Biomedizinisches Centrum München der Ludwig-Maximilians-Universität München

Dissertation
zum Erwerb des Doctor of Philosophy (Ph.D.)
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

Towards a mechanistic understanding of the activation of ChREBP

vorgelegt von:

Mehera Elizabeth Emrich

aus:

San Diego, California, USA

Jahr:

2020

First supervisor: *Herr Prof. Dr. Andreas G. Ladurner*

Second supervisor*:* *Herr Prof. Dr. Dirk Eick*

Dean: **Prof. Dr. med. dent. Reinhard Hickel**

Datum der Verteidigung:

18.12.2020

Affidavit

Emrich, Mehera Elizabeth

Surname, first name

Biomedical Center Munich, Department of Physiological Chemistry,
Ludwig-Maximilians-Universität, Großhaderner Str. 9, 82152 Planegg-Martinsried

Address

I hereby declare, that the submitted thesis entitled

Towards a mechanistic understanding of the activation of ChREBP

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 submitted thesis or parts thereof have not been presented as part of an examination degree to any other university.

Munich, 03.07.20

Place, Date

Mehera Elizabeth Emrich

Signature doctoral candidate

List of publications to date

A. R. Alasdair, B. Banos-Pinero, P. Paschke, L. Sanchez-Pulido, A. Antonio, D. Joseph, M. Emrich, D. Leys, C. P. Ponting, I. Ahel, N. D. Lakin, "The Role of ADP-ribosylation in Regulating DNA Interstrand Crosslink Repair," J Cell Sci , vol. 129, no. 20, pp. 3845-58, 2016

Abstract

Carbohydrate response element binding protein (ChREBP) is a central metabolic transcription factor found in both the vertebrate and invertebrate lineages. ChREBP is activated by increases in environmental sugar levels, and promotes the transcription of key genes involved in lipogenesis and satiation. This sugar sensitivity is known to be mediated by the actions of the highly conserved N-terminal glucose-sensing module (GSM) of the protein. The GSM is composed of an activatory glucose-response activation conserved element (GRACE) and a repressive low glucose inhibitory domain (LID). GSM activation requires a release of GRACE repression by the LID. The mechanism of this intramolecular repression and what drives its release has not yet been demonstrated. ChREBP has been proposed to be activated by a variety of metabolites and proteins, however the extent which any given factor is required for transcriptional activation remains unclear.

My host lab has discovered that the purified GSM binds directly to glucose-6-phosphate (G6P), a molecule which has previously been demonstrated to be necessary for ChREBP activation in pancreatic β-cells. This binding involves a conformational change involving three highly conserved regions of ChREBP, including its nuclear localisation sequence. I hypothesized that the LID and GRACE directly associate under low glucose conditions and that glucose metabolism and G6P binding drives a release of LID–GRACE contacts, allowing ChREBP activation. To test this hypothesis, I performed co-immunoprecipitation, fluorescence two-hybrid and fluorescence lifetime imaging experiments in cultured U2OS and 832/13 cells. These experiments demonstrated that the LID and GRACE associate directly and that the level of association shows an inverse relation to environmental sugar levels. Furthermore, my data show that this dissociation between the LID and GRACE in high media glucose levels corresponds with an opening of the GSM, implying relevance to the full-length protein. However, my results do not indicate that G6P levels play a direct role in this intramolecular interaction. Instead, I found that MCR6, a region of the GRACE domain not strictly required for G6P binding, appears to be necessary for LID–GRACE association. These results imply that G6P binding and GSM opening may be separate, but possibly coupled, steps in the activation of ChREBP.

Increased environmental sugar levels are known to drive increases in ChREBP nuclear localisation. In this work, I sought to better characterise the role environmental sugar plays in the subcellular localisation of ChREBP. To this end, I collaborated with Dr. Imke Mandemaker to perform fluorescence recovery after photobleaching and photoactivation assays. With these experiments, we attempted to quantify how media glucose levels affect the rates of ChREBP nuclear import and export as well as nucleoplasmic mobility in live 832/13 cells. Our results found that the steady-state nuclear fractions of ChREBP are rapidly exchanged, with over 50% of the steady-state nuclear fraction of GFP-ChREBP (i.e. over 10% of total cellular GFP-ChREBP) imported every minute. Surprisingly, high glucose cells had comparable rates of nuclear import, nuclear export and nucleoplasmic mobility to those in low glucose. These data stand in contrast to the published work and my own observations showing that ChREBP nuclear occupancy increases with increased environmental sugar, and imply that about 10% of total cellular ChREBP is being rapidly cycled between the nucleus and cytoplasm. It is possible that

the mechanism driving differential nuclear localisation of ChREBP is the result of a rapid acute response and thus not observable in the steady-state conditions assayed here, or that the kinetics of this mechanism are too slow to be observed in the sub two-minute time frame of these experiments.

As unpublished hydrogen/deuterium exchange mass spectrometry (HDX-MS) assays of our collaborators show that G6P binding drives solvent accessibility changes in three highly conserved regions of ChREBP, including the nuclear localisation sequence, it was of interest to elucidate the role that each of these regions plays in the glucose-responsive activation and nuclear localisation of ChREBP. To this end, I performed targeted mutagenesis of conserved residues within these regions and assayed the effect these mutations had on glucose-responsive activity and nuclear localisation using luciferase assays and confocal microscopy in 832/13 cells. One key mutant, previously found to be incapable of G6P binding, showed abrogated luciferase reporter activity but no inhibition of glucose responsive nuclear localisation, demonstrating that nuclear localisation is separable from protein activation. Additionally, my results demonstrated that the highly conserved basic patch of the nuclear localisation sequence which associates with importin in the crystal structure is dispensable for glucose-responsive nuclear localisation. Finally, I show that the universally conserved core of MCR6, the GRACE region required for LID–GRACE association, plays a repressive role in ChREBP activation.

Taken together my data imply a two-step model of ChREBP activation. Rapid steady-state cycling of ChREBP between the cytoplasm and the nucleus allows for immediate responses to changes in environmental sugar levels. Upon increased media glucose, G6P binding and altered interactions with an additional glucose-responsive factor or factors are together required for protein activity.

