005). 10 flies were transferred to a 50mL vial marked with a 5cm height mark using an air pipette. 8 vials were together placed in a rack specially designed to hold the vials. The rack together with the vials was tapped firmly to the bottom 3 times to bring all the flies to the bottom of the vials. Number of flies that climbed past the 5cm mark in 3 seconds were measured using a video camera. Percentage climbed plotted was measured as the number of flies that climbed 5cm in 3 seconds over total number of flies in a vial.

5.3.3. RT-qPCR

RNA extraction, cDNA preparations and RT-qPCR were all performed as described in **chapter 1**.

5.3.4. Generation of anti-CrebA antibody

Guinea-pig polyclonal anti-CrebA antibodies were generated against the whole CrebA protein (clone from Drosophila Genomics Resource Center, 1623052). The recombinant protein was expressed in a pETM11 vector in E.Coli at 37 degrees

Celsius for 3 hours. Expressed protein was collected from E,Coli and purified using a histidine tag column. Histidine tag was removed using a TEV cleavage site in between the CrebA protein and the histidine tag. TEV protease was incubated with the purified protein overnight at 4 degrees Celsius. Purified antigen was shipped to Eurogentec, Belgium where the animals were immunized and sera was collected under contract number DE17051.

5.3.5. Western Blotting

Western blots were performed as described in section 4.3.10. 5 fly heads were loaded for all westerns shown. Other than the Guinea-Pig anti-CrebA antibody generated, rabbit polyclonal anti-CrebA (1:200, DSHB) and anti-H2AZ (1,5:000) (Schauer et al., 2013) were used. Rabbit anti-apoltp was kindly provided by the lab of Suzan Eaton and used at 1:2000.

5.3.6. Immunoprecipitation

Immunoprecipitation assays were performed using Drosophila embryo extracts prepared as previously described (Becker and Wu, 1992). 100 grams of embryos were collected, washed and homogenized in homogenization buffer [15mM Hepes pH 7.6, 10mM KCl, 5mM MgCl2, 0.5mM EGTA pH 8.0, 0.1mM EDTA pH8.0 supplemented fresh with 1mM DTT, 0.2mM PMSF, 1mM NaMBS, 1µg aprotinin, 1µg leupeptin, 1µg pepstatin]. Nuclei were pelleted at 10,000g for 15 minutes and components of the nuclei precipitated using $(NH_4)_2SO_4$. 1ul of antibody was incubated overnight with the 500µg of the extract at 4ºC on a rotating wheel. 20ul of washed protein A Dynabeads were added and incubated at 4ºC for 3 hours. Beads were washed on a magnetic rack while on ice. Fraction of the reaction was used for western blots while the rest was further processed for mass spectrometry. Pre-immune sera from the same animals was used as a control.

5.3.7. ChIP-seq

ChIP experiments were done as previously described in Schauer et al., 2013. Using a combination of 630 and 400 microns sieves, a total of 1000-1500 heads were separated from frozen flies with constant quick freezing and vortexing. Heads were homogenized in homogenization buffer [350 mM sucrose, 15 mM HEPES pH 7.6, 10 mM KCl, 5 mM MgCl2, 0.5 mM EGTA, 0.1 mM EDTA, 0.1% Tween, with 1 mM DTT

and Protease Inhibitor Cocktail (PIC) (Roche) added immediately prior to use] at 4 degrees Celsius. Homogenization was performed at 2000rpm for a total of 20 times (up and down motion). Solution was then fixed using 1% formaldehyde and quenched with 2.5M glycine. Nuclei were filtered using a 60 microns nylon filter (Millipore, SCNY00060), washed 3 times with RIPA (150 mM NaCl, 25 mM Tris-HCl pH 7.6, 1 mM EDTA, 1% Triton-X, 0.1% SDS, 0.1% DOC, with protease inhibitors added prior to use) and sonicated using Branson 250 (2 cycles, intensity 5, pulsing 60 s) and Covaris S220 (PIP150, DC20, CPB200, time 10 min) sonicators. Soluble chromatin was taken as supernatant after centrifugation at full speed for 10 minutes. Ten ug aliquots were prepared and stored at -80ºC.

Ten µg chromatin was used for a single ChIP experiment. Chromatin was pre-absorbed with Sepharose protein A beads equilibrated with RIPA buffer containing 1 µg/µl salmon sperm DNA and 1 µg/µl BSA for 1 hour at 4ºC. Chromatin was incubated with the relevant antibody overnight on a rotating wheel at 4ºC. 10% input material was separated at this point and later pooled to form one sample per condition/time-point for sequencing. Pull-downs were performed using 20µl of equilibrated beads using either anti-CrebA (Guinea-Pig, own stock) or anti-RPB3 (Schauer et al., 2013). Immunoprecipitated DNA was purified using 1.8x AGENCOURT AMPURE XP magnetic beads (Beckmann-Coulter). At least 2 biological replicates were used for each experiment. In some cases, additional technical replicates were done in order to obtain the required amount of chromatin.

The ChIP-seq libraries were prepared with 1 ng of ChIP and input DNA with NEBNext®Ultra II DNA Library Prep Kit for Illumina® according to the manual instructions. The libraries were barcoded using NEBNext® Multiplex Oligos Set 1 and 2 and sequenced at LAFUGA at the Gene Center (LMU) using an Illumina® HiSeq1500 sequencer.

5.3.8. ChIP-seq analysis

Raw sequencing files were demultiplexed using the relevant barcode information and checked for quality using fastqc. ChIP-seq 50bp single end reads were aligned to the reference *Drosophila melanogaster* (dm6) using bowtie2 (version 2.2.9). SAM output files were converted to BAM files for further processing using samtools. macs2 (version 2.1.1) was used to call peaks that were above the fold change cut-off of 10 over the input. Reads under the regions were calculated using subread (version 1.6.2).

BAM files were read in to R using readGAlignmentPairs from the GenomicAlignment package to make coverage vectors. Coverages were normalized by dividing by total number of reads and multiplied by a million. Normalized coverages were centered in windows around peak summits using the coverageWindowsCenteredStranded from the tstools package. Heatmaps were created using R base graphics. Regions were associated to nearest genes using RGmatch. Differences between samples were called using a generalized linear regression model which was made for the nutritionally responsive (1588) genes using the MORE (Multi-Omics Regulation) package in R (https://github.com/ConesaLab/MORE). Gene expression data was used as the response variable. Counts under associated peaks (if any) to these genes for bound CrebA and Pol2 signal over gene body minus TSS was used as the two explanatory matrices. Because Pol2 signal preceded mRNA signal in terms of time, a second model was built with a time shift i.e., Pol2 signal at 4 hours matched with mRNA signal at 6 hours in order to pick out such regulators. Positive Pol2 regulators were selected based on FDR adjusted p-value < 0.05. GO term analysis was done using a Fischer's Exact Test and the database from biomaRt (version 2.38) in R. Bound CrebA measurement was calculated from number of reads present in peaks regions annotated to their respective genes. Pol2 over gene body or Pol2 minus TSS measurements were calculated as the number of reads over the entire gene minus the TSS (first 200bp) or only over TSS (first 200bp) respectively.

5.4. Results

5.4.1. Feeding regulates the *Drosophila* transcription factor, CrebA

In order to understand what could potentially drive the coordinated control of the ER protein sorting machinery genes, I used our transcriptomics dataset to search for changes in the expression of transcription factor at earlier timepoints. I hypothesized that a single regulator might be responsible for this coordinated control of ER protein sorting machinery genes.

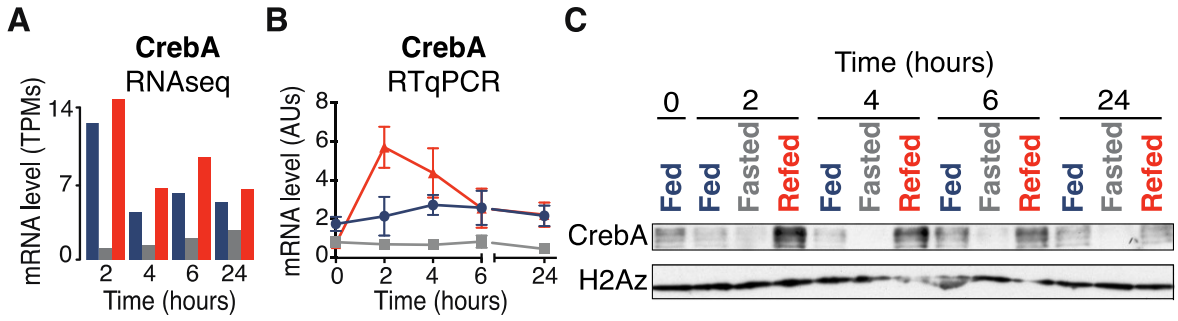

Figure 5.1 Feeding regulates the *Drosophila* transcription factor, CrebA

(A) Barplot showing CrebA expression levels plotted as transcripts per million (TPM) during the refeeding timecourse (n=2) **(B)** Normalized CrebA levels measured by RT-qPCR and normalized to the housekeeping gene, H2A.Z. RT-qPCR data are means ±SD for n=3-5. Each n represents mRNA extracted from 30 female fly heads collected on a different day. **(C)** Representative western blots showing CrebA and H2A.Z (loading control) protein levels from female fly heads.

Transcript of CrebA, the *Drosophila* ortholog of the Creb3 family of transcription factors, was identified to be transiently upregulated after feeding (Figure 5.1A). As discussed earlier, CrebA has been implicated in regulating the core ER protein sorting machinery genes during development (Fox et al., 2010) thus making it a likely contender for regulating such genes in response to nutrients. RNA-seq data was complemented with RT-qPCR data and replicated the 2-hour transient upregulation of CrebA transcript upon refeeding (Figure 5.1B). Changes in mRNA levels translated to CrebA protein levels showing similar kinetics upon feeding / fasting (Figure 5.1C). Interestingly, multiple CrebA bands were seen in western blot experiments suggesting the protein may be post-translationally modified. Overall, the quick and transient activation of CrebA together with previous data showing regulation of some ER protein sorting machinery genes positions CrebA as a prime candidate to regulate these genes in response to nutrients.

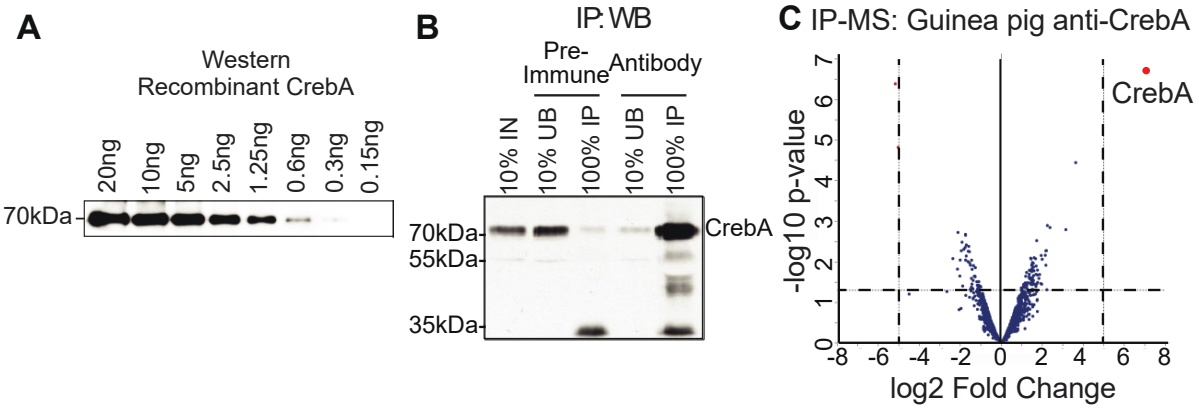

Figure 5.2 Polyclonal Guinea Pig CrebA antibody is sensitive and specific to CrebA.

(A) Recombinantly expressed CrebA runs at 70kDa and is detectable to sub- nanogram levels using Guinea Pig anti-CrebA antibody. **(B)** Western blot loaded with Immunoprecipitation (IP) pull downs of CrebA using *Drosophila* embryo extracts. Preimmune serum (control) was used as a control. **(C)** Volcano plot showing the analysis of LC-MS based quantification of IP experiments shown in (B). Data plotted are Log Fold Change intensities from anti-CrebA pull down over pre-immune serum pull down for n=3.

To test this hypothesis, I monitored CrebA activity via its binding to chromatin using genome-wide chromatin-immunoprecipitation assays (ChIP-seq) of CrebA. In order to do so, I first generated a polyclonal CrebA antibody. I purified CrebA protein by expressing a CrebA-His tag containing construct in *E. Coli*. The His-tag was cleaved off using a TEV cleavage site in between. Purified protein was then sent to Eurogentec where rabbits and guinea pigs were injected with the antigen. The antibodies obtained were then validated first using western blots on the purified protein. The generated antibody in guinea pig was able to detect purified CrebA protein to sub-nanogram levels (Figure 5.2A). Interestingly, even the bacterial purified CrebA protein migrated at 70kDa compared to its molecular weight of 55kDa suggesting this could be due to the natural migration dynamics of the protein in SDS-PAGE. The antibody was then further validated using immunoprecipitation mass spectrometry (IP-MS) assays. CrebA antibody was able to successfully pull down CrebA protein from transcriptionally active embryo extracts (Figure 5.2B). MS on these samples identified CrebA peptides as the most enriched compared to control pull-downs (Figure 5.2C).

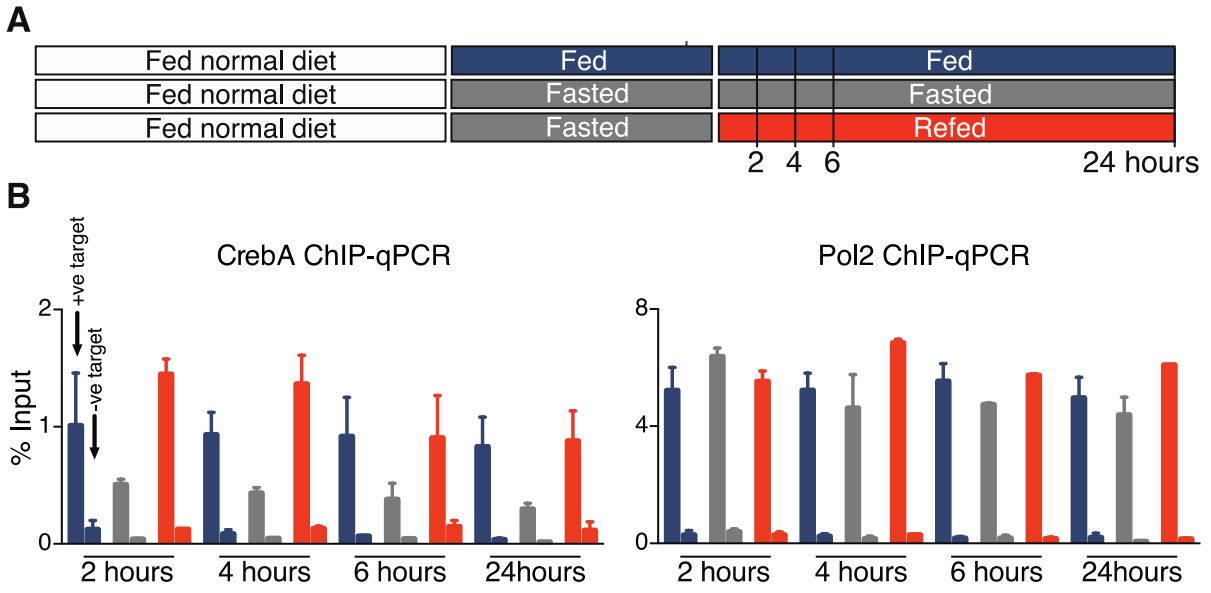

Figure 5.3 ChIP-qPCR using CrebA and Pol2 antibodies shows feeding dependent changes in CrebA binding but not in Pol2 occupancy.

(A) Feeding regiment outline used for sample collections for ChIP analysis. **(B)** ChIP-qPCR data plotted as % input calculated for enrichment over positive (Sec61B for CrebA and fas2 for Pol2) and negative (sequence upstream 1kb of respective promoters) targets. Data plotted for n=2±SEM.

Next, I assayed for genome-wide CrebA binding using Chromatin Immunoprecipitation followed by qPCR over potential target genes. I prepared chromatin from normally fed, fasted or refed female fly heads at 2, 4, 6 and 24 hours (Figure 5.3A) as done previously (chapter 4). I hypothesized that CrebA binding would reflect changes in CrebA protein levels in different nutritional states. CrebA was previously reported to bind in the promoter region of Sec61β (Fox et al., 2010). Therefore, I primers over the Sec61β promoter as a positive target sequence and a sequence 1kb upstream as a negative target sequence to check for CrebA binding. Interestingly, CrebA binding accurately reflected CrebA protein levels both consistently low in the fasted state and peaking at 2 hours after refeeding (Figure 5.3B, left panel). I confirmed that this response was specific to CrebA binding and not with global chromatin changes by assaying for Pol2 occupancy over the fas2 gene. I used an antibody against the RPB3 subunit of Pol2. Since the antibody does not specifically detect any paused or elongation phosphorylation signals, I expected my ChIP-seq data to detect any Pol2 occupancy signal within the detection limit of ChIP-seq technology. ChIP-qPCR data on a known Pol2 binding site did not show any changes upon refeeding (Figure 5.3B, right panel).

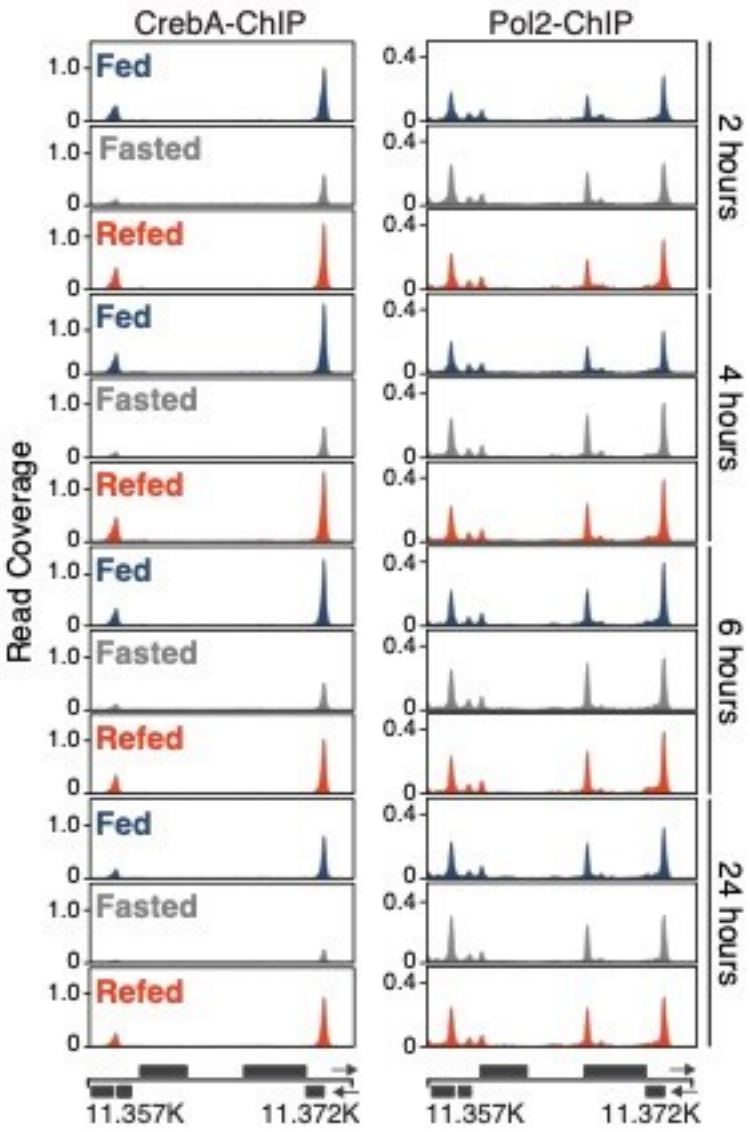

Figure 5.4 CrebA binding depends on feeding state.

Coverage profiles shown for CrebA ChIP-seq (left) and Pol2 ChIP-seq (right) along the feeding regiment. Region plotted is 2R:11.357.693-11.372.933.

To assess global CrebA binding changes upon feeding, I sequenced CrebA ChIP samples using next generation sequencing. As expected, we found that global CrebA binding was highly dependent on feeding time and status (Figure 5.4). Pol2 ChIP-seq from the same chromatin samples showed no significant changes in paused Pol2 occupancy in the same region indicating that only CrebA occupancy is affected in these samples.

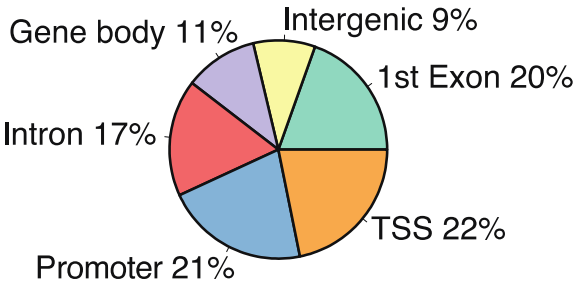

Figure 5.5 CrebA binding sites mainly localize to promoter regions.

Genomic distribution of CrebA bound regions.

Next, I called peaks from all sequenced CrebA ChIP experiments to define a peak set that contained all possible regions of interest (ROIs) using macs2 (Zhang et al., 2008). I used a fold change threshold of 10 (IP/Input) to select for the most convincing CrebA ROIs. These included 404 CrebA binding sites across the whole genome. As expected, CrebA binds primarily at promoters, transcription start sites (TSS) or within the 1st exon of the gene (Figure 5.5).

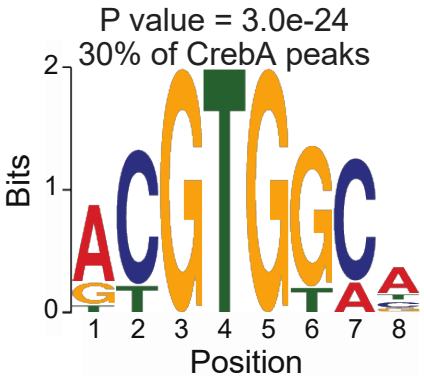

Figure 5.6 Known CrebA consensus motif is enriched under the peaks.

MEME-ChIP analysis of 404 CrebA bound regions. P-value obtained from MEME-ChIP suggesting with a high probability that at least 30% of total CrebA peaks contain the motif.

I performed an unbiased *de novo* search to identify if I can identify the CrebA consensus motif. CrebA peak regions were enriched in the previously identified CrebA motif (Figure 5.6) (Abrams and Andrew, 2005; Johnson et al., 2020; Nitta et al., 2015). These results further confirmed that these experiments efficiently pulled down and enriched CrebA bound DNA regions.