Zusammenfassung

Carbohydrate response element binding protein (ChREBP) ist ein zentraler metabolischer Transkriptionsfaktor, der bei Wirbeltieren und Tieren wirbelloser Abstammung vorkommt. ChREBP wird bei Zunahme des Nahrungszuckers in der Umwelt aktiviert und fördert die Transkription von wichtigen Genen, die an der Lipogenese und Sättigung beteiligt sind. Diese Zuckerempfindlichkeit wird durch das amino-terminale glucose-sensing module (GSM) des Proteins vermittelt. Das GSM besteht aus einem aktivierenden glucose-response activation conserved element (GRACE) und aus einer repressiven glucose inhibitory domain (LID). Eine Aufhebung der Repression des GRACE von der LID wird für die GSM Aktivierung benötigt. Es ist noch nicht bekannt wie der Mechanismus hinter dieser intramolekularen Repression funktioniert, oder welche Faktoren zu ihrer Aufhebung führen. Eine Vielzahl von Stoffwechselprodukten und Proteinen sind in der Aktivierung von ChREBP impliziert. In welchem Maße ein gegebener Faktor für die transkriptionelle Aktivität von ChREBP benötigt wird, ist derzeit noch unklar.

In anderen Versuchen meines Gast Labors wurde entdeckt, dass das gereinigte GSM direkt Glukose-6-phosphat (G6P) bindet. G6P ist ein Stoffwechselprodukt das, wie zuvor gezeigt, erforderlich für die Aktivierung des ChREBP-Proteins in pankreatischen beta-Zellen ist. Diese Bindung

schließt eine Konformationsänderung in drei konservierten Bereichen des ChREBP-Proteins ein, beziehungsweise die Kernlokalisationssequenz des Proteins. Daraus ergibt sich die Hypothese, dass LID und GRACE in direkter Verbindung zueinanderstehen, wenn die Glukosekonzentration des Wachstumsmediums niedrig ist. Zusätzlich verursachen der Glukosestoffwechsel und die Bindung von G6P eine Freilassung der LID-GRACE Wechselwirkung, wenn die Glukosekonzentration erhöht ist, was die Aktivierung von ChREBP veranlasst. Um diese Hypothese zu testen, habe ich Co-Immunpräzipitation, Fluoreszenz zwei-Hybrid, und Fluoreszenzlebensdauer-Experimente in angezüchteten U20S und 832/13 Zellen durchgeführt. Diese Experimente haben gezeigt, dass zwischen LID und GRACE eine direkte Interaktion besteht und weiterhin, dass die Höhe der Interaktion umgekehrt proportional zur Glukosekonzentration des Wachstumsmediums ist. Meine Daten zeigen, dass diese Dissoziation zwischen LID und GRACE aufgrund der Erhöhung der Glukosekonzentration mit einer Öffnung des ganzen GSMs übereinstimmt, was eine funktionale Relevanz für das volle Protein andeutet. Dennoch deuten meine Daten nicht darauf hin, dass die Konzentration des G6P eine direkte Rolle bei dieser Wechselwirkung spielt. Stattdessen scheint ein konserviertes Motiv der GRACE-Domäne namens MCR6, das nicht für die Bindung des G6Ps nötig ist, für die LID-GRACE Verbindung erforderlich zu sein. Diese Resultate deuten an, dass die Bindung von G6P und die Eröffnung des GSM voneinander getrennte und möglicherweise gekoppelte Schritte bei der Aktivierung des ChREBP-Proteins darstellen.

Es ist bekannt, dass Erhöhungen des Glukosespiegels stärkere Kernlokalisation von ChREBP verursachen. Ich habe versucht, die Rolle der Glukosekonzentration bei der Lokalisation des ChREBP-Proteins zu charakterisieren. Zu diesem Zweck habe ich mit Dr. Imke Mandemaker zusammengearbeitet, um Fluoreszenz-recovery Assays bei Bleichung und Lichtaktivierung durchzuführen. Mit diesen Assays bestimmten wir den Effekt der Glukosekonzentration des Wachstumsmediums auf die Kernlokalisations- und Kernexportraten sowie die nukleoplasmische Mobilität des ChREBP-Proteins in lebenden Zellen. Unsere Resultate zeigen, dass kernlokalisiertes ChREBP im Gleichgewichtszustand schnell ausgetauscht wird. Genauer gesagt, wird pro Minute mehr als 50% des kernlokalisierten ChREBPs bzw. mehr als 10% des gesamten ChREBPs in die Zelle importiert. Überraschenderweise haben Zellen im Wachstumsmedium mit erhöhter Glukosekonzentration und Zellen im Wachstumsmedium mit niedriger Glukosekonzentration vergleichbare Kernlokalisations- und Kernexportraten sowie nukleoplasmische Mobilität. Diese Daten stehen im Gegensatz zu veröffentlichten Ergebnissen und zu meinen eigenen Resultaten. Diese zeigen, dass die Kernlokalisation von ChREBP mit einem glukosekonzentrationserhöhten Wachstumsmedium stimuliert wird. Außerdem deuten diese Resultate an, dass ungefähr 10% des gesamten ChREBP in der Zelle in einem schnellen Zyklus zwischen Kern und Zytoplasma abläuft. Es ist möglich, dass der Mechanismus, welcher die differenziale Kernlokalisation von ChREBP hervorbringt, selbst durch eine schnelle akute Reaktion verursacht wird und deshalb nicht bei dem Gleichgewichtszustand, der hier eingesetzt wurde, beobachtbar war, oder, dass die Kinetik dieses Mechanismus zu langsam ist, um in der Zeitskala dieser Experimente sichtbar zu sein.

Unpublizierte Wasserstoff-Deuterium-Austausch-Massenspektrometrie (HDX-MS) Assays eines Kollaborationspartners zeigen, dass die Bindung von G6P Veränderungen der Zugänglichkeit des

Lösungsmittels in drei konservierten Bereichen des ChREBP-Proteins einschließlich der Kernlokalisationsequenz des Proteins verursacht. Daher beabsichtigten wir, die Rolle, welche diese konservierten Bereiche bei der glukosereagierenden Aktivierung und Kernlokalisierung ChREBP-Proteins spielen, aufzudecken. Zu diesem Zweck habe ich ortsspezifische Mutagenese konservierter Aminosäurereste durchgeführt und die Wirkung dieser Veränderungen in der glukosereagierenden Aktivität und der Kernlokalisierung durch Luziferase-Assays und Konfokal Mikroskopie in 823/13 Zellen untersucht. Eine Hauptmutante, die nicht in der Lage ist eine G6P-Bindung einzugehen, wies eine aufgehobene Luziferase-Reporteraktivität aber keine Inhibition der glukosereagierenden Kernlokalisierung des Proteins auf. Darüber hinaus beweisen meine Resultate, dass das konservierte basische Motiv der Kernlokalisationssequenz des ChREBP-Proteins, das sich mit Importin in der Kristallstruktur assoziiert, unnötig für die glukosereagierende Kernlokalisierung ist. Schließlich zeige ich, dass das konservierte Kernmotiv des MCR6, die Region der GRACE-Domäne, welche für die LID-GRACE Interaktion erforderlich ist, eine repressive Rolle bei der ChREBP-Aktivierung spielt.

Zusammengenommen deuten meine Daten auf ein Zweistufenmodell der Aktivierung von ChREBP hin. Schnelle steady-state Zirkulation von ChREBP zwischen Kern und Zytoplasma ermöglicht sofortige Reaktionen auf Veränderungen in der Glukosekonzentration des Wachstumsmediums. Bei Erhöhung des Glukosespiegels sind G6P-Bindung und veränderte Interaktionen mit einem weiteren oder mehreren glukosereagierenden Faktor bzw. Faktoren für die Proteinaktivität erforderlich.

Acknowledgements

I would like to thank my supervisor Prof. Dr. Andreas Ladurner. Thank you for fostering scientific exploration and leading a department in which constructive feedback and new perspectives are always available. Thank you as well for your supervision and invaluable feedback throughout this project.

Thank you to the Mondo team, Dr. Carla Margulies, Hui-Lan Huang, Thomas Pysik, and William Menzer for the productive discussions and for asking the right questions. Thank you to Master student Giorgi Beroshvili whose microscopy work was invaluable. Thank you to Amgen Scholar Pedro Garcia for your microscopy and cloning assistance. Thank you to my TAC members Prof. Dr. Dirk Eick, Prof. Dr. Heinrich Leonhardt, and Prof. Dr. Eckhard Wolf for their support and guidance throughout this project.

Thank you to all the members of the Ladurner and Margulies labs. Without the support of all of you I would not have made it to this point. I would especially like to thank Dr. Magdalena Murawska and Dr. Imke Mandemaker for all of their big picture feedback. I would additionally like to thank Imke for her microscopy work included in here. Together we can figure out what this protein is doing. Thank you to Julia Preisser and Christine Kotthoff for all of the assistance along the way (Go team). I would also like to thank Dr. Giuliana Möller for helping me get started in the lab and for making sure the beginning of my time here went as smoothly as possible.

Thank you to Professor Christopher Newgard and Lisa Poppe from Duke University for providing the 832/13 cells without which this project would not be possible. Thank you to Professor Lawrence Chan and Professor Benny Hung-Junn Chang at the Baylor College of Medicine for providing the vectors used in my luciferase assays.

Thank you to Dr. Steffen Dietzel and Dr. Andreas Thomae of the BMC bioimaging core facility for getting me started with fluorescent lifetime imaging. I would especially like to thank Mariano Gonzalez Pisfil for your microscope training and for your expert advice on how a lifetime experiment should be run. Thank you to Dr. Tobias Straub of the BMC bioinformatics core facility. Your assistance and advice made much of my image analysis possible. Thank you Dr. Lisa Richter from the BMC flow cytometry core facility for the efficient monoclonalisation.

Finally, I would like to thank the members of the department who I have not named. Thank you for the countless helpful discussions and thought-provoking feedback.

Contents

List of Figures

List of Tables

Abbreviations

2DG 2-deoxyglucose
AMPK AMP-activated protein kinase
BAT brown adipose tissue
bHLH basic helix-loop-helix
ChOREs Carbohydrate Response Elements
ChREBP carbohydrate response element binding protein
CoIP Co-immunoprecipitation
CRM1 chromosomal maintenance 1
DBD DNA binding domain
DNL *de novo* lipogenesis
F26BP fructose-2,6-bisphosphate
F2H fluorescence two-hybrid
F6P fructose-6-phosphate
FBPase fructose 1,6-bisphosphatase
FGF21 fibroblast growth factor 21
Fluc firefly luciferase
FRAP fluorescence recovery after photobleaching
FRET-FLIM fluorescence resonance energy transfer-fluorescence lifetime imaging
G6P glucose-6-phosphate
GLP1 Glucagon-like peptide 1
GRACE glucose-response activation conserved element
GSM glucose-sensing module
HDX-MS Hydrogen/deuterium exchange mass spectrometry
HNF4α hepatic nuclear factor 4α
HRP horseradish peroxidase
ITC isothermal titration calorimetry
LB lysogeny broth
LID low glucose inhibitory domain
LMB leptomycin B
LXR liver X receptor
MCR Mondo conserved region
MLX Max-like protein X
MST microscale thermophoresis
NAFLD nonalcoholic fatty liver disease
NCOR nuclear receptor co-repressor
NES nuclear export sequence
NLS nuclear localisation sequence
PBS phosphate-buffered saline
PDB protein data bank
PFK Phosphofructokinase
PFKFB 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase
PKA cAMP dependent protein kinase
PMSF phenylmethylsulfonyl fluoride
PMT photomultiplier tube
PNK Polynukleotide Kinase
PP2A Protein Phosphatase 2A
PPARα Peroxisome proliferator-activated receptor α
PUFAs polyunsaturated fatty acids
RLU relative light units
Rluc renilla luciferase
RXR retinoid X receptor
SDS-PAGE sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SNP single nucleotide polymorphism
SREBP-1c sterol regulatory element-binding protein 1c
T2D Type 2 diabetes
TBS tris-buffered saline
TBST tris-buffered saline 0.2% Tween-20

TCSPC time-correlated single-photon counting **X5P** xylulose 5-phosphate
WAT white adipose tissue

Chapter 1

Introduction

1.1 Carbohydrate metabolism is central to metabolic health

1.1.1 Metabolic homeostasis is maintained by the concerted actions of key tissues

Carbohydrate metabolism is central to metabolic health. Upon carbohydrate consumption in mammals, circulating blood glucose levels increase, driving insulin release by the β-cells of the pancreas. Circulating insulin promotes blood glucose uptake by organs and tissues such as the liver and adipose tissue. The liver takes up circulating glucose and stores it as glycogen via glycogenesis or fat via *de novo* lipogenesis (DNL). Adipose tissue also takes in excess blood glucose and converts it to fats by DNL (Figure 1.1a). In this way, a spike in blood sugar is prevented in the fed state and energy reserves are created for times of energy deficit. In the early fasted state, low blood glucose levels lead to the secretion of glucagon by pancreatic α-cells. Glucagon promotes glycogenolysis by the liver, the glucose product of which is released into the blood, preventing a critical drop in blood sugar. Once liver glycogen is expended during sustained periods of energy deficit, the liver uses gluconeogenesis to allow continued glucose production and release. Similarly, a drop in circulating insulin releases inhibition of adipose tissue lipolysis. The subsequent breakdown of triacylglycerides generates glycerol and fatty acids. The former acts as a gluconeogenic precursor in liver, while the latter are used as an energy source in muscle (Figure 1.1b). During a prolonged fast, gluconeogenic substrates become scarce and the liver switches to ketone body production, allowing the continued function of glucose-dependent organs such as the brain. This cyclic process of hormonally triggered metabolite uptake and release allows for the maintenance of energetic homeostasis under changeable metabolic conditions [1, 2].