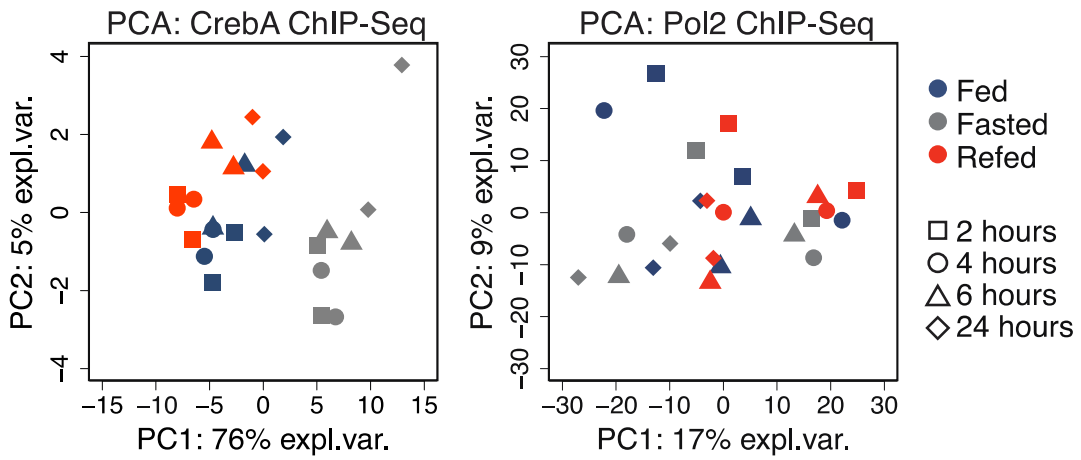

Figure 5.7 Principal component analysis changes in CrebA binding but not in Pol2 occupancy.

PCA plots for CrebA ChIP-seq (left) and Pol2 ChIP-seq (right) data depicting the first two components. Plotted signals are log2 transformed normalized counts under peaks.

Next, I quantified number of reads under all CrebA and Pol2 ChIP ROIs and performed a principal component analysis (Figure 5.7). Overall, biological replicates clustered together showing high reproducibility samples. Additionally, PC1 for CrebA ROIs which accounted for 76% explained variability, clearly separated the fed conditions from fasted samples indicating major differences were condition dependent. PC2 separated earlier timepoints in the refed state from the later ones. Pol2 ROIs however, showed no overall differences between the three different conditions.

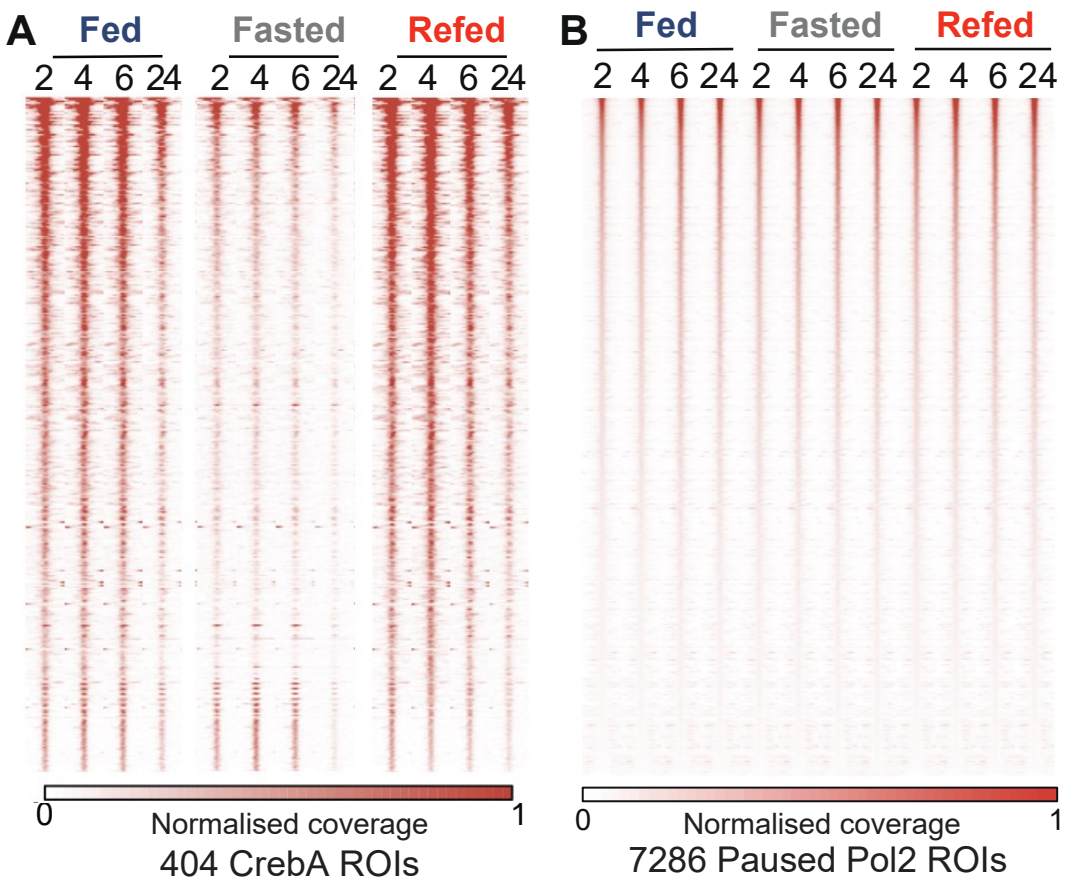

Figure 5.8 Global CrebA occupancy peaks at 4 hours after refeeding.

(A) Heatmap showing CrebA occupancy centered around +/-750 bp for the 404 CrebA peak regions from female fly heads along the feeding regiment. Centering was performed around the max peak region presented in decreasing order based on the signal at refed 4-hours. **(B)** Similar heatmap as in (A) plotted for data from pull downs using an antibody against the Rpb3 subunit of Pol2 complex. Plotted Pol2 ChIP-seq regions are +/-750bp of 7286 Pol2 called peaks ordered in decreasing order based on the signal at refed 2-hours.

To further assess CrebA and Pol2 global occupancy over its ROIs, I plotted reads under all ROIs as a heatmap centered on peak summits. Alterations of CrebA protein levels in response to nutrients results in change of bound CrebA on its target regions (Figure 5.8A). Highest occupancy was seen at 4 hours after refeeding. Are these differences in CrebA occupancy a result of differences in chromatin preparations? To answer this question, I performed a control ChIP-seq experiment from the same chromatin using a RNA Polymerase II antibody. Bound RNA Pol II was seen to be consistently bound across different region of interests in different feeding conditions (Figure 5.8B).

5.4.2. CrebA mediates the control of ER protein sorting machinery gene expression upon nutrition

CrebA binding to its targets is dependent on the nutritional state. Next, using the same CrebA ChIP-seq data, I asked if CrebA bound regions were enriched with target genes part of the ER protein sorting machinery genes.

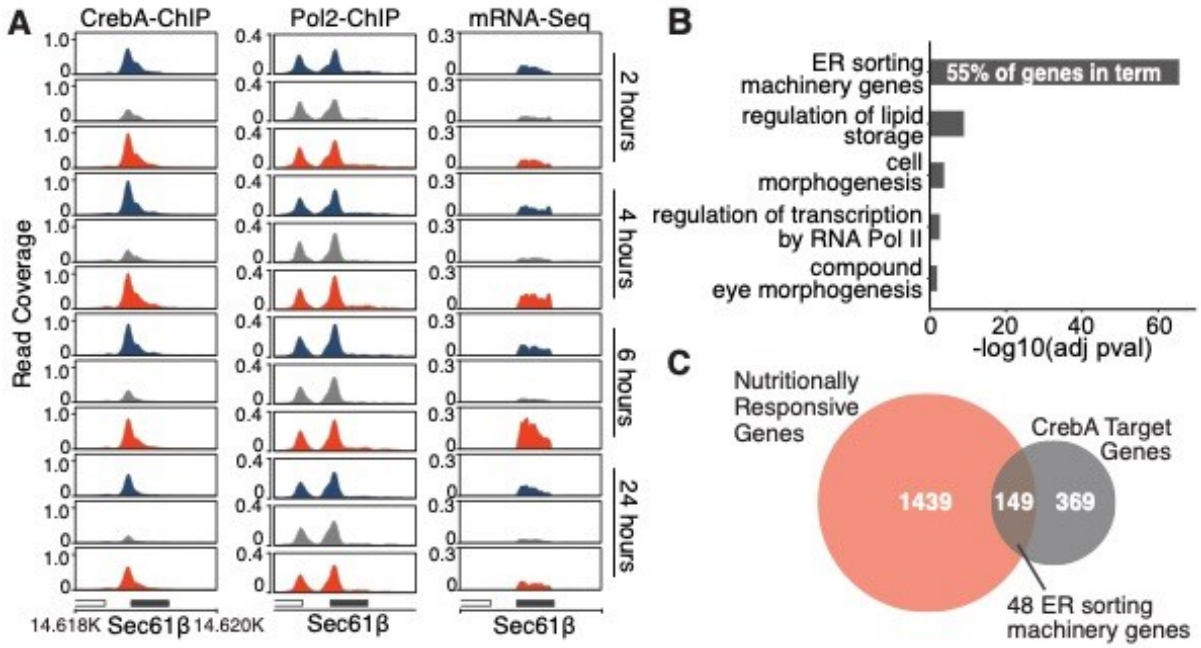

Figure 5.9 CrebA targets are enriched for ER protein sorting machinery genes.

(A) Coverage profiles plotted for CrebA ChIP-seq (left), Pol2 ChIP-seq (middle) and RNA-seq (right) profiles from female fly heads over Sec61β. Region plotted is 2R:14.617.400-14.620.853. **(B)** Gene ontology analysis of all CrebA targets using a Fischer's-Exact test. Top 5 significant GO-terms are shown from a list of GO terms identified with FDR <0.05. **(C)** Overlap of 1588 nutritionally-responsive transcripts and 518 CrebA target genes shows an overlap of 149 genes. 48 of the 149 overlap genes belong to the ER protein sorting machinery genes.

Promoter regions of all three subunits of the Sec61 translocon were highly bound by CrebA in a feeding dependent manner (Figure 5.9A). Pol II promoter peaks did not show any changes in response to feeding. In order to identify which other ER protein sorting machinery genes were bound by CrebA, I annotated the 404 CrebA binding sites to identify 518 CrebA target genes that had CrebA bound peaks within 1500bp of the core promoter region. An unbiased GO term analysis of these genes revealed the ER protein sorting machinery genes as the most significant hit with 55% of genes annotated within the term present in our dataset (Figure 5.9B). Interestingly, not all CrebA targets were nutritionally regulated. Analysis showed that only 149 CrebA targets overlapped with the 1588 nutritionally responsive genes from the

transcriptomics data (Figure 5.9C). Remarkably, these were highly enriched with the ER protein sorting machinery genes. Out of the 149 genes, 48 were genes involved in secretion via the ER. Furthermore, all 149 overlapping genes were upregulated upon refeeding. Together, this data suggested that CrebA acts as a master regulator of the ER protein sorting machinery genes in response to nutrients.

5.4.3. Pol2 signal over gene body can be an accurate predictor of gene expression

Production of mRNA depends on the rate of RNA polymerase 2 elongation over gene bodies. Paused Pol2 in *metazoans* is an important prerequisite for quick and transient control of transcription and is extremely stable as shown with extensive kinetic studies done *in vivo* (Buckley et al., 2014). This mechanism allows for controlled transcriptional regulation by transcription factor induced entry of paused Pol2 in to elongation.

I sought to ask if Pol2 signal over the gene bodies can be used as a mark for active transcription. In order to do so, I asked if the analysis of Pol2 signal over gene bodies excluding the paused promoter peaks could reveal any significant changes that correlate with changes in gene expression.

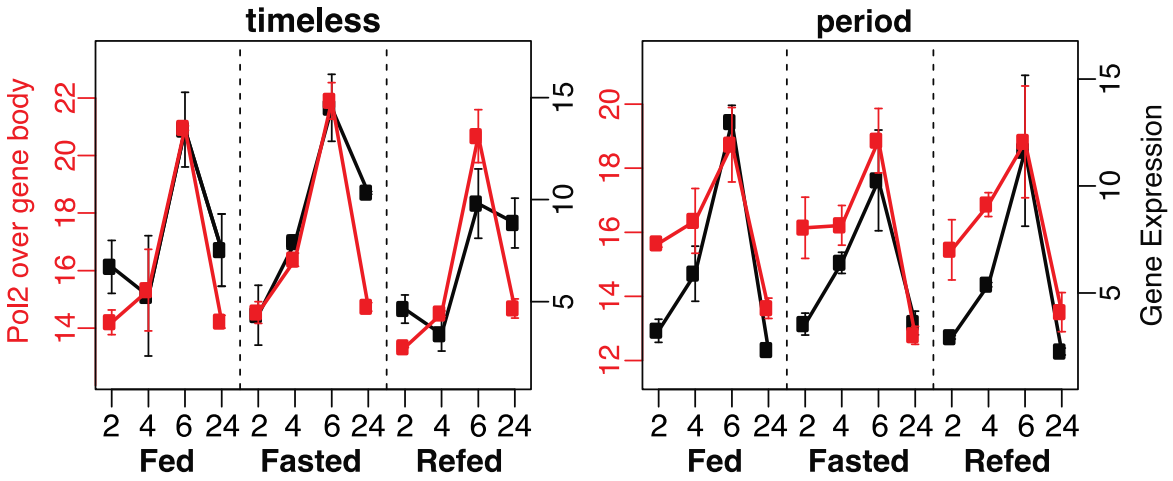

Figure 5.10 Gene expression and Pol2 elongation changes of *timeless* and *period* genes, two circadian rhythm regulators.

Both timeless and period show a similar circadian dependent oscillation of gene expression and Pol2 elongation changes. Red indicates Pol2 signal over gene body while black indicates mRNA expression levels. Data plotted is normalized counts.

Drosophila, similar to mammals, exhibits a circadian rhythm modulated by the light and dark cycle. Timeless and period are two key genes that are required for this

circadian rhythmicity in fruit flies (Sehgal et al., 1995). I asked if, using our detailed timecourse analysis, I can detect changes in mRNA and Pol2 elongation over these circadian regulators. Interestingly, gene expression levels of both *timeless* and *period* transcripts oscillated in a circadian manner that was independent of feeding (Figure 5.10). Additionally, Pol2 occupancy over these genes also showed a similar oscillation giving the first insight that this measurement can be used as a marker for active transcription.

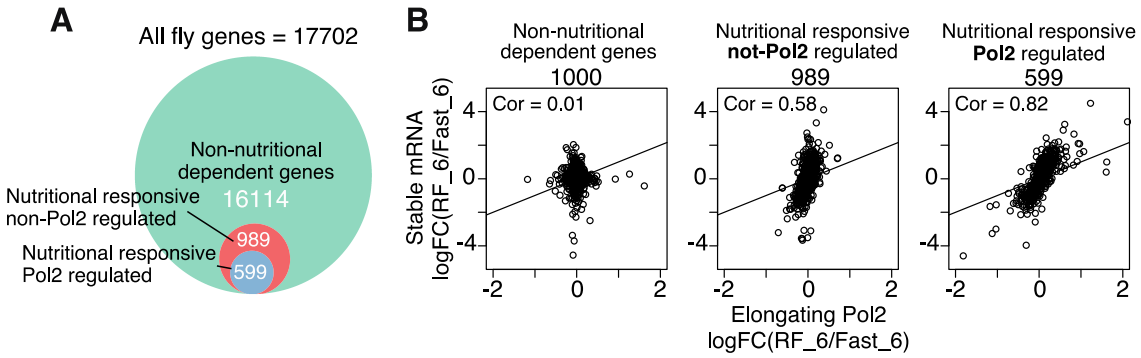

Figure 5.11 Pol2 elongation changes over gene body is an accurate predictor of mRNA changes.

(A) Venn diagram showing the total number of genes in flies (17702) out of which 16114 are not nutrition dependent (Non-nutritional dependent genes), 989 are differentially expressed but not significant in the GLM model (Nutritional responsive not-Pol2 regulated) and 599 are differentially expressed and significant in the GLM model (Nutritional responsive Pol2 regulated). (B) Log fold change differences between refed 6-hours over fasted 6-hours of stable mRNA level (y-axis) and Pol2 signal over gene body (x-axis). Corresponding values for randomly selected 1000 non-nutritional genes (left), nutritional non-regulators (middle) and nutritional positive regulators (right) identified using the MORE generalized linear regression model. Spearman correlation values shown.

Next, I asked if global changes in mRNA levels can be explained by changes in Pol2 elongation. In order to do so, I used the already identified 1588 set of feeding dependent genes and asked if Pol2 changes over the gene bodies are able to predict the changes in mRNA levels. I hypothesized that the 1588 nutritionally responsive genes would show similar changes in Pol2-ChIP signal when looked over the gene bodies. For this purpose, I used the MORE package in R and built a generalized linear regression model (GLM) to statistically model if changes in gene expression upon nutrition can be explained as a change in Pol2 signals over gene bodies. Interestingly, out of the 1588 genes that changed with nutrition, Pol2 changes in 599 of those genes showed similar patterns and were called out to be significant under the GLM model (Figure 5.11A). I called these genes nutritional responsive Pol2 regulated genes. But what happens to Pol2 signal over the other nutritional responsive genes that are not

picked by the model as being regulated by Pol2 signal (989 genes referred to as Nutritional Responsive not-Pol2 regulated genes)? To asses this, I plotted the changes between mRNA and Pol2 using correlation plots. As a negative control, I randomly selected 1000 non-nutrition dependent genes (no changes in mRNA levels). The non-nutrition dependent genes showed no correlation between the changes in the mRNA vs Pol2 (Spearman correlation = 0.01; Figure 5.11B). As expected, the genes that were significantly regulated at the mRNA level and also Pol2 signal (Nutritional Pol2 regulated genes) showed a high correlation (0.82). Interestingly, the 989 genes (Nutritional Responsive not-Pol2 regulated genes) that were not picked out by the GLM model also showed a relatively high correlation (0.58). Clearly, most of the gene expression changes during the timecourse can be explained using the Pol2 elongation over gene bodies but not all of them are significant when assessed using the GLM model. This could be due to the lack of resolution of the ChIP-seq technology or other factors such as RNA degradation that may also have a role in determining the stable mRNA levels for these nutritionally responsive genes. However, it is clear that based on the large amount of input data used to build the GLM model, it can pick out very small significant changes in Pol2 elongation.

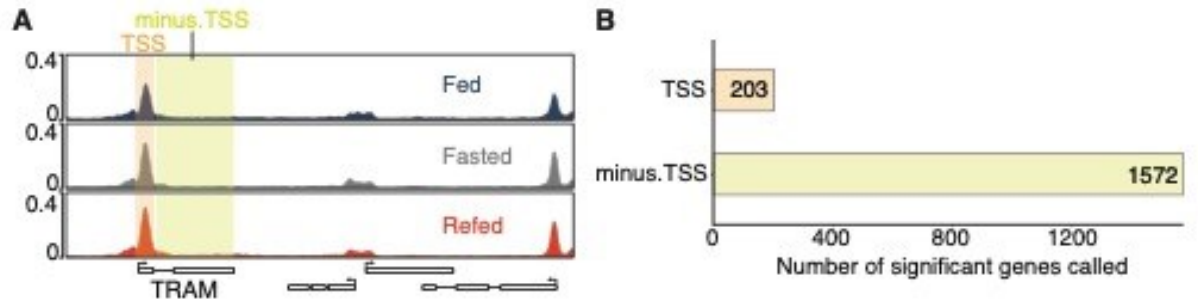

Figure 5.12 Pol2-ChIP changes are specific to elongating polymerase.

(A) Division of regions of paused and elongating polymerase shows differences only in signal over gene body. (B) Number of significant genes called by the GLM model when compared between fasted and fed conditions. Only 203 genes are called to be significant when Pol2-ChIP signal over TSS was used compared to the 1572 genes called when signal over the gene bodies was used.

All of the above analysis was done using the Pol2 signal only over the gene bodies. This was calculated by subtracting the first 200bp (TSS) from each gene. In order to further strengthen these findings, I asked if the changes are particularly specific to elongating Pol2 or if there are also changes over the paused Pol2. For this, I calculated Pol2-ChIP signals over the paused Pol2 regions (first 200bp of each gene; Figure 5.12A). Based on a manual search through the genome, I hypothesized that

Pol2-ChIP changes are specific to elongating polymerase. Using the same GLM model, I asked over how many genes Pol2-ChIP signal is significantly different between fed and fasted conditions. Interestingly, only a small number of changes were called significant in paused Pol2 state (203) compared to the elongating polymerase (1572) (Figure 5.12B). This suggested that the changes in Pol2-ChIP are primarily specific to the elongating state rather than the paused state. Overall, these data suggest that the Pol2 ChIP-seq in gene bodies can be used as a mark for active transcription for the nutritional responsive genes.

5.4.4. CrebA drives transcriptional expression by regulating RNA polymerase II elongation

In this chapter, I show that CrebA is acutely regulated by nutrition upon which it binds its target genes to regulate transcription. Interestingly, only 149 of the 518 CrebA target genes identified are regulated by nutrition. I asked if these nutritionally responsive targets of CrebA show differences in active transcription levels (Elongating Pol2) when compared with the CrebA targets that are not nutritionally responsive (Figure 5.13A). I calculated Pol2 differences over the gene bodies between fasted and refed samples for all genes in each group. Plotting this data between the two groups revealed a clear and significant difference between the nutritional and non-nutritional gene groups (Figure 5.13B). Interestingly, this change was time-dependent. A very small difference at 2-hours increased at 4 and 6 hours and went back to no change at 24 hours. Overall, this suggested that upon early CrebA binding Pol2 elongation was significantly increased over nutritionally responsive CrebA targets.

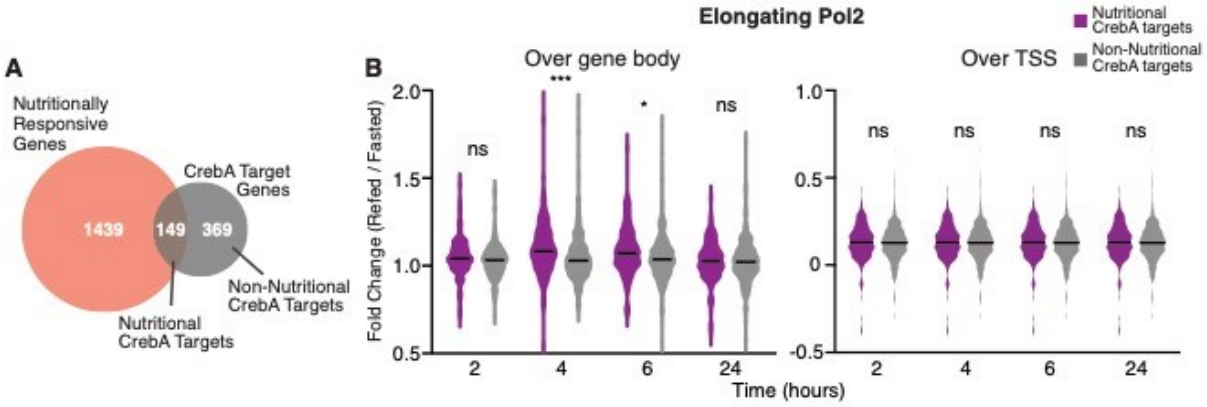

Figure 5.13 CrebA target genes are regulated by Pol2 elongation.

(A) Venn diagram showing nutritional and non-nutritional CrebA targets. (B) Volcano plots showing the fold change differences between fasted and refed elongating Pol2 over gene body

minus TSS (left) and over TSS (right) during the time-course in nutritional CrebA targets (purple) and non-nutritional CrebA targets (grey).

Next, I asked if CrebA mediated activation of the ER protein sorting machinery genes correlates with Pol2 elongation and transcription. I used the CrebA peak regions annotated to the 48 ER protein sorting machinery genes and plotted these signals together with elongating Pol2 over these genes (minus TSS) and the transcript data (Figure 5.14).

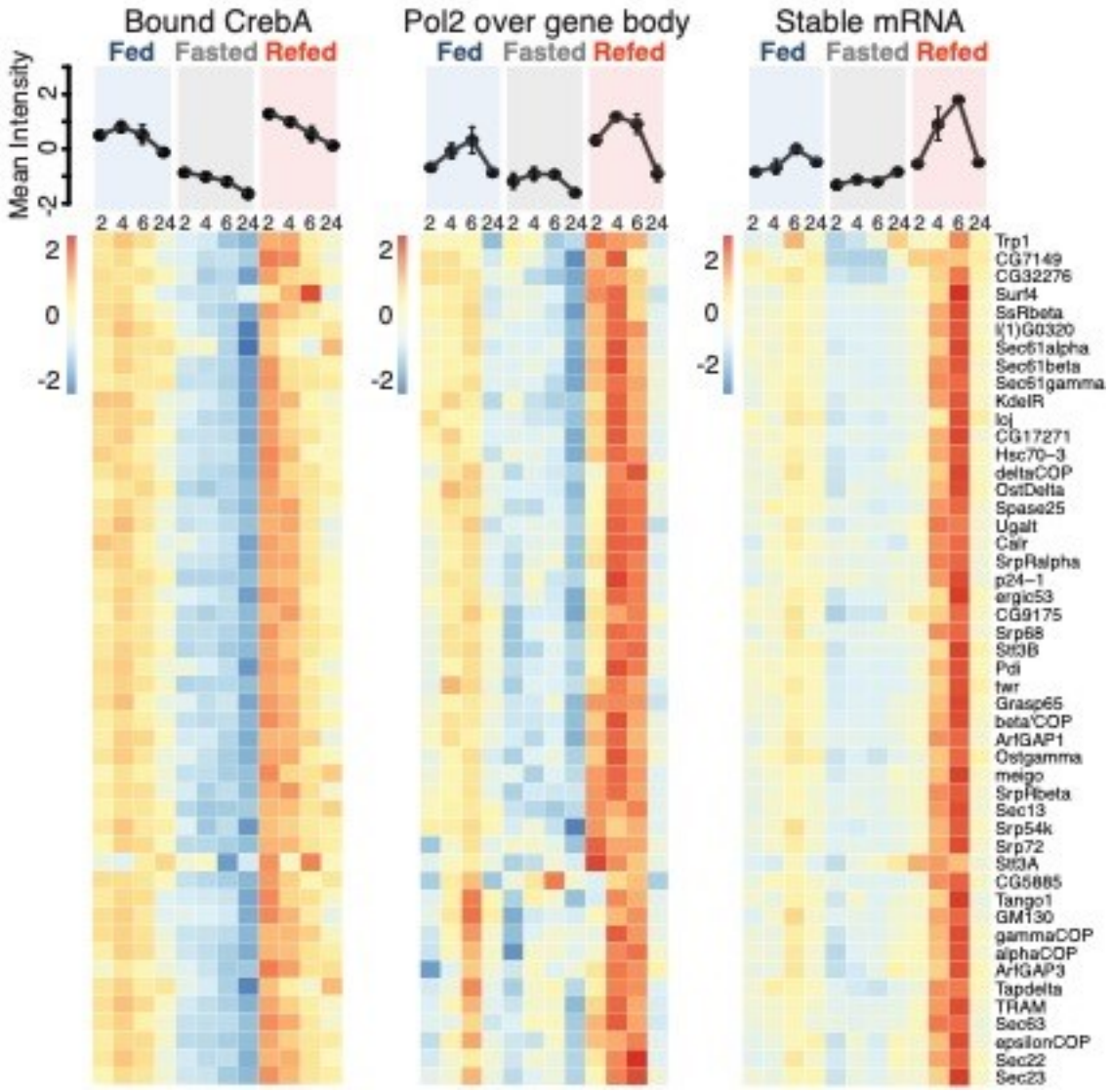

Figure 5.14 CrebA initiates the transcription of ER protein sorting machinery genes in response to nutrients.

Heatmap of CrebA binding, Pol2 signal over gene body and mRNA levels plotting individual z-scores for the 48 ER protein sorting machinery genes. Heatmaps are hierarchically clustered by rows performed on Pol2 signals. Line graphs on top show mean z-scores for all genes.

Collectively, CrebA binding over the 48 ER protein sorting machinery genes was the highest at 2 hours. Pol II levels over gene bodies of these genes peaked at 4 hours

and mRNA levels followed at 6 hours (Figure 5.14). Interestingly, this data was able to capture the time-dependent changes in CrebA driven and Pol II mediated transcription control of the ER protein sorting machinery genes. Next, I asked if the Pol II ChIP-seq data can be used to assess if CrebA drives Pol II recruitment or Pol II elongation.

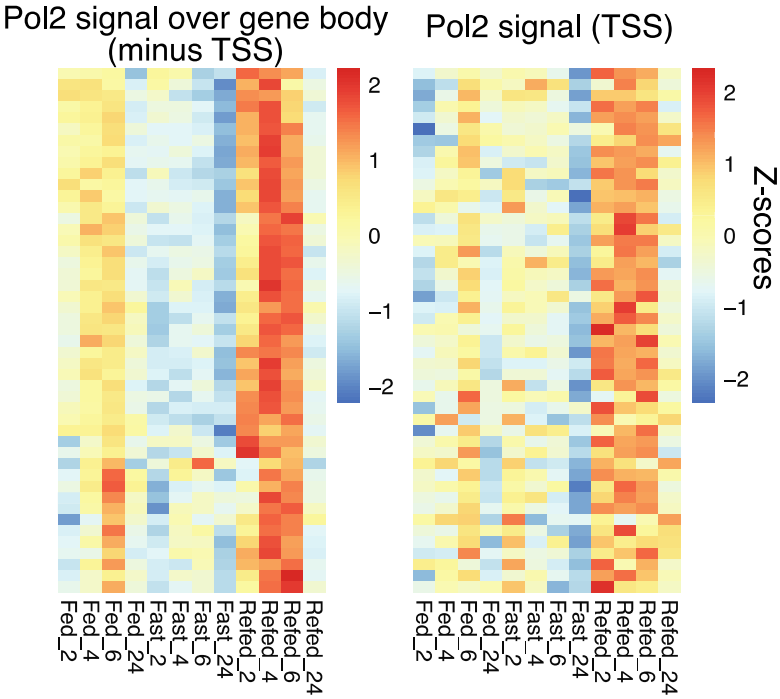

Figure 5.15 Pol2 changes are concentrated over gene bodies.

Heatmap showing mean z-scores plotted for each of the 48 ER protein sorting machinery genes of Pol2 signal over gene body (left) without the 200bp around the TSS and only at the TSS (right). Hierarchical clustering order same as in Figure 5.14.

As discussed earlier, Pol II can be in a poised state before elongating over gene body to transcribe DNA in to mRNA. Therefore, I dissected Pol II signal in to two different measurements i.e., Pol II over gene body and Pol II over TSS. Using heatmap plots over the ER protein sorting machinery genes, I analyzed Pol2 levels over gene body which showed more dramatic changes compared to the ones over TSS. This suggested that CrebA may act as a release switch for Pol II elongation. Clearly, this data offers the most comprehensive time- and condition-dependent insight in to how CrebA acts as key metabolic regulator of the core ER protein sorting machinery genes.

5.4.5. CrebA overexpression mimics satiation

Organisms have developed mechanisms of satiation that induce cessation of feeding once it has ingested enough nutrients. These processes are usually mediated by endocrine hormones that stimulate signaling to the brain. The female-specific independent of transformer (fit) is one such hormone in *Drosophila* that induces satiety (Sun et al., 2017). But how are these processes regulated at the transcriptional level?

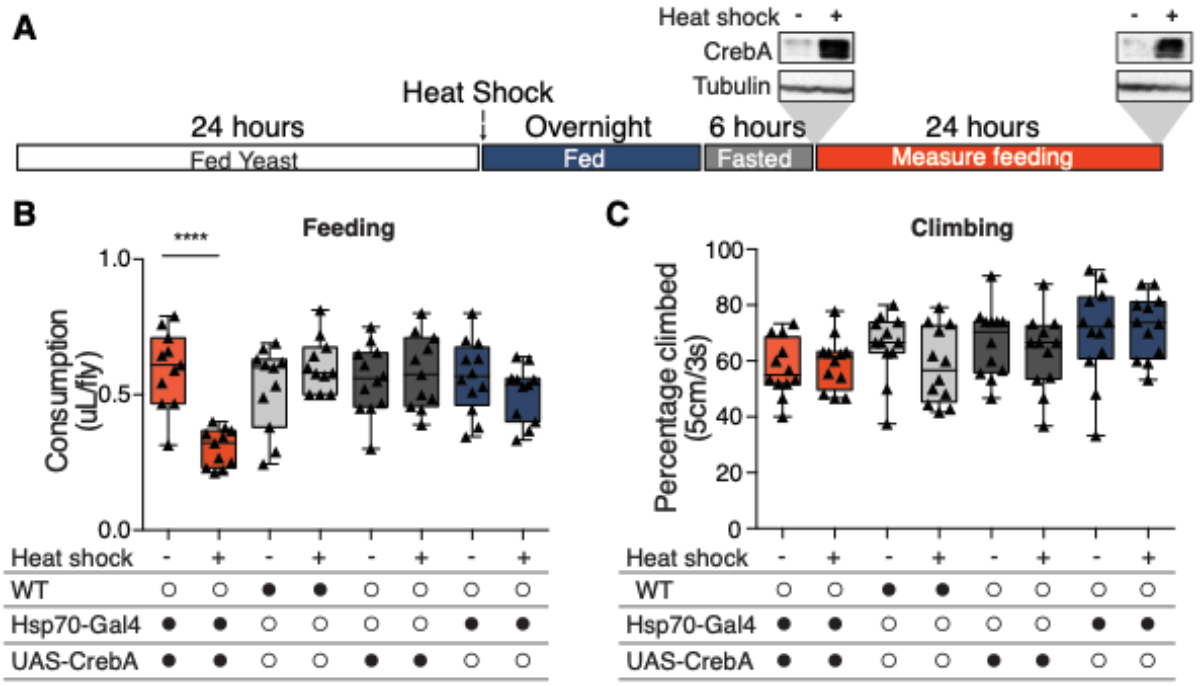

Figure 5.16 CrebA levels alter feeding in flies.

(A) A schematic representation of fly treatment over the course of the experiment. Flies were heat-shocked for 20 minutes in glass vials by placing them in a water bath at 36ºC. **(B)** Food consumption upon overexpression of CrebA was measured after 24 hours using the CAFE assay. Data (n = 12 biological replicates each representing 10 female flies) plotted as box-and-whiskers plot (min to max), analyzed with Kruskal-Wallis and post-hoc Dunn's multiple comparisons test (*p<0.03;** p<0.002; ***p<0.0002). **(C)** Fly climbing ability was measured using the RING assay. Data (n=12) plotted as box-and-whiskers plot (min to max).

Data on CrebA mRNA and protein levels showed that the transcription factor is clearly upregulated upon initial refeeding. Since CrebA acts as a major regulator of the secretory machinery including the secretion of signaling hormones / neuropeptides, I hypothesized altering CrebA levels exogenously may affect fly feeding behavior. Overexpressing CrebA levels, for example, may mimic a "refed" fly state which results in less feeding over time. In order to study this, I used the bipartite Hsp70-Gal4 system (Brand and Perrimon, 1993). Two different fly strains, UAS-CrebA and UAS-CrebARNAi were used to overexpress or knockdown CrebA protein upon heat shock in adult flies. Flies were initially fed on standard fly food for 24 hours, heat shocked and allowed to recover for 16 hours overnight (Figure 5.16A). In order to induce hunger and be able to measure food consumption, I fasted flies for 6 hours before measuring feeding with CAFÉ (see chapter 1). Overexpression of CrebA was confirmed at the start and end of feeding measurements was probed using western blot experiments. Intriguingly, flies overexpressing CrebA ate significantly less than the control flies

(Figure 5.16B). Control groups included non-heat shocked genetically identical flies, wild-type flies and only Hsp-70 Gal4 or UAS-CrebA flies. All control flies did not show any differences in feeding with or without heat shock. Further, in order to confirm flies overexpressing CrebA were not sick or had any performance issues, I assessed their ability to climb using a climbing assay (Figure 5.16C). There were no differences observed in climbing ability between different genetic groups and conditions. Collectively, these results suggested CrebA suppresses fly food consumption.

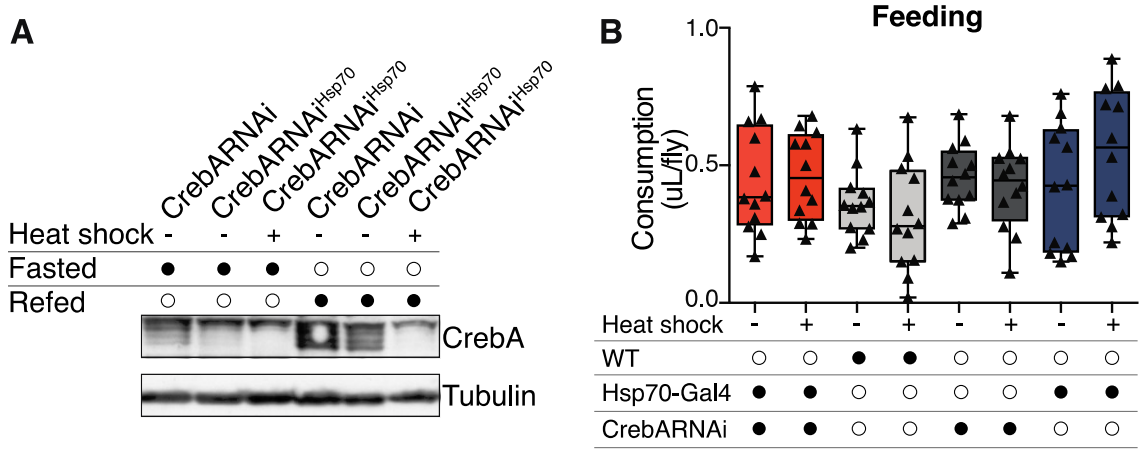

Figure 5.17 Absence of CrebA does not affect feeding.

(A) Western blot showing CrebA levels upon Hsp70-Gal4 driven UAS-CrebARNAi in fasted and refed animals. **(B)** Food consumption of flies at 24 hours upon knockdown of CrebA. Data plotted are for 12 biological replicates each representing 10 female flies.

Next, I assessed if CrebA knockdown flies would show an opposite phenotype and eat more than the controls. Again, I used Hsp70-Gal4 driver to knock down CrebA using a short hairpin CrebA RNAi fly line (Dietzl et al., 2007). I confirmed the efficiency of the knockdown using a western blot (Figure 5.17A). When flies were refed, CrebARNAi was able to efficiently knock down CrebA. However, in the absence of a heatshock there was already a significant decrease in CrebA protein levels suggesting that the Hsp70-Gal4 is leaky at 25 degrees Celsius. In fasted state, these differences were less strong due to already low levels of CrebA reported (Figure 5.1). Interestingly, I did not see any differences in feeding upon CrebA knockdown suggesting CrebA may only function in regulating downstream anorexigenic signals.

5.4.6. CrebA is predominantly expressed in fat body cells

Drosophila contains multiple high secretory organs that have high CrebA expression. But which cell-type(-s) play a physiological role in regulating secretory demand in response to feeding? All our data was collected from experiments done on fly heads which constitutes multiple cell types including neurons, glia and fatbody cells. In order to dissect the cell-type specific role of CrebA, I used fly genetic tools to knock down CrebA in different head cell types and probe for CrebA protein levels.

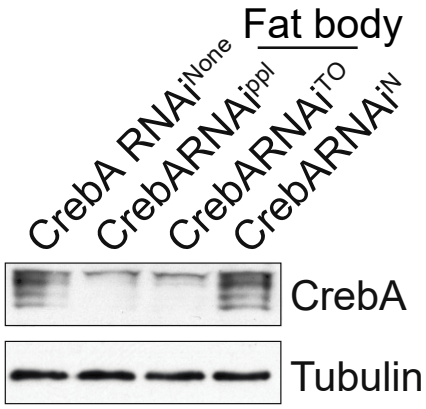

Figure 5.18 CrebA protein is mainly expressed in the fatbody.

Western blot showing CrebA and Tubulin protein levels upon the cell-type specific knockdown of CrebA. TO and ppl were used as fatbody drivers. Nsyb (N) was used as a neuronal driver.

Fly head samples were collected at refed 2-hours in order to elevate CrebA to the highest levels. I used two fat-body drivers takeout (TO) (Dauwalder et al., 2002) and pumpless (ppl) (Zinke et al., 1999) known to have distinct fat-body expression and one neuronal driver, n-synaptobrevin (N). These drivers were used to drive the expression of the previously used UAS-CrebAshRNAi to knock down CrebA protein. CrebA protein levels were drastically reduced in fat-body driven knock down flies when compared to no driver sample (Figure 5.18). Neuronal knock down did not show any differences to the control suggesting CrebA is mainly expressed in the fat-body cells of the head.

5.5. Discussion

In this chapter, I identify a nutritionally regulated transcriptional mechanism triggered to meet the secretory demands of an organism. Clearly, CrebA is a critical component of this pathway to likely avoid ER stress. Absence of ER protein sorting machinery genes including the Sec61 translocon and genes that form part of the ER-Golgi transport vesicles can activate the UPR pathway in the ER (Rios-Barrera et al., 2017). Our data identifies a majority of these critical components transcriptionally regulated by CrebA in response to nutrition. Taken together, CrebA forms part of a key regulatory node in managing ER stress and a cell's secretory demand in response to feeding.

I reported that CrebA mRNA and protein are quickly and transiently up-regulated upon feeding. However, what exactly triggers this transcriptional activation is unclear. Interestingly, western blot analysis showed CrebA protein running as a pack of multiple separate bands (Figure 5.1, 5.17A and 5.18) which suggests that it may be post-translationally modified as its mammalian orthologs (Khan and Margulies, 2019). It would be interesting to dissect if there is any post-translation control that allows CrebA protein to be transiently regulated upon feeding. Future studies focusing on CrebA pull down and/or immunohistochemistry experiments from fly head extracts can further enhance our understanding of this very interesting control of a transcription factor upon feeding.

Deeper analysis of CrebA targets that are dependent on nutrition in terms of their correlation to Pol2 elongation over gene bodies revealed that these set of targets are more likely to be transcribed by Pol2 (Section 5.4.4). With the evidence that these changes are specific to Pol2 over gene bodies and not over TSS, it suggests that CrebA more likely plays a role in releasing a poised Pol2 in response to a nutritional cue. This further suggests that the regulation of CrebA in response to nutrients is the key modulator of regulating these set of genes including the ER protein sorting machinery genes. It is likely that there are other factors involved that work together with CrebA in bringing this response. With genome-wide methodologies such as ATAQ-Seq, studying chromatin changes that occur upon nutrition can give key insights in to the regions and underlying DNA sequences that possibly are landing regions for other transcription factors.

CrebA's regulation of the ER protein sorting machinery genes in response to nutrients likely impacts the repertoire of secreted proteins. We know many secreted

proteins such as Fit and sNPF that can alter feeding behavior. Using the data collected, I identified changes in secreted levels of some of these feeding altering proteins. I also observe that exogenous increase in the expression of CrebA is sufficient to suppress feeding (Figure 5.16). How does CrebA suppress feeding behavior? Many interesting candidates can be studied that may act as the downstream actor for mediating feeding behavior. Fit is one such anorexic signal (Sun et al., 2017) that may be regulated by CrebA. Our data also revealed Fit being transiently regulated by feeding. Feeding behavior experiments using Fit mutants in a fly overexpressing CrebA may answer if Fit is a direct target of CrebA to mediate feeding behavior.

Beyond hormones, this work identified changes in hemolymph Apoltp levels upon feeding. However, changes in Apoltp hemolymph were only seen to be feeding dependent. Western blot analysis of hemolymph extracted from flies with knocked down CrebA or Sec61α, a major secretory pathway gene did not reveal reproducible changes to apoltp protein levels. Further experiments need to be done to fully understand if there are compensatory mechanisms at play in regulating secretion of lipoproteins in to the hemolymph. This would be important as growing evidence for apolipoproteins suggests their role in regulating feeding behavior. If CrebA regulates levels of circulating lipoproteins, it could reveal an important role for it as lipoproteins are key to maintaining organismal physiology and health. Although I show that CrebA can alter feeding behavior, precisely how CrebA links signaling mechanisms to feeding behavior remains to be an unanswered question. Future work using fly genetics and manipulation of expression and secretion of potential signaling molecules in CrebA gain or loss of function background can help pin down the precise signaling pathway at play here.

In summary, I specifically identify a clear and important function of CrebA in mediating cellular and organismal responses to fasting and (re-) feeding, by directly impacting the expression and function of ER components. This regulation provides feeding control possibly via regulating secretion of key hormones.

6. Mammalian Creb3L proteins regulate the ER protein sorting machinery genes in response to feeding

6.1. Summary

In previous chapters, I described the crucial role of CrebA in the transcriptional control of the ER protein sorting machinery genes and the physiological effects it has on feeding, secretion and lipid metabolism. Mammals contain five closely related CrebA orthologues. However, not much is known if these are similarly regulated in response to nutritional intake.

In this chapter, I report on the conserved CrebA orthologues in mammals known as the Cyclic-AMP Responsive Element Binding (Creb3L) transcription factor family. I describe that these mammalian proteins share a high sequence similarity in the DNA binding and activation domains with differences in how they are processed post-translationally compared to CrebA. Interestingly, I show that the most closely related Creb3L1 and Creb3L2 proteins are regulated by feeding in mice livers as confirmed with two independently done studies. As with CrebA, Creb3L proteins are regulated at the transcriptional level upon feeding. Furthermore, I show that the ER protein sorting machinery genes are also regulated upon feeding and are co-regulated. Using published ChIP-seq profiles for Creb3L1 and Creb3L2, I show that their binding is enriched in the ER protein sorting machinery genes. Additionally, Creb3L1 and Creb3L2 bind the same DNA motif as was described for CrebA. Lastly, I confirm the binding to the promoter of one of the ER protein sorting machinery genes, Sec24C in mice livers.

Overall, I show that feeding regulates expression of CrebA orthologues. Furthermore, I show that both Creb3L1 and Creb3L2 are highly enriched over promoters of ER protein sorting machinery genes whose expression levels are up-regulated upon refeeding in a similar fashion as Creb3L proteins.

6.2. Introduction

The fruit fly has contributed significantly to our understanding of developmental, cell and neurobiology. Although *Drosophila* and humans diverged millions of years ago, many of the metabolic organs in flies share the same function; the gut absorbs nutrients and the fatbody – fly equivalent of liver and adipose tissue – senses and stores nutrients (Colombani et al., 2003). *Drosophila's* shared disease genome with mammals together with a short lifespan, low maintenance costs and the availability of powerful genetic tools make it a perfect candidate for studying metabolism. Evidence of a crucial role for CrebA in maintaining lipid metabolism through the action on the ER protein sorting machinery genes provided a platform to study if this pathway is conserved in mammals.