1.1.2 Nutrient-sensitive transcription factors coordinate the role of the liver in metabolic homeostasis

Metabolically regulated transcription factors coordinate the role of the liver in blood sugar homeostasis. Key among these are carbohydrate response element binding protein (ChREBP), liver X receptor (LXR), sterol regulatory element-binding protein 1c (SREBP-1c) and Peroxisome proliferator-activated receptor α (PPARα). The former three proteins are activated in the fed state

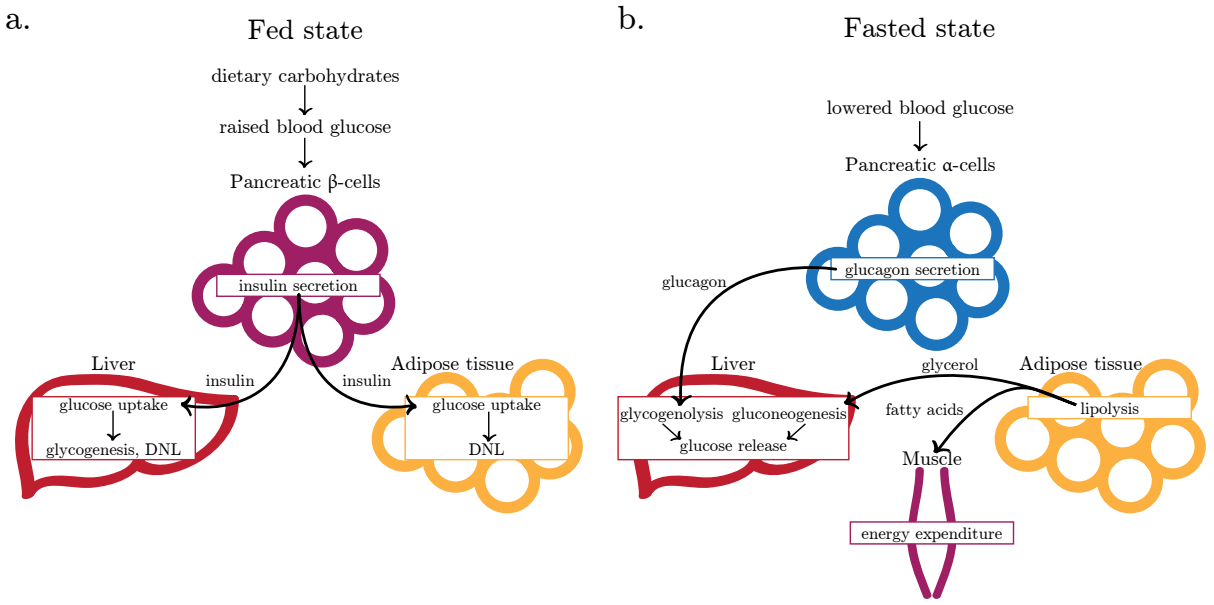


Figure 1.1: **Metabolic homeostasis is maintained by the concerted actions of key tissues**
a. Blood glucose levels rise after a carbohydrate containing meal. Increased blood glucose concentration drives insulin release from the β-cells of the pancreas. Circulating insulin promotes blood glucose uptake by the liver and adipose tissue. In the liver, imported glucose is processed via glycogenesis and DNL, allowing storage as glycogen and fat, respectively. In the adipose tissue, imported glucose is processed by DNL and stored as fat. **b.** In the fasted state, decreases in blood glucose levels stimulate the release of glucagon from the pancreatic α-cells. Glucagon promotes glycogenolysis in the liver, allowing the release of hepatic glucose into the bloodstream. Hepatic gluconeogenesis enables continuous glucose release following the exhaustion of glycogen stores. Decreases in circulating insulin levels enable adipose tissue lipolysis, causing the release of glycerol and fatty acids. Glycerol acts as a gluconeogenic precursor in the liver, while fatty acids are used as an energy source by the muscle in the fasted state.

and drive energy storage by promoting DNL. ChREBP and SREBP-1c are activated by sugar signalling, ChREBP via cytoplasmic glucose metabolism, and SREBP-1c through long-range insulin signalling [3, 4]. LXR is activated by direct binding to oxysterol cholesterol derivatives and, in addition to its activation of DNL, prevents cholesterol accumulation in the liver through the inhibition of cholesterol biosynthesis and activation of cholesterol conversion to bile acids [5, 6]. PPARα, unlike ChREBP, LXR and SREBP-1c, inhibits DNL and promotes β-oxidation and ketogenesis. It is activated by binding to fatty acids, either released by lipolysis in the low energy state or as the result of a high fat diet [7].

These factors functionally interact with each other and other metabolic regulators to coordinate the hepatic response to nutrient availability. LXR activates the transcription of both ChREBP and SREBP-1c [8, 9, 10, 11, 12]. ChREBP in turn promotes transcription of its constitutively active isoform, ChREBPβ [13, 11]. Together these proteins initiate a global transition to lipogenic metabolism in response to nutrient surfeit. In response to a high fat diet or the fasted state, PPARα counteracts this lipogenic programme, inhibiting transcription of ChREBP target genes [14, 15] and preventing LXR-driven activation of the SREBP-1c promoter [16].

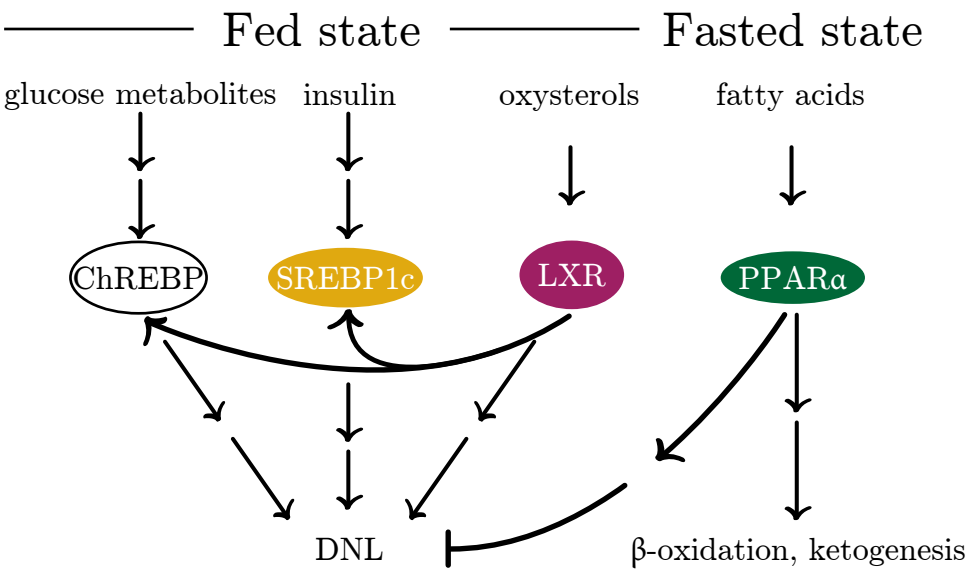


Figure 1.2: **Nutrient-sensitive transcription factors coordinate the role of the liver in metabolic homeostasis**
ChREBP, SREBPP-1c and LXR drive hepatic DNL in the fed state. The three transcription factors are activated by various metabolic signals, ChREBP by glucose metabolites, SREBP-1c by insulin signalling and LXR by oxysterol cholesterol derivatives. Activated LXR promotes transcription of ChREBP and SREBPP-1c, strengthening the anabolic response of the liver to the fed state. In the fasted state, fatty acids released by lipolysis agonise PPARα, which promotes β-oxidation and ketogenesis.