Mice liver is an important metabolic organ involved in regulation of nutrients. Many obesity related pathologies such as dyslipidemia and non-alcoholic fatty liver disease (NAFLD) are associated to liver injury (Ahn and Sundaram, 2019). Liver functions through regulating many key metabolic processes. For example, liver houses dozens of carbohydrate metabolism enzymes that are turned on or off depending on signals initiated by changes in blood glucose levels. Furthermore, liver is extremely active in lipid metabolism metabolizing triacylglycerides to produce energy. Clearly, evolution of higher organisms to compartmentalize such metabolic processes in to organs such as the liver provide for healthy responses to environmental changes.

Upstream of these key metabolic processes, many hepatocyte-specific transcription factors play a key role to regulate the expression of enzymes involved in glycolysis and lipogenesis. One such transcription factor called the carbohydrate responsive element binding protein (ChREBP), potentially activated via direct allosteric binding to Glucose-6-phosphate (G6P), regulates transcriptional control of genes involved in glycolysis and *de novo* lipogenesis (Ortega-Prieto and Postic, 2019). Expression of HNF4a, another hepatocyte-specific transcription factor, through its interaction with ChREBP and independent functions improves the metabolic profile and rescues dyslipidemia (Huang et al., 2020, Mol Therapy NA). These transcription factors including many other perform the crucial processes to maintain healthy liver function.

Creb3L transcription factors lie at a unique regulatory node that links cell's secretory capacity with cellular and environmental stresses. Creb3L3/CrebH is

expressed highly in the liver and regulates metabolic genes involved in gluconeogenesis (Chin et al., 2005; Kim et al., 2017; Lee et al., 2010) and genes involved in lipid storage and transport (Xu et al., 2015). Creb3L3 also regulates the transcription of genes encoding for lipoprotein particles such as ApoC2 and ApoA4 (Dandekar et al., 2016; Xu et al., 2014; Zhang et al., 2012). Interestingly, as contrary to CrebA's upregulation upon feeding, Creb3L3 is upregulated only with fasting and high-fat diet (Danno et al., 2010; Lee et al., 2010; Vecchi et al., 2014; Zhang et al., 2012). Surprisingly, not much is known how other Creb3L family members respond to nutritional cues. All Creb3L family members are expressed in multiple secretory tissues. The presence of multiple Creb3L proteins in mammals hints at a possible cross-talk between these proteins.

In order to better understand the regulation of mammalian Creb3L proteins in response to feeding, I sought to understand if 1) other members of Creb3L family are regulated by feeding and 2) upon this regulation they regulate the secretory capacity of cells they are expressed in. A better understanding of this pathway would reveal if there is a previously unknown additional layer of complexity that allows these proteins to coordinate responses to feeding.

6.3. Methods

6.3.1. RNA and ChIP-seq analysis

RNA was collected and cDNA prepared as described previously in section 4.3.2, 4.3.3 and 4.3.4.

Count matrix was downloaded from the submitted repository GSE118973. Differential expression analysis was performed using maSigPro. Gene expression levels for transcripts were plotted using the count tables after library normalization. Creb3L1 ChIP-seq was downloaded as BAM files from encodeproject.org with the accession ENCFF950GGY. Peaks from the BAM files were called using macs2 (version 2.1.1) with a fold change cut-off of 7 and default parameters. Background signal from each sample was used by macs2 as control to find peaks with fold greater than 7. Nearest genes to identified regions were annotated using RGmatch. Peaks within 1500bp of promoter regions of a gene were included in the downstream analysis. GO terms were used were from the biomaRt database (version 2.38) in R. FDR adjusted p-values were obtained for GO terms using a Fischer's Exact Test.

6.3.2. Mice liver analysis

Mice experiments were performed in collaboration with Paula Ortega-Prieto in the lab of Catherine Postic.

C57BL/6 male mice were fasted for 24 hours and refed with standard chow diet plus 20% glucose solution for 0, 0.5, 2 or 4 hours and sacrificed always at noon to avoid circadian effects. A total of 6 mice were used per timepoint / condition and sacrificed together on the same day. Weight and glucose measurements were performed before sacrifice using a scale and a blood glucose meter. Livers collected were snap-frozen immediately after dissection and stored at -80ºC until used. 15mg of mice liver was weighed and RNA was collected as mentioned above. RT-qPCR was done using primers amplifying small regions of Creb3L1 and Creb3L2. Data was normalized to the house-keeping gene TATA-binding protein (TBP). For westerns, whole-cell protein lysates were prepared from 20mg of liver using a lysis buffer (BIORAD) containing 1X TRITON-X. Protein concentration was measured and samples were analyzed with western blots as described in section 4.3.10. Commercially available mouse monoclonal anti-Creb3L2 (Merck, MABE1018) and anti-GAPDH (Genetex, GT239) antibodies were used.

6.3.3. Cell culture experiments

cDNA prepared from mice liver was used to clone out Creb3L2 coding sequence using PCR and introduced in a pEGFP-C1 (Clonetech) vector. HEK293T cells were transfected with 10µg of vector and protein lysates were prepared 2 days after. GFP tagged Creb3L2 expression was confirmed under a microscope and western blot analysis using antibodies against GFP (developed in-house).

6.4. Results

6.4.1. Mammalian CrebA orthologues are highly conserved in DNA binding and activation domains

Drosophila and mammals share many metabolic and nutrient-regulated hormonal signaling pathways. To understand if feeding dependent CrebA control of ER and lipid metabolism is conserved in mammals, I performed a deep sequence comparison analysis between CrebA and its Creb3L mammalian orthologs. Mammals possess five CrebA orthologs (Khan and Margulies, 2019) (Figure 3.8). Previous reports based on blast and sequence analysis reported Creb3L transcription factors are highly related to *Drosophila* CrebA (Fox et al., 2010; Abel et al., 1992; Smolik et al., 1992). CrebA however, does not contain the transmembrane domain (green) meaning it does not get cleaved at the ER. CrebA, Creb3L1, Creb3L2, and Creb3 contain an interaction site for HCF (shown in black circle) which, in turn, recruits chromatin factors such as DNA modifying enzymes that may suggest the type of chromatin changes that these transcription factors could bring about. Creb3L3 is modified by a number of posttranslational modifications, including phosphorylation (yellow circle), acetylation (red circle), and N-linked glycosylation (gray circle) that has been shown to regulate its activity. Phosphorylation in the basic domain (yellow circle shown under the sequences) is conserved in all proteins which may suggest that this post-translational control also occurs for other members and *Drosophila* CrebA.

Creb3L1 and Creb3L2 are however, the closest mammalian orthologs of CrebA based on sequence similarity in these important DNA binding and activation domains. They share approximately 25% sequence similarity overall with 97% similarity (84% identity) and 79% similarity (71% identity) within the DNA binding domains (Fox et al., 2010). Sequence similarity is further complemented by evidence that mammalian orthologs are able to activate the expression of fly CrebA target genes (Fox et al., 2010; Barbosa et al., 2013). Further, I have previously explored the many secretory roles of Creb3 proteins in mammals (Khan and Margulies, 2019). Since I was interested in the role of Creb3 proteins in response to nutrition and found CrebA to be nutritionally regulated in the fly fat body, I chose to study Creb3 proteins in mice liver, a metabolically highly active organ.

6.4.2. Mammalian Creb3L proteins are regulated by feeding

To determine if Creb3L1 and Creb3L2 are regulated by feeding, I used an already published RNA-seq dataset that quantified mRNA levels in mice livers upon early refeeding of 0, 0.5, 1, 2 and 4 hours (Brandt et al., 2018). Interestingly, both Creb3L1 and Creb3L2 were quickly up regulated upon feeding (Figure 6.1). While Creb3L1 expression levels slowly increased with feeding over the 4 hours, Creb3L2 transcript was recorded to peak at 2 hours after refeeding. Creb3L3 was down regulated upon feeding as previously described.

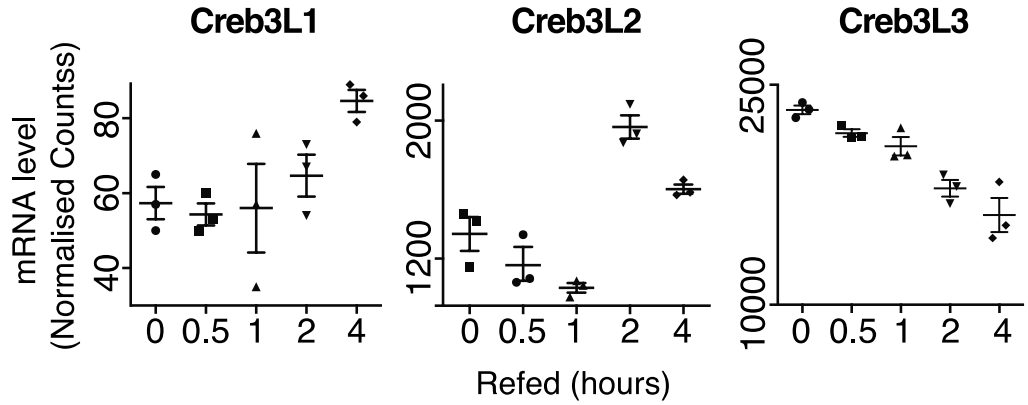

Figure 6.1 Feeding regulates the expression of Creb3L transcription factors.

Creb3L1, Creb3L2 and Creb3L3 mRNA transcript levels in mice liver upon early refeeding. Data shown are normalized counts from previously published dataset and plotted as means ±SEM for n=3 individual mice biological replicates (Brandt et al., 2018).

To further confirm these data in an independent study, I performed a similar refeeding timecourse experiment in mice. Normally fed mice were fasted for a total of 24 hours after which they were refed for 0, 0.5, 2 and 4 hours and samples collected (Figure 6.2A). Fasting period was adjusted so that all mice were collected at the same time to avoid circadian regulation dependent changes. Before sacrifice, mice were weighed and pricked for a blood glucose measurement. A total of 6 mice were used for each timepoint / condition that were all sacrificed all together on one day. Interestingly, blood glucose levels were elevated within 0.5 hours of refeeding suggesting that mice were able to feed properly within these short time periods (Figure 6.2B). Weight measurements also showed a steady increase over the 4 hours (Figure 6.2B).

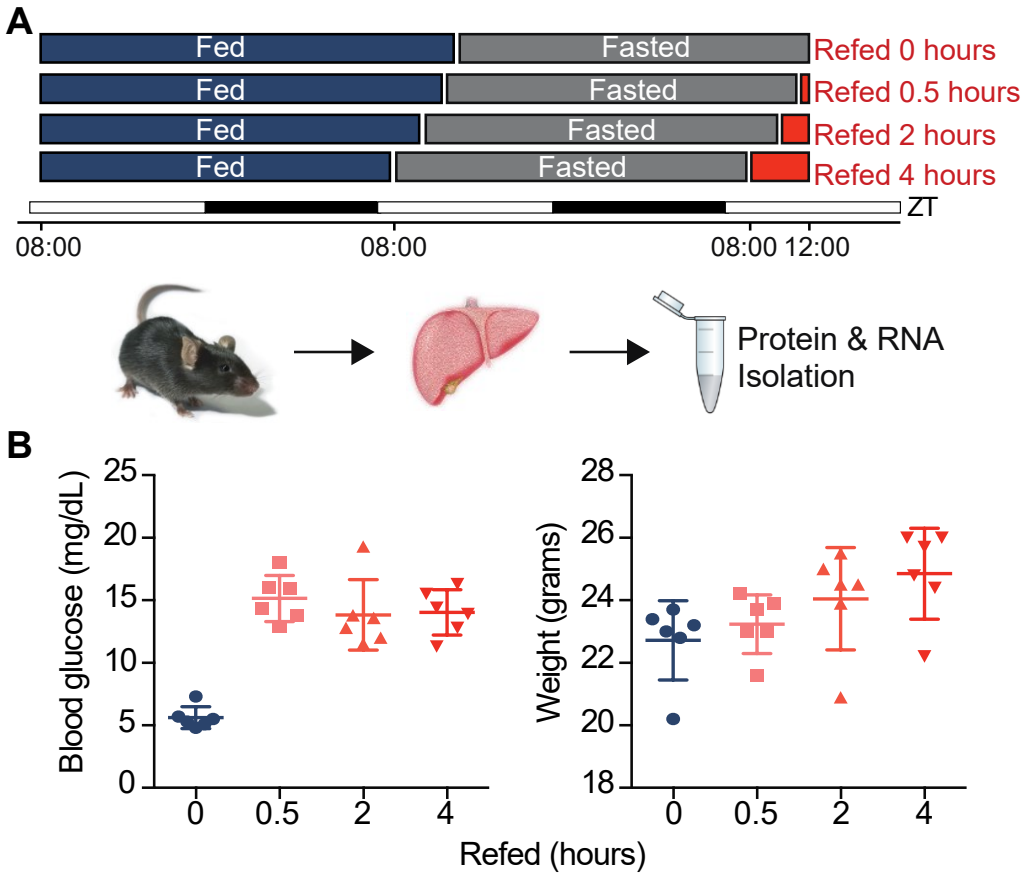

Figure 6.2 Scheme illustrating the feeding paradigm administered to mice.

(A) Normally fed mice were fasted for 24 hours. Liver samples were collected at refed 0, 0.5, 2 and 4 hours. Fasting was started accordingly so that samples were collected at the same time to account for circadian differences. ZT = zeitgeber time. **(B)** Blood glucose (left) measured from blood and weight (right) measurements confirmed that mice were able to feed quickly after a period of starvation (n=6 with each n representing data from a single mice sacrificed together with all other mice).

Next, I collected mRNA from these liver samples to assess the levels of Creb3L transcripts. I confirmed the increased levels of Creb3L1 and Creb3L2 upon refeeding (Figure 6.3). Interestingly, RT-qPCR data corroborated exactly with the mRNA seq levels. Creb3L1 transcript slowly went up with refeeding while Creb3L2 peaked at 2 hours.

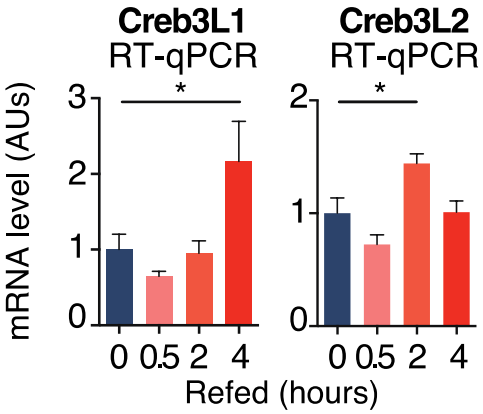

Figure 6.3 Creb3L1 and Creb3L2 transcripts up regulated upon refeeding confirming RNA-seq findings.

mRNA levels of Creb3L1 and Creb3L2 in mice liver upon refeeding measured by RT-qPCR and normalized to a housekeeping TBP gene. Data plotted are absolute units measured from standard curve (n=6 biological replicates) and analyzed with one-way ANOVA with Tukey correction (*p<0.05).

In order to check if the mRNA changes translate to protein level, I first set up western blots from mice liver. As mammalian Creb3L transcription factors are processed post-translationally (Figure 3.8), I expected two separate bands for each of the precursor and cleaved forms of the protein. To this end, I cloned Creb3L2 cDNA in to a pEGFP-C1 plasmid to express a GFP tagged version of Creb3L2 (Figure 6.4A). The plasmid was transfected in to human HEK293T cells, confirmed for GFP expression and cell extracts were prepared. Upon a western blot analysis, both precursor and cleaved forms were detected at their corresponding sizes when probed with antibodies against Creb3L2 and GFP (Figure 6.4B). This suggested that the Creb3L2 antibody can specifically detect both forms of Creb3L2.

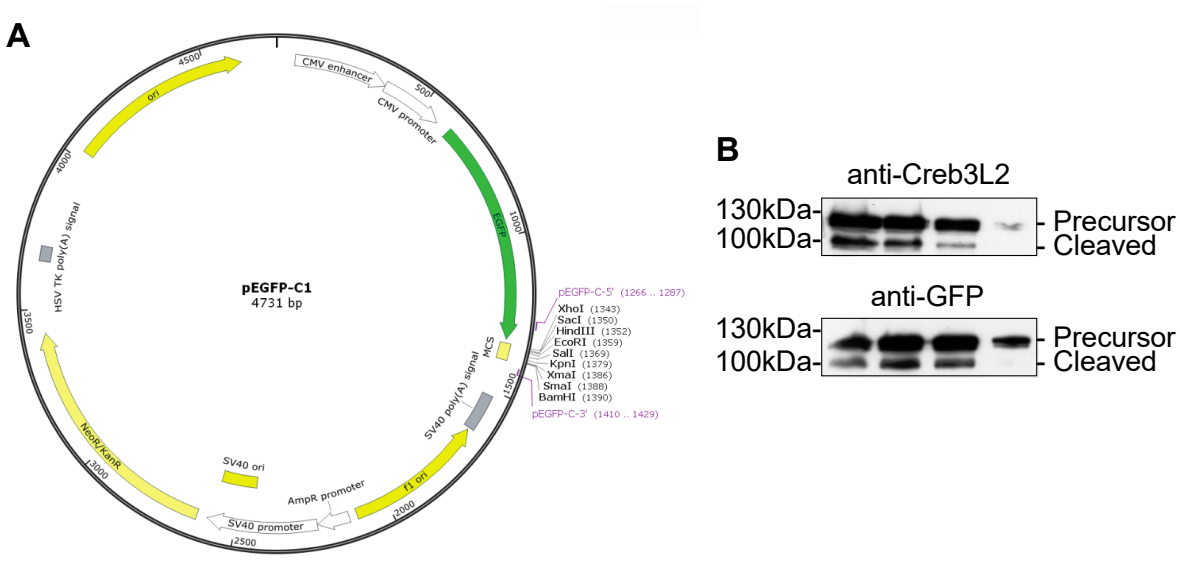

Figure 6.4 GFP tagged Creb3L2 is cleaved and processed.

(A) Plasmid map showing pEGFP-C1 vector which was used as a backbone to insert the Creb3L2 cDNA. **(B)** Western blots showing precursor and cleaved forms of Creb3L2 protein expressing in HEK293T cells.

Both CrebA transcript and protein were shown to be nutritionally regulated. Next, I asked if Creb3L1 or Creb3L2 levels are regulated at the protein level. I prepared extract from the same conditions / timepoints from mice liver. Western blot experiments against Creb3L2 protein indicated that its protein levels are also transiently regulated upon refeeding (Figure 6.5A), recapitulating CrebA regulation by nutrients in the fly. Mammalian Creb3 proteins are also regulated post-transcriptionally. Interestingly, both cleaved and uncleaved forms of Creb3L2 were seen to be nutritionally regulated (Figure 6.5A and 6.5B). Together, this suggests that Creb3L2 levels are upregulated at the transcript level resulting in increased protein levels.

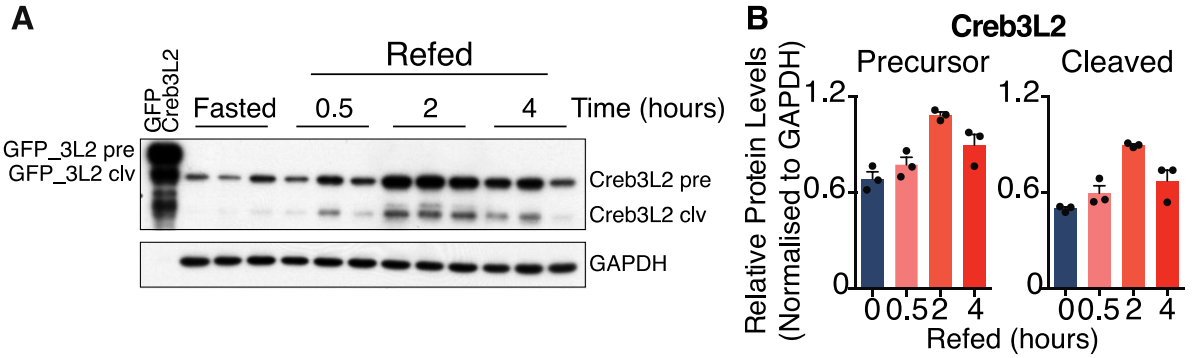

Figure 6.5 Both Creb3L2 precursor and cleaved forms are up regulated with refeeding.

(A) Western blot probing for levels of Creb3L2 along the refeeding time-course in mice liver (3 biological replicates for each time point). GAPDH was used as a loading control. **(B)** Barplot showing normalized levels of Creb3L2 precursor and cleaved forms. Data plotted are biological replicates n=3 ± SEM. GAPDH was used for normalization.

6.4.3. Mammalian Creb3L transcription factors regulate the ER protein sorting machinery genes

A number of regulatory pathways are conserved between mammals and flies (Lang et al., 2017). Next, I asked whether the ER protein sorting machinery genes are also regulated by nutrition similar to what I observed in flies. I used the previously published RNA-seq data set from mice liver (Brandt et al., 2018) to ask what happens to the expression of these genes upon early feeding. Differential gene expression analysis and unsupervised clustering of this dataset revealed three major clusters (Figure

6.6A). The expression levels of the first cluster with 275 genes were already upregulated at 0.5 hours after refeeding. A second cluster with 1053 genes was downregulated with refeeding and contained genes involving carbohydrate and lipid breakdown metabolism required in fasting. Interestingly, the third cluster with 1691 genes was upregulated with refeeding up until 4 hours and was highly enriched with the ER protein sorting machinery genes (Figure 6.6B). This expression profile matched with the upregulation of these genes in *Drosophila*. Based on these data, I postulated that similar to *Drosophila*, the Creb3L family of proteins likely regulates the ER protein sorting machinery genes in response to nutrients.

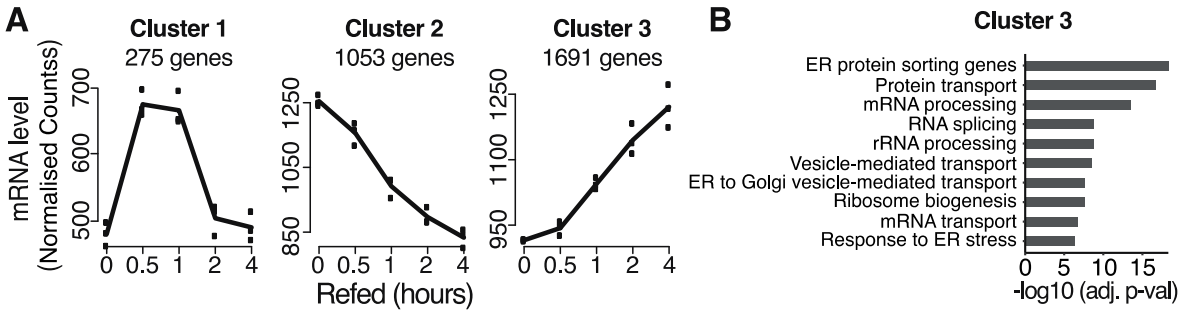

Figure 6.6 Feeding regulates the mammalian ER protein sorting machinery genes.

(A) Clustering analysis of the differentially expressed transcripts along the refeeding timecourse performed using maSigPro from the RNA-seq dataset (Brandt et al., 2018) reveals 3 major clusters. **(B)** Gene ontology analysis of 1691 genes in cluster 3 using a Fischer's Exact test.

To further explore this possibility, I used previously published Creb3L ChIP-seq datasets to ask if Creb3L transcription factors bind the promoters of secretory machinery genes. Creb3L1 and Creb3L2 ChIP-seq dataset from human K562 and mouse pituitary cell lines (Consortium, 2012; Davis et al., 2018; Khetchoumian et al., 2019) identified 1137 and 1818 targets of Creb3L1 and Creb3L2 respectively that were enriched with targets encoding for the early ER protein sorting machinery genes (Figure 6.7B). Sec24C, a core component of the CopII vesicle, was a target of both Creb3L1 and Creb3L2, establishing Sec24C as a target of mammalian Creb3L (Figure 6.7A). Furthermore, I performed a gene ontology analysis on targets of Creb3L1 and Creb3L2 proteins (Figure 6.7B). Interestingly, both Creb3L family members had gene ontology terms containing the ER protein sorting machinery genes as the most highly enriched GO terms. These included the genes required for COPI and COPII vesicle formation and transport.

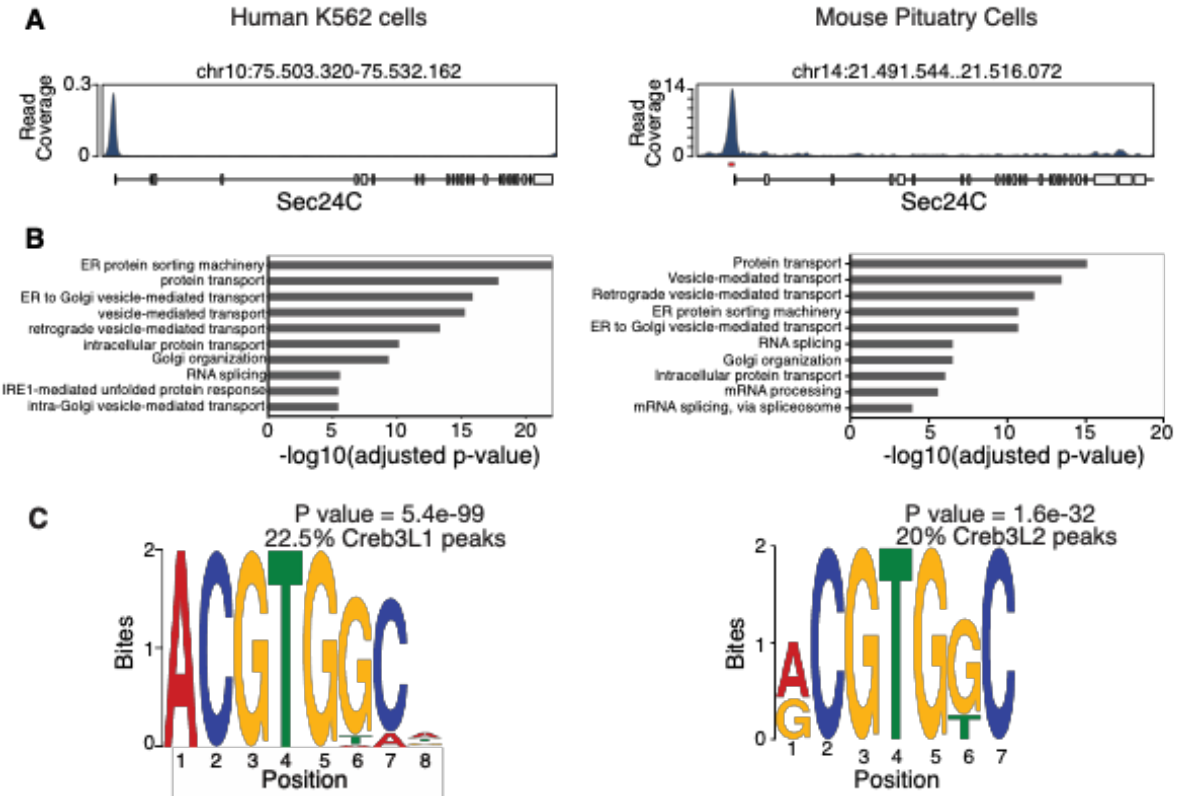

Figure 6.7 Mammalian Creb3L2 binds the ER protein sorting machinery genes.

(A) Coverage tracks showing Creb3L1 (left) and Creb3L2 (right) peaks over the Sec24C promoter (Consortium, 2012; Davis et al., 2018; Khetchoumian et al., 2019). **(B)** Gene ontology analysis of all Creb3L1 and Creb3L2 target genes using Fischer's Exact test from the ChIP-seq data. **(C)** Motif identified in Creb3L1 or Creb3L2 bound regions using MEMEChIP. At least 20% of Creb3L1 and Creb3L2 peak regions are enriched with the conserved motif.

In addition, I found that Creb3L1 and Creb3L2 bind the same DNA motif (Figure 6.8C) containing very similar basic regions. Interestingly, these motifs are very similar to the one found for CrebA (Figure 5.6) from our own experiments suggesting conservation between the mammalian and fly DNA motifs.

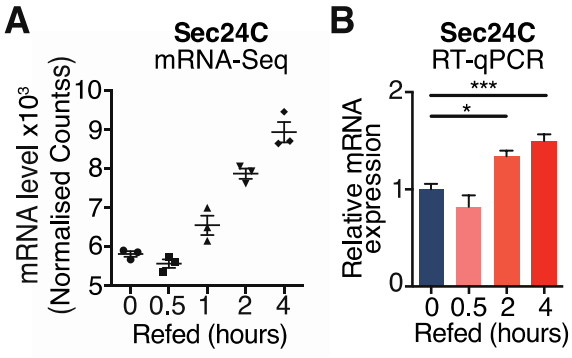

Figure 6.8 Sec24C gene expression is up regulated upon refeeding.

(A) Sec24C mRNA levels are upregulated upon refeeding (Brandt et al., 2018). Data plotted are n=3 from the RNA-seq dataset and (B) RT-qPCR confirms upregulation upon refeeding (n=6). Data plotted min to max and analyzed using one-way ANOVA with Tukey's test compared with all groups (***p<0.001 and *p<0.05).

To dissect if binding of Creb3L transcription factors on ER protein sorting machinery genes resulted in an increase in transcript levels, I checked for mRNA levels of Sec24C. Transcript levels of Sec24C quantified from the RNA-seq data clearly suggested that Sec24C is upregulated in response to feeding (Figure 6.8A). These results were then confirmed with a RT-qPCR experiment using the mice livers we had collected. Again, the RT-qPCR measured transcript levels corroborated highly with RNA-seq levels initially going down at 0.5 hours and then upregulated with subsequent feeding (Figure 6.8B).

Published Creb3L ChIP-seq datasets had confirmed Sec24C as a definite target of Creb3L transcription factors. However, these experiments were done in Human K562 (Creb3L1) or mouse pituitary cells (Creb3L2). To confirm these findings in mice livers, and better dissect the feeding dependent of the ER protein sorting machinery, I sought to ChIP Creb3L2 in mice livers. I observed that Creb3L2 specifically bound to the promoter of Sec24C and did not show any binding to a region 3kb upstream of this promoter (Figure 6.9).

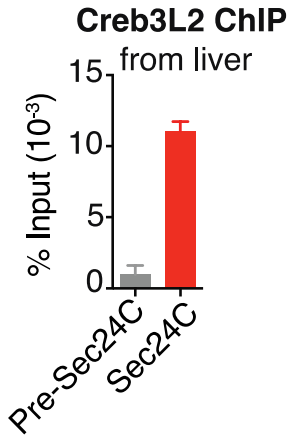

Figure 6.9 Creb3L2 binds Sec24C promoter in mouse liver.

Creb3L2 ChIP-qPCR from mice liver chromatin over the Sec24C peak marked in red Figure 6.8A. A region 3kb upstream was amplified and used as control.

Taken together, these results demonstrate that the mammalian Creb3L members depict similar functions in response to nutrients and likely function in a similar way as identified in *Drosophila* to regulated the secretory capacity of cells.

6.5. Discussion

In this chapter, I extended the findings in *Drosophila* to mammals and showed that mammalian CrebA orthologues are regulated by feeding and in turn may coordinate the expression of the same ER protein sorting pathway as in flies. Previously, only

Creb3L3 was reported to be expressed in mice liver. Our findings uncover a potentially new role for Creb3L1 and Creb3L2 in mice liver where they may regulate the ER protein sorting machinery upon changes in nutrition.

I report that both Creb3L1 and Creb3L2 are regulated transcriptionally within 2 hours of refeeding. This was evident by mRNA analysis and no change seen in the ratio between the precursor and cleaved forms of Creb3L2 protein. Interestingly, post-translational modifications reported for Creb3L3 suggest that there is an additional layer of control. Phosphorylation of Creb3L3 was shown to regulate both ER anchored proteolysis (Zheng et al., 2016) and protein turnover (Barbosa et al., 2015; Barbosa et al., 2017; Cheng et al., 2016). If these modifications play a role in regulating post-translation control of other proteins still remains to be elucidated. An interesting modification residue for such control would be the conserved phosphorylation residues in bZip domain of Creb3L3. Two conserved serine residues in all Creb3L proteins and CrebA are reported to be phosphorylated by glycogen synthase kinase 3 beta (GSK3B). Decrease in GSK3B activity results in a decrease in glycogen synthesis in mice livers and muscles and increase in blood sugar levels or hyperglycemia (Ali et al., 2001). Interestingly, cellular activity of GSK3b is regulated by insulin signaling (Souder and Anderson, 2019). It would be interesting to elucidate if such kinase mediated control of Creb3L proteins provides for an additional layer of control in protein turnover and activity.

Using a previously published RNA-seq dataset, I show that the ER protein sorting machinery genes are co-regulated upon refeeding. Upon an unsupervised clustering analysis, I find the ER protein sorting machinery genes enriched in a cluster that is slowly up regulated and peaks at 4 hours. However, based on the time resolution of the experiment it is unclear if the upregulation of these genes goes beyond the 4-hour time point. Regardless, I show that the Creb3L transcription factors bind the promoters of these set of genes and are their likely transcriptional regulators. Enrichment of the Creb3L binding site in published ChIP-seq datasets further corroborates this hypothesis as most ER protein sorting machinery have the canonical Creb3L binding motif. ChIP-qPCR experiments in mice livers also support these findings as Sec24C, a core component of the CopII vesicle was strongly bound by Creb3L2. However, more detailed time-resolution ChIP-seq experiments are needed

to be done to understand if refeeding also coordinates the binding of these transcription factors to its target genes.

7. Discussion and Perspectives

Altering nutrition states inflict a wide array of changes to organismal physiology ranging from absorption of nutrients, their delivery and storage. It is commonly understood that higher organisms achieve metabolic homeostasis by three broad mechanisms. The first entails the allosteric control of key enzymes involved in cellular metabolic processes that are activated upon ligand binding. The second mechanism involves the posttranslational control of proteins via proteolytic cleavage, glycosylation, acetylation and phosphorylation to quickly and transiently affect enzyme activity and stability. And third by transcriptional control to bring about more stronger and long-term changes by regulating expression of genes required for metabolic readjustment. Our understanding of these three broad mechanisms suggests that they benefit by functioning in coordination with each other to maintain metabolic homeostasis.

Fruit flies provide us an opportunity to study complex feeding regulated mechanisms with an array of available genetics and proteomics tools that can be rather easily “programmed” to suit the experimental needs. Specifically, gene expression control requires complementary pieces of information i.e., 1) events upstream of transcriptional activity such as an activating signal, 2) the molecular action of transcription factors and 3) the events downstream that turn on the transcription of groups of genes targeted. Using a nutritional paradigm, we systematically profiled gene activity in response to nutrients and uncovered a role for the evolutionary conserved transcription factor, CrebA and its mammalian orthologs. I show that both fly and mammalian proteins regulate cellular secretory capacity in response to nutrients. In response to acute refeeding, CrebA is upregulated to bind and regulate the transcription of the ER protein machinery genes via the release of Pol II from a paused state.

In this thesis, I presented the findings on a novel pathway that functions in response to nutritional intake to reorganize the metabolic state of the organism. Using mice liver, a highly active metabolic organ, I extended the analysis to show that the mammalian orthologs of CrebA are also regulated in a similar manner and bind similar target genes.

7.1. Nutritional regulation of CrebA/Creb3L transcription factors

CrebA/Creb3L transcription factors are quickly and transiently up regulated with refeeding normal food. Experiments in both flies and mice suggest that this upregulation is quick and transient with protein levels peaking at 2 hours after refeeding. These levels then slowly return back to normally fed levels after 24 hours. How these changes are regulated upstream is still not clearly understood.

Experiments presented in this thesis were done with animals fed standard diet, a mixture of carbohydrates, proteins and fats. Thus, it cannot be determined if a specific nutrient regulates CrebA/Creb3L levels. Preliminary work in the lab has shown that both sugars and protein sources can up regulate CrebA levels in flies. Future work can explore if there is a nutrient specificity to this pathway or a general response to food intake that in turn coordinates the ER to be able metabolize incoming nutrients.

The upregulation of CrebA and Creb3L2 transcript levels upon refeeding suggests that there is transcriptional control of these genes. So far, there are no known transcriptional regulators of these transcription factors. Analysis of their promoter region did not show any conserved binding sites for transcription factors. Creb3L3, however, has been shown to be up regulated by free fatty acids (FA) (Danno et al., 2010). Notably, Danno et al. showed that the Creb3L3 gene contains a peroxisome proliferator responsive element (PPRE) and that the administration of PPARα agonists increased Creb3L3 transcript levels. However, another study (Gentile et al., 2010) showed that the FA mediated upregulation of Creb3L3 transcript levels does not depend on PPARα activity but is instead induced by PI3K signaling. Further studies are needed to understand this complex transcriptional control of CrebA/Creb3L transcription factors.

Post-translation control of CrebA/Creb3L transcription factors is an area which is better understood. All mammalian forms are cleaved at serine protease sites upon activation and anchoring to the ER, after which the N-terminal part is translocated in to the nucleus. Additionally, Creb3L proteins have been shown to post-translationally phosphorylated, acetylated and glycosylated (Chan et al., 2010; Kim et al., 2010; Zheng et al., 2016). A specifically interesting residue that is phosphorylated is the S260 located in the basic domain which is conserved across species (Figure 6.1). Currently, it has only been shown that Creb3L3 is post-translationally regulated site to induce intramembrane proteolysis and translocation in to the nucleus. If other Creb3L

proteins including CrebA are phosphorylated are regulated via this event still remains to be elucidated. Westerns performed in flies show a series of different sized bands suggesting that there is some post-translation control. Immunoprecipitation experiments can further explore this possibility and elucidate if these mechanisms are conserved between family members and across species. These processes are likely events that are important in fine tuning of CrebA/Creb3L transcription factor activity.

7.2. Cell-type specifity of CrebA/Creb3L regulation

The data from *Drosophila* presented in this thesis were done using the head tissues. *Drosophila* head contains multiple cell types. The main part is occupied by the different compounds of eye that are comprised of photoreceptor cells (Pichaud et al., 2001). The other major part of the head cavity is occupied by the *Drosophila* brain. Most abundant are cell types present in the brain are neuronal cells followed by glial cells. Additionally, the brain is surrounded by fat body cells. As discussed before, fat body has a role in regulating lipid metabolism and secretion of key hormones and pheromones in to the hemolymph (Arrese and Soulages, 2010).

Western blots done on fly heads determined that CrebA protein levels depended on the feeding state of flies. In order to understand which cells types express CrebA, I performed a CrebA RNAi experiment knocking down CrebA in either the fat body or neurons. Interestingly, only a fat body knockdown and not a neuronal CrebA knockdown resulted in the loss of CrebA protein when whole heads were loaded suggesting most of the CrebA protein expression came from the fat body (Figure 5.18). Previously, CrebA has been reported to be highly expressed in high secretory demand tissues such as the salivary glands. Considering the fat body also needs to secrete hormones for metabolic control, CrebA's high expression in the fat body suggests that it could serve as a key regulator of secretory control of metabolic hormones upon food intake. Further, expression profiling of CrebA mRNA in different tissues also suggests that it is abundantly expressed in the larval and adult fat bodies (Flybase; Figure 7.1). Other tissues which have high CrebA mRNA expression are the male and female reproductive tissues, also cells that have a high secretion load.

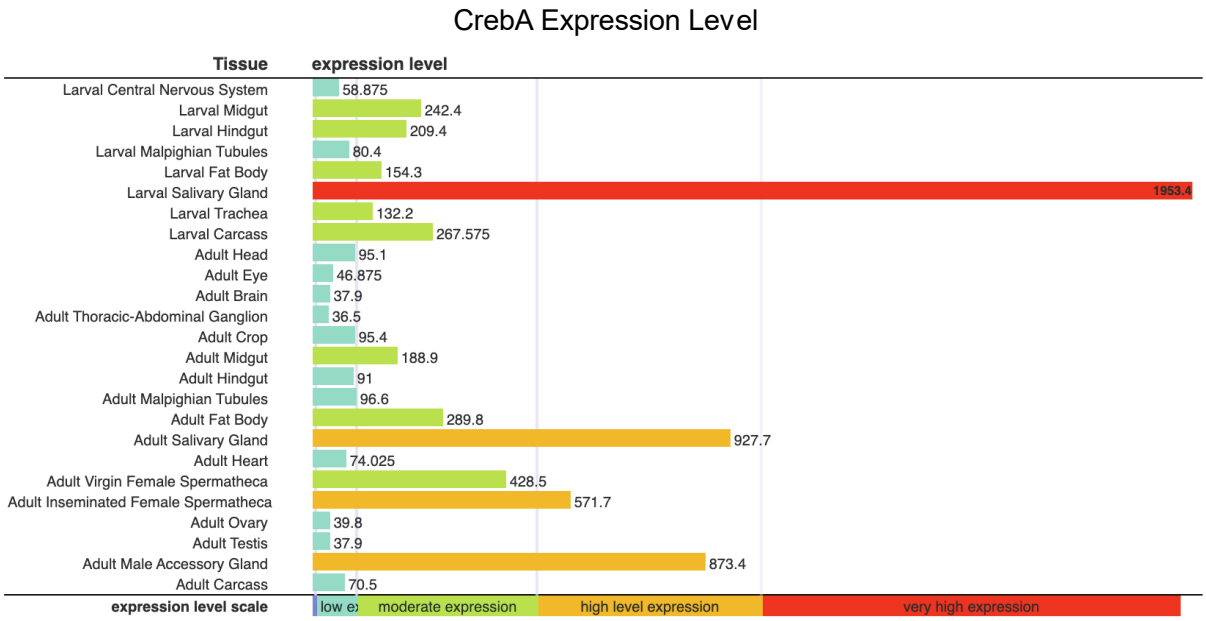

Figure 7.1 Expression level of CrebA mRNA in different tissues.

CrebA is highly expressed in secretory tissues such as the larval and adult salivary glands, adult fat body and reproductive organs. Data shown are scaled z-scores. Source: Flybase.org.

Mammals contain five CrebA ortholog genes suggesting that there is a complex interplay of these proteins. Creb3L3, being the most studied, is highly expressed in the liver where its expression regulates key pathways in glucose and lipid metabolism (Nakagawa et al., 2016). Creb3L3 is also abundantly expressed in the small intestine which has been shown to regulate cholesterol metabolism (Kikuchi et al., 2016). Other Creb3L proteins are only expressed in low level. Clearly, Creb3L3 expression in these metabolic tissues has a significant role in maintaining metabolic homeostasis.

Other Creb3L proteins are expressed in different secretory tissues. Creb3L1 and Creb3L2 are expressed in bone and cartilage forming cells and their absence leads to defects in these processes (Murakami et al., 2009; Saito et al., 2009). Additionally, Creb3L1 is also highly expressed in metabolic tissues such as the stomach and pancreas (Figure 7.2). However, its function in these tissues remains to be elucidated. It would be interesting to dissect if Creb3L1 function corroborates our findings by functioning as feeding dependent regulator of secretory control in these tissues.

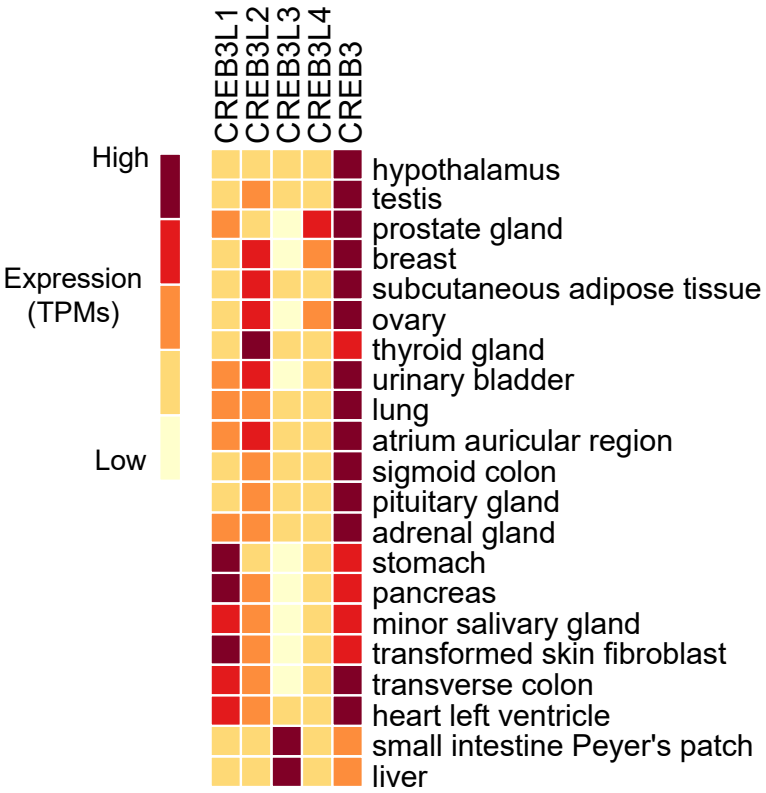

Figure 7.2 Expression level of Creb3L transcription factors mRNA in different tissues.

Creb3L transcription factors are highly expressed in secretory and metabolic tissues.

7.3. CrebA may regulate RNA Pol II elongation over its target genes

Transcription of genes requires the formation and release of the RNA Polymerase II on the protein coding DNA sequence. Transcription factors play a key role in mediating multiple steps that are required during this process. As discussed earlier (see section 3.4), Pol II can be poised in a paused state requiring a signal to elongate and transcribe DNA in to mRNA. While the CrebA ChIP-seq data identified promoter regions CrebA binds, Pol II ChIP-seq data was able to serve two purposes. First, it showed that the difference observed in CrebA binding in different feeding conditions was not due to differences in chromatin preparations. Second, the data was able to pull out differences in Pol II elongation over nutritionally responsive genes. This was evident from very few differences observed when TSS localized paused Pol II peaks were compared between fasted and fed states (Figure 5.13). Pol II signal over gene bodies were significantly more different corroborating the mRNA levels from our transcriptomics data. However, not all nutrient dependent transcripts showed differences in Pol II occupancy over gene bodies. This could be attributed to the lack of sensitivity associated with ChIP-seq technology. Use of other technologies more appropriate such as NETseq may be more effective in determining “live” transcription

occurring from RNA Pol II. It is interesting to note that 149 CrebA targets that are nutritionally regulated showed a marked difference in Pol II occupancy over gene bodies and not over TSS between different feeding conditions. Phosphorylation of the Ser 5 CTD residue is a key characteristic of elongating Pol2 which may aid in breaking Pol2 contacts with promoter-bound factors (Liu et al., 2005, Mol Cell Biol). Previous work has suggested that transcriptional activators can promote this phosphorylation event by signaling components of the pre-initiation complex (Spilianakis et al., 2003). Pull-down experiments with antibodies against CrebA/Creb3L transcription factors can help further dissect if they function by interacting with components that regulate Ser 5 phosphorylation of the CTD. Based on these findings, I hypothesize that CrebA may act as a Pol II release switch for its target to initiate transcription. Future work may help us understand how CrebA interacts with the paused polymerase transcriptional machinery to cause this release and broaden our understanding of transcription factor mediated control of Pol II in response to external stimuli. Furthermore, Pol2 ChIP-seq experiments done in mammals will also give insight if Creb3L mediated transcription release is regulated in a similar manner.