LXR and PPARα are nuclear receptor proteins which interact with retinoid X receptor (RXR) to regulate the transcription of target genes. Their activity is regulated not only by the presence of ligand, but by interactions with other proteins such as the nuclear receptor co-repressor (NCOR) complex, which can interact with the apo forms of both LXR and PPARα and block their activity [17, 18]. This interaction is regulated by both nuclear receptor ligand (and therefore nutrient) availability and the modification of interacting protein functionality by external nutrient signals. For example, although NCOR, a subunit of the NCOR complex, is capable of inhibiting both PPARα and LXR in the basal state, insulin signalling drives its phosphorylation and release of LXR. This regulation of activity and specificity is additionally moderated by direct associations between nuclear receptor proteins. LXR and PPARα directly interact in a manner regulated by the availability of their respective ligands. LXR-PPARα association and agonist levels modulate the ability of these proteins to bind and activate their respective DNA response elements [19, 20, 21, 22]. In this way, metabolic cues regulate nuclear receptor functions at various levels, allowing both for fine-tuning and redundancy in the response to nutrient availability.

In light of this complexity of regulatory signalling, it is not surprising that the interactions between the lipogenic transcription factors ChREBP, LXR and SREBP-1c and PPARα involve more than straightforward inhibition of the former proteins by the latter. PPARα has been shown to promote the transcription of SREBP-1c [23] and, more generally, PPARα and lipogenic transcription factors often share target genes. For example, both ChREBP and PPARα promote the transcription of the microsomal triglyceride transfer protein, a key catalyst in the synthesis of the lipoprotein particles required for hepatic triglyceride export [24, 25]. More strikingly, ChREBP

and PPARα together are required for the glucose-stimulated transcription of the hepatokine fibroblast growth factor 21 (FGF21). This hormone has varied functions including the promotion of glucose uptake in peripheral tissues, the downregulation of sucrose feeding and the inhibition of DNL [26, 27].

This interconnectedness is in part mediated by additional regulatory factors such as nuclear receptor protein hepatic nuclear factor 4α (HNF4α). Unlike other nuclear receptors, HNF4α activity does not appear to be affected by its proposed ligand [28]. HNF4α is required for both the fasting-induced expression of PPARα and the glucose-induced expression of ChREBP [29, 11] and plays a complex role in the regulation of PPARα, whose transactivation activity it appears to promote or inhibit in a target-dependent manner [29, 30]. In contrast, the reported interactions between HNF4α and ChREBP are purely activatory. HNF4α binds ChREBP at the promoter of ChREBP target genes and promotes their activation [31, 11]. This direct ChREBP-nuclear receptor interaction is not exclusive to HNF4α. In fact, there is evidence that competition between ChREBP and LXR agonists regulates LXR specificity [32].

LXR and ChREBP physically interact to activate ChREBP target genes even in the absence of an LXR response element. This interaction and activation of ChREBP-specific target genes is inhibited by LXR agonists, demonstrating an agonist binding driven switch of LXR function from co-activation of ChREBP to activation of LXR response-element containing genes [32]. Whether this represents a unique interaction or a more general mechanism of nuclear receptor regulation by ChREBP remains unknown.

1.1.3 PKA and AMPK are central metabolic regulators with tissue-specific activity

Nutrient responsive transcription factors are, of necessity, regulated by the nutrient availability of the cells they occupy. In a practical sense, the nutrient availability in these cells is dependent on the rates of both nutrient import and nutrient metabolism. This connection between import, metabolism and regulation can be seen in the intracellular processing of glucose. Glucose is imported into the cell by selective transporters. Once inside the cell it is phosphorylated to glucose-6-phosphate (G6P) by hexokinase and can then be converted into glycogen, enter the pentose phosphate pathway, or be processed by glycolysis. The products of glycolysis can function as the substrates for DNL or ketogenesis, or generate biosynthetic building blocks and ATP by way of the tricarboxylic acid cycle and the electron transport chain. These pathways are regulated by the energy state of the organism which, in addition to direct signalling by individual metabolites, exemplified by the nuclear receptors described above, is largely signalled by nutrient-sensitive post-translational modifications. Key among these metabolic regulatory mechanisms are the post-translational modifications catalysed by cAMP dependent protein kinase (PKA) and AMP-activated protein kinase (AMPK). The balance of these is controlled by the intracellular metabolic state which can be signalled by distinct molecules in different organs and tissues.

As described previously, the role of the liver in glucose metabolism is, broadly, to maintain

blood glucose homeostasis by releasing glucose or ketone bodies under low energy conditions and taking up circulating glucose in situations of excess. This homeostatic function is reflected in the metabolic response of the liver to low energy states. In the fasted state, hepatic cAMP and PKA orchestrate a global metabolic shift. Glucagon signalling drives an increase in hepatic cAMP [33], which activates PKA to promote glucose release and inhibit energy storing mechanisms. This effect is achieved by a variety of mechanisms including the phosphorylation of LXR. Phosphorylated LXR cannot heterodimerise with RXR, and is therefore prevented from activating target lipogenic genes such as SREBP-1c. Additionally, LXR phosphorylation recruits corepressor NCOR, further subduing lipogenic protein activity [34]. PKA mediated inhibition of hepatic energy storage is complemented by activation of glycogenolysis and gluconeogenesis, driving a complete metabolic shift [35]. In this way, PKA initialises a shift in hepatic metabolism during the low energy state which leads to glucose release and maintenance of stable blood sugar.