7.4. Cross-talk between CrebA/Creb3L and other factors?

Transcription factors are composed of transactivation domains that are responsible for protein-protein interactions with other transcriptional co-factors. Similarly, Creb3L transcription factors interact with themselves and other co-factors to regulate transcription of their downstream targets. Interestingly, CrebA, Creb3L1, Creb3L2 and Creb3 all contain a linear motif that allows them to interact with the host cell factor (HCF). HCF, in turn, recruits chromatin factors such as the methyltransferase enzymes (Set1 and MLL; (Wysocka and Herr, 2003)) and a demethylases (Liang et al., 2009), acetyltransferases (Dou et al., 2005; Guelman et al., 2006; Smith et al., 2005) and deacetylase (Wysocka and Herr, 2003). Interaction of Creb3L transcription factors with HCF is also particularly interesting because HCF is also metabolically regulated by the control of O-GlcNAc transferase in response to nutritional and hormonal cues (Capotosti et al., 2011). In the future, it would be interesting to study these interactions *in vivo* in response to nutritional cues and how it may impact gene expression changes driven by these transcription factors.

In this thesis, I dissect the nutrient-driven role of *Drosophila* CrebA and mammalian Creb3L transcription factors play in regulating the ER protein sorting machinery. What role does this pathway play in regulating organismal physiology? One potential answer from literature is to avoid ER stress. The unfolded protein response (UPR) plays a key role in avoiding ER stress by managing ER folding capacity. Some studies suggest a link between ER stress, UPR and the Creb3L family via the transcription factor Xbp1. Xbp1 has been reported to bind the promoter regions of both Creb3L1 and Creb3L2 in mice liver (Liu et al., 2019). Interestingly, similar to our findings for Creb3L transcription factors, Xbp1 binding was also nutritionally regulated. Furthermore, the same study also reported nutrition-dependent binding of Xbp1 on genes involved in COPII vesicle formation, an integral part of the ER protein sorting pathway (Liu et al., 2019). Likely, ER stress triggered by excess nutritional influx, initiates the UPR pathway by activating Xbp1 resulting in a coordinated control of the Creb3L and ER protein sorting machinery genes. As a result, ER stress is avoided. Clearly, the ability of Creb3L transcription factors to regulate such important ER pathways is therefore crucial especially in context of nutritional response. Further studies are needed in different cell-type / organisms to understand if these proteins play a similar role under different circumstances.

More recent studies have uncovered further evidence of interaction between CrebA/Creb3L and Xbp1 pathways. Genome-wide binding data of CrebA in flies suggests its binding at the Xbp1 promoter (Johnson et al., 2020). Interestingly, I also see CrebA binding in our ChIP-seq dataset. Apart from genome-wide studies, Xbp1 was also shown to be transcriptionally regulated by CrebA via transcription reporter assays (Johnson et al., 2020). Taken together, CrebA's regulation of Xbp1 in flies clearly suggests that CrebA combines with other key regulators to meet the secretory demands of ER, the cell and the organism. Interestingly, where CrebA drives transcription of Xbp1 in response to ER stress, CrebA may repress Xbp1 expression in response to bacterial infection (Troha et al., 2018) indicating that there might be a more complex interaction at play depending on cell-type/organism. A better understanding of this interaction may help us gain further insights on how a key regulatory pathway mediates organismal physiology in response to environmental stress.

One of the members of the Creb3L transcription factor family, Creb3, interacts with another protein called the Creb3 regulatory factor (CrebRF). CrebRF is induced by ER stress, interacts with Creb3 and promotes its proteosomal degradation (Audas et al., 2008). Interestingly, CrebRF has been implicated in many metabolic roles. A missense variation in the CrebRF region was associated with higher body mass index and an increased risk of obesity in a human population (Loos, 2016; Minster et al., 2016). Confirmatory experiments conducted in preadipocytes with wild-type and the missense variant revealed a CrebRF dependent levels of key metabolic enzymes, lipid accumulation and basal glycolysis (Minster et al., 2016). *Drosophila* also contain a CrebRF ortholog REPTOR which has also been implicated in adaptation to nutritional stress in TORC1 signaling (Tiebe et al., 2015). Both CrebRF and REPTOR are highly induced by starvation and may negatively regulate other Creb3L/CrebA transcription factors. Immunoprecipitation experiments in flies and mammals will further improve our understanding of feeding dependent regulation.

8. List of Materials and Reagents

Material / Reagent	**Supplier**
Sucrose (20%)	Roth, Cat # 4661.3
Yeast Extract (5%)	Serva, Cat# 24540.3
Red Dye (5%)	McCormick
Trizol	Life Technologies, Cat#15596018
Glycogen (200mg/mL)	Thermofischer, Cat#10814010
TURBO DNA-free DNAse	Ambion, M1907
Random Primers	Thermo, 48190011
SuperscriptIII Polymerase	Thermo, 18080085
SYBR Green Master Mix	Life Technologies
Preomics iST sample preparation kit	Preomics, Cat# P.O.00027
Protease Inhibitor Cocktail (PIC)	Roche
60 microns nylon filter	Millipore, SCNY00060
Sepharose protein A beads	GE Healthcare
protein A Dynabeads	Thermo
Salmon Sperm DNA	Sigma Aldrich
AGENCOURT AMPURE XP magnetic beads	Beckmann-Coulter
NEBNext®Ultra II DNA Library Prep Kit	NewEngland Biolabs
NEBNext® Multiplex Oligos	NewEngland Biolabs
Lysis Buffer (Mice Experiments)	BIORAD
pEGFP-C1 Vector	Clonetech
CrebA Genetic Construct	Drosophila Genomics Resource Center, 1623052
SDS-PAGE Protein Ladder	Thermo
HRP 1:1 substrate solution	Merck
Sodium Chloride (NaCl)	Sigma Aldrich
Potassium Chloride (KCl)	Sigma Aldrich
Magnesium Chloride ($MgCl_2$	Sigma Aldrich
Tris-HCl	Sigma Aldrich
EDTA	Sigma Aldrich
EGTA	Sigma Aldrich
Triton-X	Sigma Aldrich
SDS	Sigma Aldrich
DOC	Sigma Aldrich
Antibodies	
Anti-CrebA Guinea Pig	Protein purified In-house, animals inoculated by Eurogentec
Anti-Apoltp Antibody	Susan Eaton Lab
Anti-CV-d Antibody	Bruno Lemaitre Lab
monoclonal anti-Creb3L2	Merck, MABE1018
anti-GAPDH antibody	Genetex, GT239
Anti-RPB3 (RNA-Polymerase II)	Schauer et al., 2013
Fly Strains	
2202U w^{1118}	Boynton and Tully, 1992
Hsp70-Gal4	Brand and Perrimon, 1993
nSyb-Gal4	Bloomington stock center, 458
ppl-Gal4	Bloomington stock center, 58768
TO-Gal4	Dauwalder et al., 2002
UAS-CrebA	Rose et al., 1997
UAS-CrebAshRNAi	Vienna stock center, 330349

UAS-Sec61alpha RNAi	Vienna stock center, 109660

9. List of Figures

10. List of Abbreviations

ER: Endoplasmic Reticulum

CrebA: cAMP Responsive Element Binding Protein A

RNA-seq: Ribonucleic acid sequencing

ChIP: Chromatin Immunoprecipitation

CNS: Central Nervous System

TOR: Target of Rapamycin

TAG: Triacylglyceride

LD: Lipid Droplets

cAMP: Cyclic Adenosine Monophosphate

PKA: Protein Kinase A

ATGL: Adipose Triglyceride Lipase

iCa: Intracellular Calcium

SOCE: Store-Operated Calcium Entry

IPC: Insulin Producing Cell

NPF: Neuropeptide F

APC: Adipokinetic Hormone Producing Cells

DGAT: Diacyleglycerol acetyltransferases

HDL: High Density Lipoproteins

VLDL: Very Low-Density Lipoproteins

SRP: Signal Recognition Particle

TRAP: Translocon Associated Protein

OST: Oligosaccharyl Transferase Complex

TRAM: Translocating-Chain Associated Membrane

SPC: Signal Peptidase Complex

COP: Coat Protein

TF: Transcription Factor

PDX1: Pancreatic Duodenum Homeobox 1

HNF: Hepatocyte Nuclear Factor

ChREBP: Carbohydrate-Responsive Element Binding Protein

ChoRE: Carbohydrate Responsive Element

AAR: Amino Acid Response

GCN2: General Control Nonderepressible 2

ATF4: Activating Transcription Factor 4
PPAR: Peroxisome Proliferator-Activated Receptor
PPRE: PPAR Responsive Element
RXR: Retinoid X Receptor
LBD: Ligand Binding Domain
SREBP: Sterol Regulatory Element Binding Protein
CRE: cAMP Responsive Element
ATB: Adjacent To bZip
TNF: Tumor Necrosis Factor
GET: Guided Entry of Tail-Anchored
PCR: Polymerase Chain Reaction
RT-qPCR: Reverse Transcription Quantitative PCR
LC-MS: Liquid Chromatography Mass Spectrometry
TGF: Transforming Growth Factor
FIT: Female-specific Independent of Transformer
CAFÉ: Capillary Feeding
WT: Wild Type
PCA: Principal Component Analysis
ATP: Adenosine Triphosphate
GO: Gene Ontology
ERAD: ER-Associated Protein Degradation
RING: Rapid Iterative Negative Geotaxis
PIC: Protease Inhibitor Cocktail
ROI: Region of Interest
IP: Immunoprecipitation
Pol2: RNA Polymerase II
GLM: Generalized Linear Regression Model
GFP: Green Fluorescent Protein
GSK: Glycogen Synthase Kinase
FA: Fatty Acids
HCF: Host Cell Factor
CTD: C Terminal Domain

11. Bibliography

Abel, T., Bhatt, R., and Maniatis, T. (1992). A Drosophila CREB/ATF transcriptional activator binds to both fat body- and liver-specific regulatory elements. Genes Dev *6*, 466-480.

Abrams, E.W., and Andrew, D.J. (2005). CrebA regulates secretory activity in the Drosophila salivary gland and epidermis. Development *132*, 2743-2758.

Ahn, J.C., and Sundaram, V. (2019). Obesity and Liver Decompensation. Clin Liver Dis (Hoboken) *14*, 12-15.

Al-Anzi, B., Sapin, V., Waters, C., Zinn, K., Wyman, R.J., and Benzer, S. (2009). Obesity-blocking neurons in Drosophila. Neuron *63*, 329-341.

Ali, A., Hoeflich, K.P., and Woodgett, J.R. (2001). Glycogen synthase kinase-3: properties, functions, and regulation. Chem Rev *101*, 2527-2540.

Arakel, E.C., and Schwappach, B. (2018). Formation of COPI-coated vesicles at a glance. J Cell Sci *131*.

Arrese, E.L., and Soulages, J.L. (2010). Insect fat body: energy, metabolism, and regulation. Annu Rev Entomol *55*, 207-225.

Asada, R., Kanemoto, S., Kondo, S., Saito, A., and Imaizumi, K. (2011). The signalling from endoplasmic reticulum-resident bZIP transcription factors involved in diverse cellular physiology. J Biochem *149*, 507-518.

Aviram, N., and Schuldiner, M. (2017). Targeting and translocation of proteins to the endoplasmic reticulum at a glance. J Cell Sci *130*, 4079-4085.

Baiceanu, A., Mesdom, P., Lagouge, M., and Foufelle, F. (2016). Endoplasmic reticulum proteostasis in hepatic steatosis. Nat Rev Endocrinol *12*, 710-722.

Balla, T., Kim, Y.J., Alvarez-Prats, A., and Pemberton, J. (2019). Lipid Dynamics at Contact Sites Between the Endoplasmic Reticulum and Other Organelles. Annu Rev Cell Dev Biol *35*, 85-109.

Barbosa, S., Carreira, S., Bailey, D., Abaitua, F., and O'Hare, P. (2015). Phosphorylation and SCF-mediated degradation regulate CREB-H transcription of metabolic targets. Mol Biol Cell *26*, 2939-2954.

Barbosa, S., Carreira, S., and O'Hare, P. (2017). GSK-3-mediated phosphorylation couples ER-Golgi transport and nuclear stabilization of the CREB-H transcription factor to mediate apolipoprotein secretion. Mol Biol Cell *28*, 1565-1579.

Barbosa, S., Fasanella, G., Carreira, S., Llarena, M., Fox, R., Barreca, C., Andrew, D., and O'Hare, P. (2013). An orchestrated program regulating secretory pathway genes and cargos by the transmembrane transcription factor CREB-H. Traffic *14*, 382-398.

Baumbach, J., Hummel, P., Bickmeyer, I., Kowalczyk, K.M., Frank, M., Knorr, K., Hildebrandt, A., Riedel, D., Jackle, H., and Kuhnlein, R.P. (2014). A Drosophila in vivo screen identifies store-operated calcium entry as a key regulator of adiposity. Cell Metab *19*, 331-343.

Becker, P.B. (2002). Nucleosome sliding: facts and fiction. EMBO J *21*, 4749-4753.

Becker, P.B., and Wu, C. (1992). Cell-free system for assembly of transcriptionally repressed chromatin from Drosophila embryos. Mol Cell Biol *12*, 2241-2249.

Ben Aicha, S., Lessard, J., Pelletier, M., Fournier, A., Calvo, E., and Labrie, C. (2007). Transcriptional profiling of genes that are regulated by the endoplasmic reticulum-bound transcription factor AIbZIP/CREB3L4 in prostate cells. Physiol Genomics *31*, 295-305.

Berriot-Varoqueaux, N., Aggerbeck, L.P., Samson-Bouma, M., and Wetterau, J.R. (2000). The role of the microsomal triglygeride transfer protein in abetalipoproteinemia. Annu Rev Nutr *20*, 663-697.

Beshel, J., Dubnau, J., and Zhong, Y. (2017). A Leptin Analog Locally Produced in the Brain Acts via a Conserved Neural Circuit to Modulate Obesity-Linked Behaviors in Drosophila. Cell Metab *25*, 208-217.

Bharucha, K.N., Tarr, P., and Zipursky, S.L. (2008). A glucagon-like endocrine pathway in Drosophila modulates both lipid and carbohydrate homeostasis. J Exp Biol *211*, 3103-3110.

Birse, R.T., Choi, J., Reardon, K., Rodriguez, J., Graham, S., Diop, S., Ocorr, K., Bodmer, R., and Oldham, S. (2010). High-fat-diet-induced obesity and heart dysfunction are regulated by the TOR pathway in Drosophila. Cell Metab *12*, 533-544.

Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development *118*, 401-415.

Brown, M.S., and Goldstein, J.L. (1997). The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell *89*, 331-340.

Brunn, G.J., Hudson, C.C., Sekulic, A., Williams, J.M., Hosoi, H., Houghton, P.J., Lawrence, J.C., Jr., and Abraham, R.T. (1997). Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin. Science *277*, 99-101.

Buchon, N., and Osman, D. (2015). All for one and one for all: Regionalization of the Drosophila intestine. Insect Biochem Mol Biol *67*, 2-8.

Buckley, M.S., Kwak, H., Zipfel, W.R., and Lis, J.T. (2014). Kinetics of promoter Pol II on Hsp70 reveal stable pausing and key insights into its regulation. Genes Dev *28*, 14-19.

Buhman, K.K., Chen, H.C., and Farese, R.V., Jr. (2001). The enzymes of neutral lipid synthesis. J Biol Chem *276*, 40369-40372.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Mol Cell *36*, 541-546.

Burnett, J.R., Shan, J., Miskie, B.A., Whitfield, A.J., Yuan, J., Tran, K., McKnight, C.J., Hegele, R.A., and Yao, Z. (2003). A novel nontruncating APOB gene mutation, R463W, causes familial hypobetalipoproteinemia. J Biol Chem *278*, 13442-13452.

Burnett, J.R., Zhong, S., Jiang, Z.G., Hooper, A.J., Fisher, E.A., McLeod, R.S., Zhao, Y., Barrett, P.H., Hegele, R.A., van Bockxmeer, F.M., et al. (2007). Missense mutations in APOB within the betaalpha1 domain of human APOB-100 result in impaired secretion of ApoB and ApoB-containing lipoproteins in familial hypobetalipoproteinemia. J Biol Chem *282*, 24270-24283.

Capotosti, F., Guernier, S., Lammers, F., Waridel, P., Cai, Y., Jin, J., Conaway, J.W., Conaway, R.C., and Herr, W. (2011). O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1. Cell *144*, 376-388.

Cases, S., Smith, S.J., Zheng, Y.W., Myers, H.M., Lear, S.R., Sande, E., Novak, S., Collins, C., Welch, C.B., Lusis, A.J., et al. (1998). Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. Proc Natl Acad Sci U S A *95*, 13018-13023.

Cases, S., Stone, S.J., Zhou, P., Yen, E., Tow, B., Lardizabal, K.D., Voelker, T., and Farese, R.V., Jr. (2001). Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members. J Biol Chem *276*, 38870-38876.

Castellanos, M., Mothi, N., and Munoz, V. (2020). Eukaryotic transcription factors can track and control their target genes using DNA antennas. Nat Commun *11*, 540.

Chan, C.P., Mak, T.Y., Chin, K.T., Ng, I.O., and Jin, D.Y. (2010). N-linked glycosylation is required for optimal proteolytic activation of membrane-bound transcription factor CREB-H. J Cell Sci *123*, 1438-1448.

Chang, T.Y., Chang, C.C., Ohgami, N., and Yamauchi, Y. (2006). Cholesterol sensing, trafficking, and esterification. Annu Rev Cell Dev Biol *22*, 129-157.

Cheng, D., Xu, X., Simon, T., Boudyguina, E., Deng, Z., VerHague, M., Lee, A.H., Shelness, G.S., Weinberg, R.B., and Parks, J.S. (2016). Very Low Density Lipoprotein Assembly Is Required for cAMP-responsive Element-binding Protein H Processing and Hepatic Apolipoprotein A-IV Expression. J Biol Chem *291*, 23793-23803.

Chin, K.T., Zhou, H.J., Wong, C.M., Lee, J.M., Chan, C.P., Qiang, B.Q., Yuan, J.G., Ng, I.O., and Jin, D.Y. (2005). The liver-enriched transcription factor CREB-H is a growth suppressor protein underexpressed in hepatocellular carcinoma. Nucleic Acids Res *33*, 1859-1873.

Chng, W.A., Sleiman, M.S.B., Schupfer, F., and Lemaitre, B. (2014). Transforming growth factor beta/activin signaling functions as a sugar-sensing feedback loop to regulate digestive enzyme expression. Cell Rep *9*, 336-348.

Colombani, J., Raisin, S., Pantalacci, S., Radimerski, T., Montagne, J., and Leopold, P. (2003). A nutrient sensor mechanism controls Drosophila growth. Cell *114*, 739-749.

Conesa, A., and Nueda, M. (2021). maSigPro: Significant Gene Expression Profile Differences in Time Course Gene Expression Data. R package version 1.64.0.

Core, L.J., and Lis, J.T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. Science *319*, 1791-1792.

Cousins, R.J. (1999). Nutritional regulation of gene expression. Am J Med *106*, 20S-23S; discussion 50S-51S.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics *13*, 2513-2526.

Cui, X., Cui, M., Asada, R., Kanemoto, S., Saito, A., Matsuhisa, K., Kaneko, M., and Imaizumi, K. (2016). The androgen-induced protein AIbZIP facilitates proliferation of prostate cancer cells through downregulation of p21 expression. Sci Rep *6*, 37310.

Dandekar, A., Qiu, Y., Kim, H., Wang, J., Hou, X., Zhang, X., Zheng, Z., Mendez, R., Yu, F.S., Kumar, A., et al. (2016). Toll-like Receptor (TLR) Signaling Interacts with CREBH to Modulate High-density Lipoprotein (HDL) in Response to Bacterial Endotoxin. J Biol Chem *291*, 23149-23158.

Danno, H., Ishii, K.A., Nakagawa, Y., Mikami, M., Yamamoto, T., Yabe, S., Furusawa, M., Kumadaki, S., Watanabe, K., Shimizu, H., et al. (2010). The liver-enriched transcription factor CREBH is nutritionally regulated and activated by fatty acids and PPARalpha. Biochem Biophys Res Commun *391*, 1222-1227.

Das, M., and Gursky, O. (2015). Amyloid-Forming Properties of Human Apolipoproteins: Sequence Analyses and Structural Insights. Adv Exp Med Biol *855*, 175-211.

Dauwalder, B., Tsujimoto, S., Moss, J., and Mattox, W. (2002). The Drosophila takeout gene is regulated by the somatic sex-determination pathway and affects male courtship behavior. Genes Dev *16*, 2879-2892.

Deato, M.D., and Tjian, R. (2007). Switching of the core transcription machinery during myogenesis. Genes Dev *21*, 2137-2149.

DenBoer, L.M., Hardy-Smith, P.W., Hogan, M.R., Cockram, G.P., Audas, T.E., and Lu, R. (2005). Luman is capable of binding and activating transcription from the unfolded protein response element. Biochem Biophys Res Commun *331*, 113-119.

Desvergne, B., Michalik, L., and Wahli, W. (2006). Transcriptional regulation of metabolism. Physiol Rev *86*, 465-514.

Desvergne, B., and Wahli, W. (1999). Peroxisome proliferator-activated receptors: nuclear control of metabolism. Endocr Rev *20*, 649-688.

Dietzl, G., Chen, D., Schnorrer, F., Su, K.C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature *448*, 151-156.

Diop, S.B., Bisharat-Kernizan, J., Birse, R.T., Oldham, S., Ocorr, K., and Bodmer, R. (2015). PGC-1/Spargel Counteracts High-Fat-Diet-Induced Obesity and Cardiac Lipotoxicity Downstream of TOR and Brummer ATGL Lipase. Cell Rep *10*, 1572-1584.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell *121*, 873-885.

Dreyer, A.P., Martin, M.M., Fulgham, C.V., Jabr, D.A., Bai, L., Beshel, J., and Cavanaugh, D.J. (2019). A circadian output center controlling feeding:fasting rhythms in Drosophila. PLoS Genet *15*, e1008478.

Ebmeier, C.C., Erickson, B., Allen, B.L., Allen, M.A., Kim, H., Fong, N., Jacobsen, J.R., Liang, K., Shilatifard, A., Dowell, R.D., et al. (2017). Human TFIIH Kinase

CDK7 Regulates Transcription-Associated Chromatin Modifications. Cell Rep *20*, 1173-1186.

Espenshade, P.J., Li, W.P., and Yabe, D. (2002). Sterols block binding of COPII proteins to SCAP, thereby controlling SCAP sorting in ER. Proc Natl Acad Sci U S A *99*, 11694-11699.

Falk, J., Rohde, M., Bekhite, M.M., Neugebauer, S., Hemmerich, P., Kiehntopf, M., Deufel, T., Hubner, C.A., and Beetz, C. (2014). Functional mutation analysis provides evidence for a role of REEP1 in lipid droplet biology. Hum Mutat *35*, 497-504.

Fan, W., Waizenegger, W., Lin, C.S., Sorrentino, V., He, M.X., Wall, C.E., Li, H., Liddle, C., Yu, R.T., Atkins, A.R., et al. (2017). PPARdelta Promotes Running Endurance by Preserving Glucose. Cell Metab *25*, 1186-1193 e1184.

Feizi, A., Gatto, F., Uhlen, M., and Nielsen, J. (2017). Human protein secretory pathway genes are expressed in a tissue-specific pattern to match processing demands of the secretome. NPJ Syst Biol Appl *3*, 22.

Feldman, M., and Richardson, C.T. (1986). Role of thought, sight, smell, and taste of food in the cephalic phase of gastric acid secretion in humans. Gastroenterology *90*, 428-433.

Fox, R.M., Hanlon, C.D., and Andrew, D.J. (2010). The CrebA/Creb3-like transcription factors are major and direct regulators of secretory capacity. J Cell Biol *191*, 479-492.

Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature *461*, 186-192.

Garcia-Barrio, M., Dong, J., Ufano, S., and Hinnebusch, A.G. (2000). Association of GCN1-GCN20 regulatory complex with the N-terminus of eIF2alpha kinase GCN2 is required for GCN2 activation. EMBO J *19*, 1887-1899.

Gargano, J.W., Martin, I., Bhandari, P., and Grotewiel, M.S. (2005). Rapid iterative negative geotaxis (RING): a new method for assessing age-related locomotor decline in Drosophila. Exp Gerontol *40*, 386-395.

Geminard, C., Rulifson, E.J., and Leopold, P. (2009). Remote control of insulin secretion by fat cells in Drosophila. Cell Metab *10*, 199-207.

Gemmer, M., and Forster, F. (2020). A clearer picture of the ER translocon complex. J Cell Sci *133*.

Gentile, C.L., Wang, D., Pfaffenbach, K.T., Cox, R., Wei, Y., and Pagliassotti, M.J. (2010). Fatty acids regulate CREBh via transcriptional mechanisms that are dependent on proteasome activity and insulin. Mol Cell Biochem *344*, 99-107.

Gerrish, K., Van Velkinburgh, J.C., and Stein, R. (2004). Conserved transcriptional regulatory domains of the pdx-1 gene. Mol Endocrinol *18*, 533-548.

Gershman, B., Puig, O., Hang, L., Peitzsch, R.M., Tatar, M., and Garofalo, R.S. (2007). High-resolution dynamics of the transcriptional response to nutrition in Drosophila: a key role for dFOXO. Physiol Genomics *29*, 24-34.

Gillon, A.D., Latham, C.F., and Miller, E.A. (2012). Vesicle-mediated ER export of proteins and lipids. Biochim Biophys Acta *1821*, 1040-1049.

Gingras, A.C., Gygi, S.P., Raught, B., Polakiewicz, R.D., Abraham, R.T., Hoekstra, M.F., Aebersold, R., and Sonenberg, N. (1999). Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. Genes Dev *13*, 1422-1437.

Gray, S.M., Meijer, R.I., and Barrett, E.J. (2014). Insulin regulates brain function, but how does it get there? Diabetes *63*, 3992-3997.

Gronke, S., Mildner, A., Fellert, S., Tennagels, N., Petry, S., Muller, G., Jackle, H., and Kuhnlein, R.P. (2005). Brummer lipase is an evolutionary conserved fat storage regulator in Drosophila. Cell Metab *1*, 323-330.

Gronke, S., Muller, G., Hirsch, J., Fellert, S., Andreou, A., Haase, T., Jackle, H., and Kuhnlein, R.P. (2007). Dual lipolytic control of body fat storage and mobilization in Drosophila. PLoS Biol *5*, e137.

Guelman, S., Suganuma, T., Florens, L., Swanson, S.K., Kiesecker, C.L., Kusch, T., Anderson, S., Yates, J.R., 3rd, Washburn, M.P., Abmayr, S.M., et al. (2006). Host cell factor and an uncharacterized SANT domain protein are stable components of ATAC, a novel dAda2A/dGcn5-containing histone acetyltransferase complex in Drosophila. Mol Cell Biol *26*, 871-882.

Guna, A., Volkmar, N., Christianson, J.C., and Hegde, R.S. (2018). The ER membrane protein complex is a transmembrane domain insertase. Science *359*, 470-473.

Hara, K., Maruki, Y., Long, X., Yoshino, K., Oshiro, N., Hidayat, S., Tokunaga, C., Avruch, J., and Yonezawa, K. (2002). Raptor, a binding partner of target of rapamycin (TOR), mediates TOR action. Cell *110*, 177-189.

Harding, H.P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M., and Ron, D. (2000). Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol Cell *6*, 1099-1108.

Harding, H.P., Zhang, Y., Zeng, H., Novoa, I., Lu, P.D., Calfon, M., Sadri, N., Yun, C., Popko, B., Paules, R., et al. (2003). An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. Mol Cell *11*, 619-633.

Haro, D., Marrero, P.F., and Relat, J. (2019). Nutritional Regulation of Gene Expression: Carbohydrate-, Fat- and Amino Acid-Dependent Modulation of Transcriptional Activity. Int J Mol Sci *20*.

Havula, E., Teesalu, M., Hyotylainen, T., Seppala, H., Hasygar, K., Auvinen, P., Oresic, M., Sandmann, T., and Hietakangas, V. (2013). Mondo/ChREBP-Mlx-regulated transcriptional network is essential for dietary sugar tolerance in Drosophila. PLoS Genet *9*, e1003438.

Hayhurst, G.P., Lee, Y.H., Lambert, G., Ward, J.M., and Gonzalez, F.J. (2001). Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. Mol Cell Biol *21*, 1393-1403.

Hegde, R.S., Voigt, S., and Lingappa, V.R. (1998). Regulation of protein topology by trans-acting factors at the endoplasmic reticulum. Mol Cell *2*, 85-91.

Hino, K., Saito, A., Kido, M., Kanemoto, S., Asada, R., Takai, T., Cui, M., Cui, X., and Imaizumi, K. (2014). Master regulator for chondrogenesis, Sox9, regulates transcriptional activation of the endoplasmic reticulum stress transducer BBF2H7/CREB3L2 in chondrocytes. J Biol Chem *289*, 13810-13820.

Hirose, Y., and Ohkuma, Y. (2007). Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. J Biochem *141*, 601-608.

Holst, D., Luquet, S., Nogueira, V., Kristiansen, K., Leverve, X., and Grimaldi, P.A. (2003). Nutritional regulation and role of peroxisome proliferator-activated receptor delta in fatty acid catabolism in skeletal muscle. Biochim Biophys Acta *1633*, 43-50.

Honti, V., Csordas, G., Kurucz, E., Markus, R., and Ando, I. (2014). The cell-mediated immunity of Drosophila melanogaster: hemocyte lineages, immune compartments, microanatomy and regulation. Dev Comp Immunol *42*, 47-56.

Hsu, H.T., Chen, H.M., Yang, Z., Wang, J., Lee, N.K., Burger, A., Zaret, K., Liu, T., Levine, E., and Mango, S.E. (2015). TRANSCRIPTION. Recruitment of RNA polymerase II by the pioneer transcription factor PHA-4. Science *348*, 1372-1376.

Huang, J.H., and Douglas, A.E. (2015). Consumption of dietary sugar by gut bacteria determines Drosophila lipid content. Biol Lett *11*, 20150469.

Hussain, M.M., Shi, J., and Dreizen, P. (2003). Microsomal triglyceride transfer protein and its role in apoB-lipoprotein assembly. J Lipid Res *44*, 22-32.

Hutton, J.C., and O'Brien, R.M. (2009). Glucose-6-phosphatase catalytic subunit gene family. J Biol Chem *284*, 29241-29245.

Iijima, K., Zhao, L., Shenton, C., and Iijima-Ando, K. (2009). Regulation of energy stores and feeding by neuronal and peripheral CREB activity in Drosophila. PLoS One *4*, e8498.

Imbalzano, A.N., Zaret, K.S., and Kingston, R.E. (1994). Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. J Biol Chem *269*, 8280-8286.

Ishikawa, T., Toyama, T., Nakamura, Y., Tamada, K., Shimizu, H., Ninagawa, S., Okada, T., Kamei, Y., Ishikawa-Fujiwara, T., Todo, T., et al. (2017). UPR transducer BBF2H7 allows export of type II collagen in a cargo- and developmental stage-specific manner. J Cell Biol *216*, 1761-1774.

Itskov, P.M., and Ribeiro, C. (2013). The dilemmas of the gourmet fly: the molecular and neuronal mechanisms of feeding and nutrient decision making in Drosophila. Front Neurosci *7*, 12.

Jang, W.G., Jeong, B.C., Kim, E.J., Choi, H., Oh, S.H., Kim, D.K., Koo, S.H., Choi, H.S., and Koh, J.T. (2015). Cyclic AMP Response Element-binding Protein H (CREBH) Mediates the Inhibitory Actions of Tumor Necrosis Factor alpha in Osteoblast Differentiation by Stimulating Smad1 Degradation. J Biol Chem *290*, 13556-13566.

Jang, W.G., Kim, E.J., and Koh, J.T. (2011). Tunicamycin negatively regulates BMP2-induced osteoblast differentiation through CREBH expression in MC3T3E1 cells. BMB Rep *44*, 735-740.

Jensen, D., and Schekman, R. (2011). COPII-mediated vesicle formation at a glance. J Cell Sci *124*, 1-4.

Johnson, D.M., Wells, M.B., Fox, R., Lee, J.S., Loganathan, R., Levings, D., Bastien, A., Slattery, M., and Andrew, D.J. (2020). CrebA increases secretory capacity through direct transcriptional regulation of the secretory machinery, a subset of secretory cargo, and other key regulators. Traffic *21*, 560-577.

Jones, J.R., Barrick, C., Kim, K.A., Lindner, J., Blondeau, B., Fujimoto, Y., Shiota, M., Kesterson, R.A., Kahn, B.B., and Magnuson, M.A. (2005). Deletion of PPARgamma in adipose tissues of mice protects against high fat diet-induced obesity and insulin resistance. Proc Natl Acad Sci U S A *102*, 6207-6212.

Kamei, Y., Ohizumi, H., Fujitani, Y., Nemoto, T., Tanaka, T., Takahashi, N., Kawada, T., Miyoshi, M., Ezaki, O., and Kakizuka, A. (2003). PPARgamma coactivator 1beta/ERR ligand 1 is an ERR protein ligand, whose expression induces a high-

energy expenditure and antagonizes obesity. Proc Natl Acad Sci U S A *100*, 12378-12383.

Kellenberger, C., Hendrickson, T.L., and Imperiali, B. (1997). Structural and functional analysis of peptidyl oligosaccharyl transferase inhibitors. Biochemistry *36*, 12554-12559.

Kennelly, J.P., and Tontonoz, P. (2022). Cholesterol Transport to the Endoplasmic Reticulum. Cold Spring Harb Perspect Biol.

Khan, H.A., and Margulies, C.E. (2019). The Role of Mammalian Creb3-Like Transcription Factors in Response to Nutrients. Front Genet *10*, 591.

Kikuchi, T., Orihara, K., Oikawa, F., Han, S.I., Kuba, M., Okuda, K., Satoh, A., Osaki, Y., Takeuchi, Y., Aita, Y., et al. (2016). Intestinal CREBH overexpression prevents high-cholesterol diet-induced hypercholesterolemia by reducing Npc1l1 expression. Mol Metab *5*, 1092-1102.

Kilberg, M.S., Shan, J., and Su, N. (2009). ATF4-dependent transcription mediates signaling of amino acid limitation. Trends Endocrinol Metab *20*, 436-443.

Kim, H., Zheng, Z., Walker, P.D., Kapatos, G., and Zhang, K. (2017). CREBH Maintains Circadian Glucose Homeostasis by Regulating Hepatic Glycogenolysis and Gluconeogenesis. Mol Cell Biol *37*, e00048-00017.

Kim, H.C., Choi, K.C., Choi, H.K., Kang, H.B., Kim, M.J., Lee, Y.H., Lee, O.H., Lee, J., Kim, Y.J., Jun, W., et al. (2010). HDAC3 selectively represses CREB3-mediated transcription and migration of metastatic breast cancer cells. Cell Mol Life Sci *67*, 3499-3510.

Kim, J., and Guan, K.L. (2019). mTOR as a central hub of nutrient signalling and cell growth. Nat Cell Biol *21*, 63-71.

Kim, S.K., and Rulifson, E.J. (2004). Conserved mechanisms of glucose sensing and regulation by Drosophila corpora cardiaca cells. Nature *431*, 316-320.

Kizer, K.O., Phatnani, H.P., Shibata, Y., Hall, H., Greenleaf, A.L., and Strahl, B.D. (2005). A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol *25*, 3305-3316.

Klemm, R.W., Norton, J.P., Cole, R.A., Li, C.S., Park, S.H., Crane, M.M., Li, L., Jin, D., Boye-Doe, A., Liu, T.Y., et al. (2013). A conserved role for atlastin GTPases in regulating lipid droplet size. Cell Rep *3*, 1465-1475.

Klok, M.D., Jakobsdottir, S., and Drent, M.L. (2007). The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review. Obes Rev *8*, 21-34.

Koerner, C.M., Roberts, B.S., and Neher, S.B. (2019). Endoplasmic reticulum quality control in lipoprotein metabolism. Mol Cell Endocrinol *498*, 110547.

Kolodziej, P.A., Woychik, N., Liao, S.M., and Young, R.A. (1990). RNA polymerase II subunit composition, stoichiometry, and phosphorylation. Mol Cell Biol *10*, 1915-1920.

Kondo, S., Murakami, T., Tatsumi, K., Ogata, M., Kanemoto, S., Otori, K., Iseki, K., Wanaka, A., and Imaizumi, K. (2005). OASIS, a CREB/ATF-family member, modulates UPR signalling in astrocytes. Nat Cell Biol *7*, 186-194.

Kondo, S., Saito, A., Hino, S., Murakami, T., Ogata, M., Kanemoto, S., Nara, S., Yamashita, A., Yoshinaga, K., Hara, H., et al. (2007). BBF2H7, a novel transmembrane bZIP transcription factor, is a new type of endoplasmic reticulum stress transducer. Mol Cell Biol *27*, 1716-1729.

Lang, S., Pfeffer, S., Lee, P.H., Cavalie, A., Helms, V., Forster, F., and Zimmermann, R. (2017). An Update on Sec61 Channel Functions, Mechanisms, and Related Diseases. Front Physiol *8*, 887.

Latchman, D.S. (1997). Transcription factors: an overview. Int J Biochem Cell Biol *29*, 1305-1312.

LeBlanc, J., and Cabanac, M. (1989). Cephalic postprandial thermogenesis in human subjects. Physiol Behav *46*, 479-482.

Lee, G., and Park, J.H. (2004). Hemolymph sugar homeostasis and starvation-induced hyperactivity affected by genetic manipulations of the adipokinetic hormone-encoding gene in Drosophila melanogaster. Genetics *167*, 311-323.

Lee, J.H., Giannikopoulos, P., Duncan, S.A., Wang, J., Johansen, C.T., Brown, J.D., Plutzky, J., Hegele, R.A., Glimcher, L.H., and Lee, A.H. (2011). The transcription factor cyclic AMP-responsive element-binding protein H regulates triglyceride metabolism. Nat Med *17*, 812-815.

Lee, M.W., Chanda, D., Yang, J., Oh, H., Kim, S.S., Yoon, Y.S., Hong, S., Park, K.G., Lee, I.K., Choi, C.S., et al. (2010). Regulation of hepatic gluconeogenesis by an ER-bound transcription factor, CREBH. Cell Metab *11*, 331-339.

Lee, S., Bao, H., Ishikawa, Z., Wang, W., and Lim, H.Y. (2017). Cardiomyocyte Regulation of Systemic Lipid Metabolism by the Apolipoprotein B-Containing Lipoproteins in Drosophila. PLoS Genet *13*, e1006555.

Lee, S.S., Pineau, T., Drago, J., Lee, E.J., Owens, J.W., Kroetz, D.L., Fernandez-Salguero, P.M., Westphal, H., and Gonzalez, F.J. (1995). Targeted disruption of the alpha isoform of the peroxisome proliferator-activated receptor gene in mice results in abolishment of the pleiotropic effects of peroxisome proliferators. Mol Cell Biol *15*, 3012-3022.

Lehmann, J.M., Moore, L.B., Smith-Oliver, T.A., Wilkison, W.O., Willson, T.M., and Kliewer, S.A. (1995). An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor gamma (PPAR gamma). J Biol Chem *270*, 12953-12956.

Lemaitre, B., and Hoffmann, J. (2007). The host defense of Drosophila melanogaster. Annu Rev Immunol *25*, 697-743.

Lemberg, M.K. (2011). Intramembrane proteolysis in regulated protein trafficking. Traffic *12*, 1109-1118.

Li, J., Song, J., Zaytseva, Y.Y., Liu, Y., Rychahou, P., Jiang, K., Starr, M.E., Kim, J.T., Harris, J.W., Yiannikouris, F.B., et al. (2016). An obligatory role for neurotensin in high-fat-diet-induced obesity. Nature *533*, 411-415.

Liang, Y., Vogel, J.L., Narayanan, A., Peng, H., and Kristie, T.M. (2009). Inhibition of the histone demethylase LSD1 blocks alpha-herpesvirus lytic replication and reactivation from latency. Nat Med *15*, 1312-1317.

Lin, Q., Ruuska, S.E., Shaw, N.S., Dong, D., and Noy, N. (1999). Ligand selectivity of the peroxisome proliferator-activated receptor alpha. Biochemistry *38*, 185-190.

Liu, Y., Liao, S., Veenstra, J.A., and Nassel, D.R. (2016). Drosophila insulin-like peptide 1 (DILP1) is transiently expressed during non-feeding stages and reproductive dormancy. Sci Rep *6*, 26620.

Loos, R.J. (2016). CREBRF variant increases obesity risk and protects against diabetes in Samoans. Nat Genet *48*, 976-978.

Lopata, A., Kniss, A., Lohr, F., Rogov, V.V., and Dotsch, V. (2020). Ubiquitination in the ERAD Process. Int J Mol Sci *21*.

Lopez, A.B., Wang, C., Huang, C.C., Yaman, I., Li, Y., Chakravarty, K., Johnson, P.F., Chiang, C.M., Snider, M.D., Wek, R.C., et al. (2007). A feedback transcriptional mechanism controls the level of the arginine/lysine transporter cat-1 during amino acid starvation. Biochem J *402*, 163-173.

Lu, R., Yang, P., O'Hare, P., and Misra, V. (1997). Luman, a new member of the CREB/ATF family, binds to herpes simplex virus VP16-associated host cellular factor. Mol Cell Biol *17*, 5117-5126.

Lu, R., Yang, P., Padmakumar, S., and Misra, V. (1998). The herpesvirus transactivator VP16 mimics a human basic domain leucine zipper protein, luman, in its interaction with HCF. J Virol *72*, 6291-6297.

Ludtke, A., Buettner, J., Wu, W., Muchir, A., Schroeter, A., Zinn-Justin, S., Spuler, S., Schmidt, H.H., and Worman, H.J. (2007). Peroxisome proliferator-activated receptor-gamma C190S mutation causes partial lipodystrophy. J Clin Endocrinol Metab *92*, 2248-2255.

Luebke-Wheeler, J., Zhang, K., Battle, M., Si-Tayeb, K., Garrison, W., Chhinder, S., Li, J., Kaufman, R.J., and Duncan, S.A. (2008). Hepatocyte nuclear factor 4alpha is implicated in endoplasmic reticulum stress-induced acute phase response by regulating expression of cyclic adenosine monophosphate responsive element binding protein H. Hepatology *48*, 1242-1250.

MacFarlane, W.M., Read, M.L., Gilligan, M., Bujalska, I., and Docherty, K. (1994). Glucose modulates the binding activity of the beta-cell transcription factor IUF1 in a phosphorylation-dependent manner. Biochem J *303 (Pt 2)*, 625-631.

Matsuda, M., Korn, B.S., Hammer, R.E., Moon, Y.A., Komuro, R., Horton, J.D., Goldstein, J.L., Brown, M.S., and Shimomura, I. (2001). SREBP cleavage-activating protein (SCAP) is required for increased lipid synthesis in liver induced by cholesterol deprivation and insulin elevation. Genes Dev *15*, 1206-1216.

Maus, M., Cuk, M., Patel, B., Lian, J., Ouimet, M., Kaufmann, U., Yang, J., Horvath, R., Hornig-Do, H.T., Chrzanowska-Lightowlers, Z.M., et al. (2017). Store-Operated Ca(2+) Entry Controls Induction of Lipolysis and the Transcriptional Reprogramming to Lipid Metabolism. Cell Metab *25*, 698-712.

Mayer, J. (1955). Regulation of energy intake and the body weight: the glucostatic theory and the lipostatic hypothesis. Ann N Y Acad Sci *63*, 15-43.

McCaughey, J., and Stephens, D.J. (2019). ER-to-Golgi Transport: A Sizeable Problem. Trends Cell Biol *29*, 940-953.

McGilvray, P.T., Anghel, S.A., Sundaram, A., Zhong, F., Trnka, M.J., Fuller, J.R., Hu, H., Burlingame, A.L., and Keenan, R.J. (2020). An ER translocon for multi-pass membrane protein biogenesis. Elife *9*.

Melville, D.B., Montero-Balaguer, M., Levic, D.S., Bradley, K., Smith, J.R., Hatzopoulos, A.K., and Knapik, E.W. (2011). The feelgood mutation in zebrafish dysregulates COPII-dependent secretion of select extracellular matrix proteins in skeletal morphogenesis. Dis Model Mech *4*, 763-776.

Minehira, K., Young, S.G., Villanueva, C.J., Yetukuri, L., Oresic, M., Hellerstein, M.K., Farese, R.V., Jr., Horton, J.D., Preitner, F., Thorens, B., et al. (2008). Blocking

VLDL secretion causes hepatic steatosis but does not affect peripheral lipid stores or insulin sensitivity in mice. J Lipid Res *49*, 2038-2044.

Minster, R.L., Hawley, N.L., Su, C.T., Sun, G., Kershaw, E.E., Cheng, H., Buhule, O.D., Lin, J., Reupena, M.S., Viali, S., et al. (2016). A thrifty variant in CREBRF strongly influences body mass index in Samoans. Nat Genet *48*, 1049-1054.

Misra, V., Rapin, N., Akhova, O., Bainbridge, M., and Korchinski, P. (2005). Zhangfei is a potent and specific inhibitor of the host cell factor-binding transcription factor Luman. J Biol Chem *280*, 15257-15266.

Mitchell, P.J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371-378.

Muniz, L., Nicolas, E., and Trouche, D. (2021). RNA polymerase II speed: a key player in controlling and adapting transcriptome composition. EMBO J *40*, e105740.