One of the central downstream regulatory factors acted on by PKA is the bifunctional enzyme 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase (PFKFB). The kinase activity of PFKFB catalyses the phosphorylation of fructose-6-phosphate (F6P) to fructose-2,6-bisphosphate (F26BP). F26BP then activates the glycolytic conversion of its precursor F6P to fructose-1,6-bisphosphate (F16BP) while inhibiting the reverse reaction via regulation of the glycolytic enzyme Phosphofructokinase (PFK) and the gluconeogenic enzyme fructose 1,6-bisphosphatase (FBPase), respectively. Conversely, while the kinase activity of PFKFB catalyses the synthesis of F26BP, the phosphatase activity of PFKFB catalyses the dephosphorylation of the hexose back to F6P, thereby removing an important activator of glycolytic flux [36]. In liver, PFKFB is phosphorylated under low energy conditions by PKA. This phosphorylation inhibits the kinase and activates the phosphatase of the bifunctional enzyme, leading to degradation of F26BP back into F6P. The decrease in F26BP levels results in the stimulation of gluconeogenic flux and an increase in cellular hexose phosphates [37]. After a meal, PKA activation decreases and Protein Phosphatase 2A (PP2A) dephosphorylates PFKFB, which is once again able to stimulate glycolytic flux by way of F26BP. This regulation of F26BP levels, not only plays a central role in the control of glycolytic flux, it is also indicated to regulate the activity of the sugar-responsive transcription factor ChREBP [38].

While PKA is central to the response of the liver to the low energy state, AMPK coordinates hepatic metabolism during the starved state. Increases in cytoplasmic AMP, which occur in conditions of extreme energy deficit, activate AMPK, which binds AMP with a low µM affinity; activated AMPK in turn inhibits DNL and gluconeogenesis and promotes beta oxidation and ketogenesis [39, 40, 41, 42, 43, 44]. In this way, hepatic AMPK promotes the synthesis of an alternative energy store when standard carbohydrates are not available and facilitates the maintenance of circulating metabolizable energy sources (Figure 1.3a).

Energy balance signalling in pancreatic β-cells is distinct from that of the liver, as these cells act to maintain glucose homeostasis by means of insulin secretion, rather than by direct glucose uptake and release. While PKA and AMPK are responsible for aspects of the metabolic response to low energy in liver, they favour opposing pathways in β-cells, where PKA is a signal of the

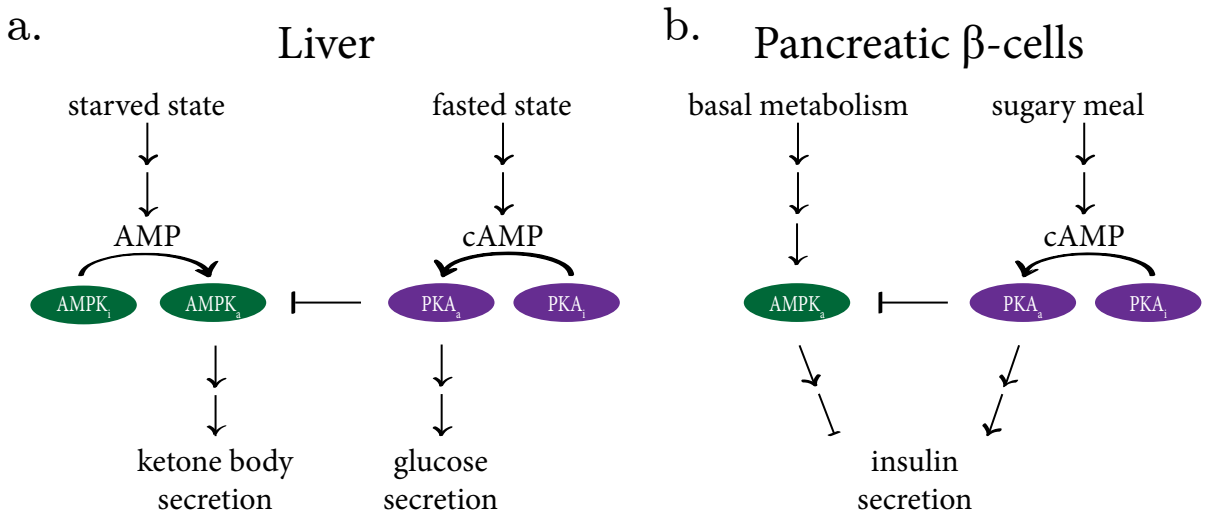


Figure 1.3: **PKA and AMPK are central metabolic regulators with tissue-specific activity**
a. In the liver, AMPK and PKA are distinct markers of low energy states. Upon activation, AMPK directs metabolism towards ketogenesis. When PKA is activated, it promotes glucose secretion. **b.** In pancreatic β-cells, AMPK is active in the basal metabolic state and inhibits insulin secretion. After a sugar-containing meal, PKA is activated and promotes insulin secretion.

fed state and AMPK regulates metabolic flux in the basal and starved states. After a meal, Glucagon-like peptide 1 (GLP1) signalling leads to an increase in cAMP levels in β-cells. The resulting activation of PKA, combined with increased postprandial glucose, promotes an influx of cytosolic Ca^{2+} which, along with cAMP itself, stimulates insulin secretion [45, 35]. The secreted insulin promotes glucose uptake and anabolic metabolism throughout the body, thereby maintaining stable blood glucose levels in the fed state [46]. As in liver, active PKA directly phosphorylates and inactivates AMPK, ensuring that both enzymes are not simultaneously active [47]. Once blood glucose returns to basal levels after a meal, glucose-driven repression of β-cell AMPK is released. AMPK then promotes the inhibition of insulin release, preventing overactive glucose uptake and hypoglycaemia [48, 49]. In this way, PKA and AMPK act in concert in β-cells to maintain glucose homeostasis by regulation of insulin secretion (Figure 1.3b). Thus, nutrient responsive post-translational modifications, together with metabolite-responsive transcription factors, act on a cellular level to maintain whole body energetic homeostasis.

1.1.4 Dysregulation or errors of metabolic homeostasis can be debilitating

Dysregulation or errors of metabolic homeostasis can be debilitating. Chronic metabolic imbalance can drive the development of traits such as obesity and insulin resistance which are associated with increased risk of cardiovascular disease and diabetes [50, 51]. These conditions, as members of the top ten leading causes of death worldwide, are two of the most devastating noncommunicable diseases facing the world today [52, 53]. Cardiovascular disease comprises disorders of the circulatory system while diabetes is a disease of ineffective or absent insulin signalling [54]. The conditions are closely linked. Type 2 diabetes (T2D), the class of diabetes characterised by

systemic insulin resistance, is a major risk factor for cardiovascular disease, the leading cause of death for adults with diabetes [55]. T2D and diabetes-associated cardiovascular disease correspond with a loss of glycaemic control and hyperglycaemia. In fact, blood sugar level is directly predictive of cardiovascular disease outcome. Despite this, deliberate glycaemic lowering of patients with T2D has surprisingly small effects on patient mortality and the extent to which blood sugar level is causal as opposed to symptomatic of disease severity is under debate [56, 57, 58]. A better understanding of the mechanisms of glucose metabolic regulation is therefore a key step on the path to more effective treatments or cures for those with diabetes and cardiovascular disease. For this reason, it is of interest to understand the role that glucose plays in intracellular metabolic networks. Key to this understanding is ChREBP, a member of the Mondo family of glucose-responsive transcription factors.

1.2 Mondo proteins are central regulators of carbohydrate metabolism

Mondo proteins are central regulators of carbohydrate metabolism. These proteins represent a highly conserved family of glucose-responsive transcription factors found throughout vertebrates and in some invertebrate species. These transcription factors contain an N-terminal sugar metabolite responsive region termed a glucose-sensing module (GSM) and a C-terminal basic helix-loop-helix (bHLH) DNA binding domain (DBD) (Figure 1.4a). Mondo proteins form an obligate heterodimer with Max-like protein X (MLX) and bind to target promoters in response to glucose signalling [59].

Mondo is found across various invertebrate lineages including the model organisms *Drosophila melanogaster* and *Caenorhabditis elegans* and is critical for sugar processing in flies. Mondo knockdown flies have elevated glycogen as well as increased circulating glucose and trehalose levels. These animals fail to develop when fed glucose, fructose, or sucrose (Hui Lan Huang, unpublished data), [62]. In these animals, the Mondo-MLX heterodimer is required for the upregulation of lipogenesis, glycolysis and the pentose phosphate pathway upon feeding and is seen as directly responsible for diet-induced synthesis of fats, reducing agents and biosynthetic precursors. Finally, many of the downstream targets of Mondo are required for sugar tolerance in flies, underscoring the importance of this protein as a core metabolic transcription factor [63].

The Mondo ortholog in *C. elegans* has been linked to metabolic signalling by way of the Insulin-like receptor. Loss of the Insulin-like receptor results in increased organismal lifespan in a manner dependent on the presence of *mxl-2*, the obligate Mondo interactor in *C. elegans* [64, 65]. Similarly, animals without the Insulin-like receptor have higher Mondo mRNA levels and increased Mondo nuclear localisation [65]. Work in my department on adult Mondo knockout flies has similarly indicated a diet-associated role of the protein in longevity (Hui Lan Huang, Man Leng Wong, Rory Beresford, unpublished data). The importance of the Mondo transcription factor is underscored by its conservation across invertebrates, and between vertebrates and invertebrates.

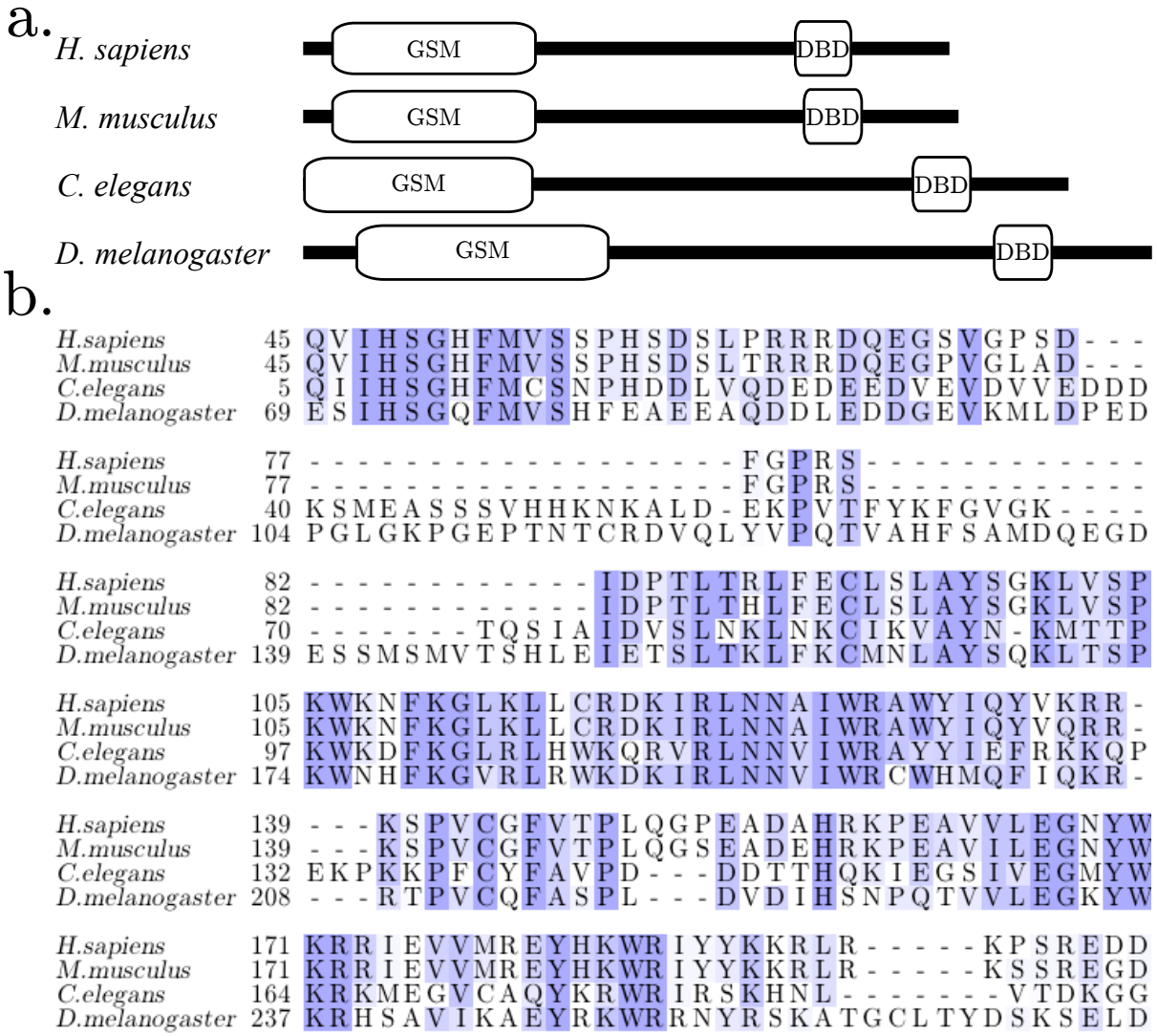


Figure 1.4: **ChREBP/Mondo is conserved between vertebrates and invertebrates** **a.** Mondo/ChREBP proteins contain an N-terminal GSM and a C-terminal bHLH DBD. Schematics of the human, mouse, *C. elegans* and *D. melanogaster* proteins are shown. **b.** An alignment of the N-terminus of the human, mouse, *C. elegans* and *D. melanogaster* Mondo/ChREBP GSM. Residue backgrounds are filled according to their BLOSUM 62 score. The alignment was performed using MUSCLE [60] with default settings in Jalview 2.8.2 [61]. Sequences used correspond to Uniprot ID Q9NP71-1, Q99MZ3-1, P41846-1 and Q9VID4-1 for the human, mouse, *C. elegans* and *D. melanogaster* proteins, respectively.