Murakami, T., Saito, A., Hino, S., Kondo, S., Kanemoto, S., Chihara, K., Sekiya, H., Tsumagari, K., Ochiai, K., Yoshinaga, K., et al. (2009). Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation. Nat Cell Biol *11*, 1205-1211.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nat Genet *39*, 1507-1511.

Nagamori, I., Yomogida, K., Ikawa, M., Okabe, M., Yabuta, N., and Nojima, H. (2006). The testes-specific bZip type transcription factor Tisp40 plays a role in ER stress responses and chromatin packaging during spermiogenesis. Genes Cells *11*, 1161-1171.

Nakagawa, Y., Satoh, A., Tezuka, H., Han, S.I., Takei, K., Iwasaki, H., Yatoh, S., Yahagi, N., Suzuki, H., Iwasaki, Y., et al. (2016). CREB3L3 controls fatty acid oxidation and ketogenesis in synergy with PPARalpha. Sci Rep *6*, 39182.

Ni, Z., Saunders, A., Fuda, N.J., Yao, J., Suarez, J.R., Webb, W.W., and Lis, J.T. (2008). P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. Mol Cell Biol *28*, 1161-1170.

Nitta, K.R., Jolma, A., Yin, Y., Morgunova, E., Kivioja, T., Akhtar, J., Hens, K., Toivonen, J., Deplancke, B., Furlong, E.E., et al. (2015). Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. Elife *4*.

Ohlsson, H., Karlsson, K., and Edlund, T. (1993). IPF1, a homeodomain-containing transactivator of the insulin gene. EMBO J *12*, 4251-4259.

Oka, Y., Asano, T., Shibasaki, Y., Lin, J.L., Tsukuda, K., Katagiri, H., Akanuma, Y., and Takaku, F. (1990). C-terminal truncated glucose transporter is locked into an inward-facing form without transport activity. Nature *345*, 550-553.

Omori, Y., Imai, J., Suzuki, Y., Watanabe, S., Tanigami, A., and Sugano, S. (2002). OASIS is a transcriptional activator of CREB/ATF family with a transmembrane domain. Biochem Biophys Res Commun *293*, 470-477.

Ortega-Prieto, P., and Postic, C. (2019). Carbohydrate Sensing Through the Transcription Factor ChREBP. Front Genet *10*, 472.

Palm, W., Sampaio, J.L., Brankatschk, M., Carvalho, M., Mahmoud, A., Shevchenko, A., and Eaton, S. (2012). Lipoproteins in Drosophila melanogaster--assembly, function, and influence on tissue lipid composition. PLoS Genet *8*, e1002828.

Pan, Y.X., Chen, H., Thiaville, M.M., and Kilberg, M.S. (2007). Activation of the ATF3 gene through a co-ordinated amino acid-sensing response programme that controls transcriptional regulation of responsive genes following amino acid limitation. Biochem J *401*, 299-307.

Park, Y., Reyna-Neyra, A., Philippe, L., and Thoreen, C.C. (2017). mTORC1 Balances Cellular Amino Acid Supply with Demand for Protein Synthesis through Post-transcriptional Control of ATF4. Cell Rep *19*, 1083-1090.

Patel, A.B., Louder, R.K., Greber, B.J., Grunberg, S., Luo, J., Fang, J., Liu, Y., Ranish, J., Hahn, S., and Nogales, E. (2018). Structure of human TFIID and mechanism of TBP loading onto promoter DNA. Science *362*.

Patel, R.T., Soulages, J.L., and Arrese, E.L. (2006). Adipokinetic hormone-induced mobilization of fat body triglyceride stores in Manduca sexta: role of TG-lipase and lipid droplets. Arch Insect Biochem Physiol *63*, 73-81.

Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. Mol Cell *23*, 297-305.

Peterson, C.L., and Workman, J.L. (2000). Promoter targeting and chromatin remodeling by the SWI/SNF complex. Curr Opin Genet Dev *10*, 187-192.

Pichaud, F., Treisman, J., and Desplan, C. (2001). Reinventing a common strategy for patterning the eye. Cell *105*, 9-12.

Pospisilik, J.A., Schramek, D., Schnidar, H., Cronin, S.J., Nehme, N.T., Zhang, X., Knauf, C., Cani, P.D., Aumayr, K., Todoric, J., et al. (2010). Drosophila genome-wide obesity screen reveals hedgehog as a determinant of brown versus white adipose cell fate. Cell *140*, 148-160.

Power, M.L., and Schulkin, J. (2008). Anticipatory physiological regulation in feeding biology: cephalic phase responses. Appetite *50*, 194-206.

Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. Nature *386*, 569-577.

Puigserver, P., and Spiegelman, B.M. (2003). Peroxisome proliferator-activated receptor-gamma coactivator 1 alpha (PGC-1 alpha): transcriptional coactivator and metabolic regulator. Endocr Rev *24*, 78-90.

Rajan, A., and Perrimon, N. (2012). Drosophila cytokine unpaired 2 regulates physiological homeostasis by remotely controlling insulin secretion. Cell *151*, 123-137.

Rios-Barrera, L.D., Sigurbjornsdottir, S., Baer, M., and Leptin, M. (2017). Dual function for Tango1 in secretion of bulky cargo and in ER-Golgi morphology. Proc Natl Acad Sci U S A *114*, E10389-E10398.

Ro, J., Harvanek, Z.M., and Pletcher, S.D. (2014). FLIC: high-throughput, continuous analysis of feeding behaviors in Drosophila. PLoS One *9*, e101107.

Rose, R.E., Gallaher, N.M., Andrew, D.J., Goodman, R.H., and Smolik, S.M. (1997). The CRE-binding protein dCREB-A is required for Drosophila embryonic development. Genetics *146*, 595-606.

Sadowski, C.L., Henry, R.W., Lobo, S.M., and Hernandez, N. (1993). Targeting TBP to a non-TATA box cis-regulatory element: a TBP-containing complex activates transcription from snRNA promoters through the PSE. Genes Dev *7*, 1535-1548.

Saito, A., Hino, S., Murakami, T., Kanemoto, S., Kondo, S., Saitoh, M., Nishimura, R., Yoneda, T., Furuichi, T., Ikegawa, S., et al. (2009). Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis. Nat Cell Biol *11*, 1197-1204.

Sakai, J., Nohturfft, A., Goldstein, J.L., and Brown, M.S. (1998). Cleavage of sterol regulatory element-binding proteins (SREBPs) at site-1 requires interaction with SREBP cleavage-activating protein. Evidence from in vivo competition studies. J Biol Chem *273*, 5785-5793.

Sampieri, L., Di Giusto, P., and Alvarez, C. (2019). CREB3 Transcription Factors: ER-Golgi Stress Transducers as Hubs for Cellular Homeostasis. Front Cell Dev Biol *7*, 123.

Sancak, Y., Peterson, T.R., Shaul, Y.D., Lindquist, R.A., Thoreen, C.C., Bar-Peled, L., and Sabatini, D.M. (2008). The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. Science *320*, 1496-1501.

Schauer, T., Schwalie, P.C., Handley, A., Margulies, C.E., Flicek, P., and Ladurner, A.G. (2013). CAST-ChIP maps cell-type-specific chromatin states in the Drosophila central nervous system. Cell Rep *5*, 271-282.

Schier, A.C., and Taatjes, D.J. (2020). Structure and mechanism of the RNA polymerase II transcription machinery. Genes Dev *34*, 465-488.

Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., Schwappach, B., and Weissman, J.S. (2008). The GET complex mediates insertion of tail-anchored proteins into the ER membrane. Cell *134*, 634-645.

Schwedes, C.C., and Carney, G.E. (2012). Ecdysone signaling in adult Drosophila melanogaster. J Insect Physiol *58*, 293-302.

Seegmiller, A.C., Dobrosotskaya, I., Goldstein, J.L., Ho, Y.K., Brown, M.S., and Rawson, R.B. (2002). The SREBP pathway in Drosophila: regulation by palmitate, not sterols. Dev Cell *2*, 229-238.

Sehgal, A., Rothenfluh-Hilfiker, A., Hunter-Ensor, M., Chen, Y., Myers, M.P., and Young, M.W. (1995). Rhythmic expression of timeless: a basis for promoting circadian cycles in period gene autoregulation. Science *270*, 808-810.

Sellin, J., Schulze, H., Paradis, M., Gosejacob, D., Papan, C., Shevchenko, A., Psathaki, O.E., Paululat, A., Thielisch, M., Sandhoff, K., et al. (2017). Characterization of Drosophila Saposin-related mutants as a model for lysosomal sphingolipid storage diseases. Dis Model Mech *10*, 737-750.

Shen, X., Ellis, R.E., Lee, K., Liu, C.Y., Yang, K., Solomon, A., Yoshida, H., Morimoto, R., Kurnit, D.M., Mori, K., et al. (2001). Complementary signaling pathways regulate the unfolded protein response and are required for C. elegans development. Cell *107*, 893-903.

Shimano, H., and Sato, R. (2017). SREBP-regulated lipid metabolism: convergent physiology - divergent pathophysiology. Nat Rev Endocrinol *13*, 710-730.

Shin, D.Y., Chung, J., Joe, Y., Pae, H.O., Chang, K.C., Cho, G.J., Ryter, S.W., and Chung, H.T. (2012). Pretreatment with CO-releasing molecules suppresses hepcidin expression during inflammation and endoplasmic reticulum stress through inhibition of the STAT3 and CREBH pathways. Blood *119*, 2523-2532.

Shurtleff, M.J., Itzhak, D.N., Hussmann, J.A., Schirle Oakdale, N.T., Costa, E.A., Jonikas, M., Weibezahn, J., Popova, K.D., Jan, C.H., Sinitcyn, P., et al. (2018). The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. Elife *7*.

Sieber, M.H., and Thummel, C.S. (2009). The DHR96 nuclear receptor controls triacylglycerol homeostasis in Drosophila. Cell Metab *10*, 481-490.

Smith, E.R., Cayrou, C., Huang, R., Lane, W.S., Cote, J., and Lucchesi, J.C. (2005). A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. Mol Cell Biol *25*, 9175-9188.

Smolik, S.M., Rose, R.E., and Goodman, R.H. (1992). A cyclic AMP-responsive element-binding transcriptional activator in Drosophila melanogaster, dCREB-A, is a member of the leucine zipper family. Mol Cell Biol *12*, 4123-4131.

Song, Y., Zhao, M., Cheng, X., Shen, J., Khound, R., Zhang, K., and Su, Q. (2017). CREBH mediates metabolic inflammation to hepatic VLDL overproduction and hyperlipoproteinemia. J Mol Med (Berl) *95*, 839-849.

Souder, D.C., and Anderson, R.M. (2019). An expanding GSK3 network: implications for aging research. Geroscience *41*, 369-382.

Spilianakis, C., Kretsovali, A., Agalioti, T., Makatounakis, T., Thanos, D., and Papamatheakis, J. (2003). CIITA regulates transcription onset viaSer5-phosphorylation of RNA Pol II. EMBO J *22*, 5125-5136.

Spitz, F., and Furlong, E.E. (2012). Transcription factors: from enhancer binding to developmental control. Nat Rev Genet *13*, 613-626.

Subramanian, M., Metya, S.K., Sadaf, S., Kumar, S., Schwudke, D., and Hasan, G. (2013). Altered lipid homeostasis in Drosophila InsP3 receptor mutants leads to obesity and hyperphagia. Dis Model Mech *6*, 734-744.

Sujkowski, A., Saunders, S., Tinkerhess, M., Piazza, N., Jennens, J., Healy, L., Zheng, L., and Wessells, R. (2012). dFatp regulates nutrient distribution and long-term physiology in Drosophila. Aging Cell *11*, 921-932.

Sun, J., Liu, C., Bai, X., Li, X., Li, J., Zhang, Z., Zhang, Y., Guo, J., and Li, Y. (2017). Drosophila FIT is a protein-specific satiety hormone essential for feeding control. Nat Commun *8*, 14161.

Sznaidman, M.L., Haffner, C.D., Maloney, P.R., Fivush, A., Chao, E., Goreham, D., Sierra, M.L., LeGrumelec, C., Xu, H.E., Montana, V.G., et al. (2003). Novel selective small molecule agonists for peroxisome proliferator-activated receptor delta (PPARdelta)--synthesis and biological activity. Bioorg Med Chem Lett *13*, 1517-1521.

Tepper, B., and IT., B. (2020). Taste, Nutrition, and Health. Nutrients *12(1)*.

Thiam, A.R., and Foret, L. (2016). The physics of lipid droplet nucleation, growth and budding. Biochim Biophys Acta *1861*, 715-722.

Thomas, H.E., Stunnenberg, H.G., and Stewart, A.F. (1993). Heterodimerization of the Drosophila ecdysone receptor with retinoid X receptor and ultraspiracle. Nature *362*, 471-475.

Tiebe, M., Lutz, M., De La Garza, A., Buechling, T., Boutros, M., and Teleman, A.A. (2015). REPTOR and REPTOR-BP Regulate Organismal Metabolism and Transcription Downstream of TORC1. Dev Cell *33*, 272-284.

Tiefenbock, S.K., Baltzer, C., Egli, N.A., and Frei, C. (2010). The Drosophila PGC-1 homologue Spargel coordinates mitochondrial activity to insulin signalling. EMBO J *29*, 171-183.

Troha, K., Im, J.H., Revah, J., Lazzaro, B.P., and Buchon, N. (2018). Comparative transcriptomics reveals CrebA as a novel regulator of infection tolerance in D. melanogaster. PLoS Pathog *14*, e1006847.

Uldry, M., Yang, W., St-Pierre, J., Lin, J., Seale, P., and Spiegelman, B.M. (2006). Complementary action of the PGC-1 coactivators in mitochondrial biogenesis and brown fat differentiation. Cell Metab *3*, 333-341.

Vattem, K.M., and Wek, R.C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc Natl Acad Sci U S A *101*, 11269-11274.

Vecchi, C., Montosi, G., Garuti, C., Corradini, E., Sabelli, M., Canali, S., and Pietrangelo, A. (2014). Gluconeogenic signals regulate iron homeostasis via hepcidin in mice. Gastroenterology *146*, 1060-1069.

Vellanki, R.N., Zhang, L., Guney, M.A., Rocheleau, J.V., Gannon, M., and Volchuk, A. (2010). OASIS/CREB3L1 induces expression of genes involved in extracellular matrix production but not classical endoplasmic reticulum stress response genes in pancreatic beta-cells. Endocrinology *151*, 4146-4157.

Vinson, C., Acharya, A., and Taparowsky, E.J. (2006). Deciphering B-ZIP transcription factor interactions in vitro and in vivo. Biochim Biophys Acta *1759*, 4-12.

Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., and Cramer, P. (2018). Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature *560*, 607-612.

Waddell, I.D., and Burchell, A. (1991). Transverse topology of glucose-6-phosphatase in rat hepatic endoplasmic reticulum. Biochem J *275 (Pt 1)*, 133-137.

Walther, T.C., and Farese, R.V., Jr. (2009). The life of lipid droplets. Biochim Biophys Acta *1791*, 459-466.

Wang, B., Moya, N., Niessen, S., Hoover, H., Mihaylova, M.M., Shaw, R.J., Yates, J.R., 3rd, Fischer, W.H., Thomas, J.B., and Montminy, M. (2011). A hormone-dependent module regulating energy balance. Cell *145*, 596-606.

Warnick, C.T., and Lazarus, H.M. (1982). Effect of a protein-free diet and fasting on RNA polymerase activity in mice. J Nutr *112*, 293-298.

Wetterau, J.R., Combs, K.A., Spinner, S.N., and Joiner, B.J. (1990). Protein disulfide isomerase is a component of the microsomal triglyceride transfer protein complex. J Biol Chem *265*, 9800-9807.

Wilfling, F., Wang, H., Haas, J.T., Krahmer, N., Gould, T.J., Uchida, A., Cheng, J.X., Graham, M., Christiano, R., Frohlich, F., et al. (2013). Triacylglycerol synthesis enzymes mediate lipid droplet growth by relocalizing from the ER to lipid droplets. Dev Cell *24*, 384-399.

Winter, W.E., and Silverstein, J.H. (2000). Molecular and genetic bases for maturity onset diabetes of youth. Curr Opin Pediatr *12*, 388-393.

Wong, A.C., Vanhove, A.S., and Watnick, P.I. (2016). The interplay between intestinal bacteria and host metabolism in health and disease: lessons from Drosophila melanogaster. Dis Model Mech *9*, 271-281.

Wu, Q., Wen, T., Lee, G., Park, J.H., Cai, H.N., and Shen, P. (2003). Developmental control of foraging and social behavior by the Drosophila neuropeptide Y-like system. Neuron *39*, 147-161.

Wu, Z., Rosen, E.D., Brun, R., Hauser, S., Adelmant, G., Troy, A.E., McKeon, C., Darlington, G.J., and Spiegelman, B.M. (1999). Cross-regulation of C/EBP alpha and PPAR gamma controls the transcriptional pathway of adipogenesis and insulin sensitivity. Mol Cell *3*, 151-158.

Wyant, G.A., Abu-Remaileh, M., Wolfson, R.L., Chen, W.W., Freinkman, E., Danai, L.V., Vander Heiden, M.G., and Sabatini, D.M. (2017). mTORC1 Activator SLC38A9 Is Required to Efflux Essential Amino Acids from Lysosomes and Use Protein as a Nutrient. Cell *171*, 642-654 e612.

Wysocka, J., and Herr, W. (2003). The herpes simplex virus VP16-induced complex: the makings of a regulatory switch. Trends Biochem Sci *28*, 294-304.

Xu, H.E., Lambert, M.H., Montana, V.G., Parks, D.J., Blanchard, S.G., Brown, P.J., Sternbach, D.D., Lehmann, J.M., Wisely, G.B., Willson, T.M., et al. (1999). Molecular

recognition of fatty acids by peroxisome proliferator-activated receptors. Mol Cell *3*, 397-403.

Xu, X., Park, J.G., So, J.S., Hur, K.Y., and Lee, A.H. (2014). Transcriptional regulation of apolipoprotein A-IV by the transcription factor CREBH. J Lipid Res *55*, 850-859.

Xu, X., Park, J.G., So, J.S., and Lee, A.H. (2015). Transcriptional activation of Fsp27 by the liver-enriched transcription factor CREBH promotes lipid droplet growth and hepatic steatosis. Hepatology *61*, 857-869.

Yabal, M., Brambillasca, S., Soffientini, P., Pedrazzini, E., Borgese, N., and Makarow, M. (2003). Translocation of the C terminus of a tail-anchored protein across the endoplasmic reticulum membrane in yeast mutants defective in signal peptide-driven translocation. J Biol Chem *278*, 3489-3496.

Yamashita, H., Takenoshita, M., Sakurai, M., Bruick, R.K., Henzel, W.J., Shillinglaw, W., Arnot, D., and Uyeda, K. (2001). A glucose-responsive transcription factor that regulates carbohydrate metabolism in the liver. Proc Natl Acad Sci U S A *98*, 9116-9121.

Yao, Z., Zhou, H., Figeys, D., Wang, Y., and Sundaram, M. (2013). Microsome-associated lumenal lipid droplets in the regulation of lipoprotein secretion. Curr Opin Lipidol *24*, 160-170.

Yarmolinsky, D.A., Zuker, C.S., and Ryba, N.J. (2009). Common sense about taste: from mammals to insects. Cell *139*, 234-244.

Yoshida, H., Matsui, T., Yamamoto, A., Okada, T., and Mori, K. (2001). XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. Cell *107*, 881-891.

Yusuf, D., Butland, S.L., Swanson, M.I., Bolotin, E., Ticoll, A., Cheung, W.A., Zhang, X.Y., Dickman, C.T., Fulton, D.L., Lim, J.S., et al. (2012). The transcription factor encyclopedia. Genome Biol *13*, R24.

Zeituni, E.M., Wilson, M.H., Zheng, X., Iglesias, P.A., Sepanski, M.A., Siddiqi, M.A., Anderson, J.L., Zheng, Y., and Farber, S.A. (2016). Endoplasmic Reticulum Lipid Flux Influences Enterocyte Nuclear Morphology and Lipid-dependent Transcriptional Responses. J Biol Chem *291*, 23804-23816.

Zhang, C., Wang, G., Zheng, Z., Maddipati, K.R., Zhang, X., Dyson, G., Williams, P., Duncan, S.A., Kaufman, R.J., and Zhang, K. (2012). Endoplasmic reticulum-tethered transcription factor cAMP responsive element-binding protein, hepatocyte specific, regulates hepatic lipogenesis, fatty acid oxidation, and lipolysis upon metabolic stress in mice. Hepatology *55*, 1070-1082.

Zhang, H., Liu, J., Li, C.R., Momen, B., Kohanski, R.A., and Pick, L. (2009). Deletion of Drosophila insulin-like peptides causes growth defects and metabolic abnormalities. Proc Natl Acad Sci U S A *106*, 19617-19622.

Zhang, K., and Kaufman, R.J. (2008). From endoplasmic-reticulum stress to the inflammatory response. Nature *454*, 455-462.

Zhang, K., Shen, X., Wu, J., Sakaki, K., Saunders, T., Rutkowski, D.T., Back, S.H., and Kaufman, R.J. (2006). Endoplasmic reticulum stress activates cleavage of CREBH to induce a systemic inflammatory response. Cell *124*, 587-599.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zheng, Z., Kim, H., Qiu, Y., Chen, X., Mendez, R., Dandekar, A., Zhang, X., Zhang, C., Liu, A.C., Yin, L., et al. (2016). CREBH Couples Circadian Clock With Hepatic Lipid Metabolism. Diabetes *65*, 3369-3383.

Zhong, S., Magnolo, A.L., Sundaram, M., Zhou, H., Yao, E.F., Di Leo, E., Loria, P., Wang, S., Bamji-Mirza, M., Wang, L., et al. (2010). Nonsynonymous mutations within APOB in human familial hypobetalipoproteinemia: evidence for feedback inhibition of lipogenesis and postendoplasmic reticulum degradation of apolipoprotein B. J Biol Chem *285*, 6453-6464.

Zinke, I., Kirchner, C., Chao, L.C., Tetzlaff, M.T., and Pankratz, M.J. (1999). Suppression of food intake and growth by amino acids in Drosophila: the role of pumpless, a fat body expressed gene with homology to vertebrate glycine cleavage system. Development *126*, 5275-5284.

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Dean's Office Medical Faculty
Faculty of Medicine

Affidavit

Khan, Haris Ahmad

Surname, first name

Grosshadernerstr. 9, 82152, Planegg-Martinsried

Address

I hereby declare, that the submitted thesis entitled

Dissecting the nutrient-driven role of Creb3L transcription factor family to coordinate ER function

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree.

Munich, 10.11.2023

Place, Date

Haris Ahmad Khan

Signature doctoral candidate

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Dean's Office Medical Faculty
Doctoral Office

Confirmation of congruency between printed and electronic version of the doctoral thesis

Doctoral candidate: Haris Ahmad Khan

Address: Grosshadernerstr. 9, 82152, Planegg-Martinsried

I hereby declare that the electronic version of the submitted thesis, entitled

Dissecting the nutrient-driven role of Creb3L transcription factor family to coordinate ER function

is congruent with the printed version both in content and format.

Munich, 10.11.2023 — Haris Ahmad Khan

Place, Date — Signature doctoral candidate