Figure 1.4b, which depicts an alignment of the N-terminus of the human, mouse, *C. elegans* and *D. melanogaster* GSM, highlights this conservation.

While invertebrates have one Mondo gene, vertebrates have two, *mlxip* and *mlxipl*, encoding the proteins MondoA and ChREBP, respectively. These paralogues differ in their tissue-specificity. MondoA is most highly expressed in muscle while ChREBP is mainly expressed in liver, small intestine, adipose tissue and pancreatic β-cells [66, 59]. Interestingly, ChREBP and invertebrate isoforms of Mondo (but not MondoA) have a conserved LxxLL nuclear receptor box binding

motif (the sequence which mediates canonical nuclear receptor-coactivator interactions [67]) which may allow it to act as a co-activator for nuclear receptor proteins such as HNF4α [31]. Despite differences at the primary sequence level and variations in tissue expression patterns, both MondoA and ChREBP show glucose-responsive transcriptional activation of target genes [68].

The glucose sensitivity of ChREBP is mediated by the regulatory interplay between the two main isoforms of the protein. ChREBPα shows glucose-regulated activity and nuclear localisation [69, 70], while ChREBPβ shows glucose-regulated *expression* and is constitutively active and nuclear localised [13]. Increased glucose levels drive the activation of ChREBPα which then activates the transcription of ChREBPβ [13]. Work in pancreatic β-cells shows that ChREBPβ downregulates ChREBPα levels, creating a negative feedback loop [71]. This combination of feed-forward and feedback regulation forms the basis of ChREBP activity levels *in vivo* and dictates the response of the protein to glycaemic changes.

ChREBP is required for successful sugar metabolism. ChREBP knockout mice die of hypothermia on a high sucrose or high fructose diet and have elevated plasma glucose levels [72]. The importance of ChREBP in sugar metabolism has been indicated in humans as well. The rs3812316 single nucleotide polymorphism (SNP) produces a Q241H variant of ChREBP that is linked to low triglyceride levels [73, 74]. An association study across multiple populations in Asia and Africa found an increased prevalence of the minor H allele in central Asian populations and it was hypothesised that this allele may have a glucose sparing function beneficial in populations whose traditional diets are extremely low in carbohydrates [74].

In addition to triglyceride levels, ChREBP function has been directly associated with glucose tolerance and insulin sensitivity in humans. Adipose ChREBP levels correlate positively with insulin sensitivity in nondiabetic adults, independent of body mass index [13]. Similarly, in adolescents with diabetes, ChREBP expression correlates positively with glucose tolerance [75]. Additionally supporting the importance of ChREBP in successful sugar metabolism is the phenotype of Williams-Beuren syndrome, a disorder caused by a chromosomal deletion of up to 17 genes including *mlxipl*. Seventy-five percent of people with this disorder have reduced glucose tolerance [76], possibly because of a constitutive ChREBP deletion.

1.3 ChREBP promotes whole-body energy homeostasis

ChREBP plays complementary roles in distinct tissues, the cumulative effect of which can be seen as the promotion of whole-body energy homeostasis. In mouse liver, ChREBP is a central regulator of both lipogenesis and glycolysis and is required for the expression of glucose regulated genes [77]. This glucose-responsive activity is promoted by LXR [78]. Hepatic ChREBP is critical

in the maintenance of blood glucose homeostasis and the introduction of constitutively-active hepatic ChREBP to full body knockout mice is sufficient to restore normoglycaemia [26]. The hepatic role of ChREBP in lipogenesis is largely mediated by its activation of SREBP-1c transcription [79]. When overexpressed in mouse liver, ChREBP causes nonalcoholic fatty liver disease (NAFLD). However, these hepatic ChREBP overexpressing mice actually experience improved insulin tolerance when on a high fat diet, indicating that, in the liver, ChREBP both promotes NAFLD and protects against insulin intolerance [80]. The converse has also been demonstrated. Liver-specific ChREBP knockout mice show resistance to hepatic steatosis caused by a high carbohydrate diet but develop insulin resistance even on a standard chow diet [81]. Thus, in the liver, ChREBP drives efficient energy storage and distribution, while maintaining organism-wide sensitivity to sugar availability.

As with hepatic ChREBP, adipose tissue ChREBP is a key mediator of carbohydrate utilisation and insulin sensitivity. The expression of adipose tissue ChREBP and ChREBP target genes appears to be directly related to glucose availability. In white adipose tissue (WAT) there is a dose-dependent linear relationship between the expression level of glucose transporter GLUT4 and that of ChREBP. Supporting the causal nature of this relationship is the fact that mice with adipose-specific knockout and overexpression of GLUT4 have reduced and elevated levels of ChREBP and ChREBP target genes, respectively [13]. Mice with an adipose-specific ChREBP knockout exhibit insulin resistance [82] whereas mice overexpressing constitutively active adipose tissue ChREBP are thin despite upregulated DNL, have increased insulin sensitivity and experience markers of WAT browning [83]. These results correlate with the discovery that ChREBP appears to play an important role in thermogenesis in brown adipose tissue (BAT), and further imply a developmental role for ChREBP in driving the conversion of WAT to BAT [84, 85].

In the small intestine, ChREBP appears to be required for fructose absorption. On a high fructose diet, intestine-specific knockout mice experience dramatic weight loss and signs of malabsorption. In contrast, liver-specific knockouts can tolerate a high fructose diet (although they develop steatohepatitis) [86, 87, 79]. Thus, the insulin-sensitising actions of hepatic and adipose tissue ChREBP are complemented by its facilitation of fructose absorption in the small intestine.

This tissue-specific evidence of ChREBP function gives clues as to the role of ChREBP at the whole-body level. The hypothermia and death seen in sugar-fed knockout mice may be a result of digestive absorption complications (resulting from the loss of intestinal ChREBP), synergising with impaired thermogenesis (due to lost adipose ChREBP) leading to weight loss and extreme hypothermia. Conversely, ChREBP can be seen to act via an intestine-adipose axis to allow successful sugar uptake and to ensure that, in the event of carbohydrate excess, metabolism is shifted to take advantage of this nutrient imbalance. Intestinal sugar absorption causes an increase in circulating carbohydrates, leading to the activation of adipose tissue ChREBP. Adipose ChREBP then promotes a shift from fat to carbohydrate metabolism, observable by an increase

in organism respiratory exchange ratio, and allows for the management of excess carbohydrates via the combined effect of increased total BAT and increased thermogenesis in BAT [81]. The known functions of ChREBP in the liver, adipose tissue and small intestine are summarized in Figure 1.5.

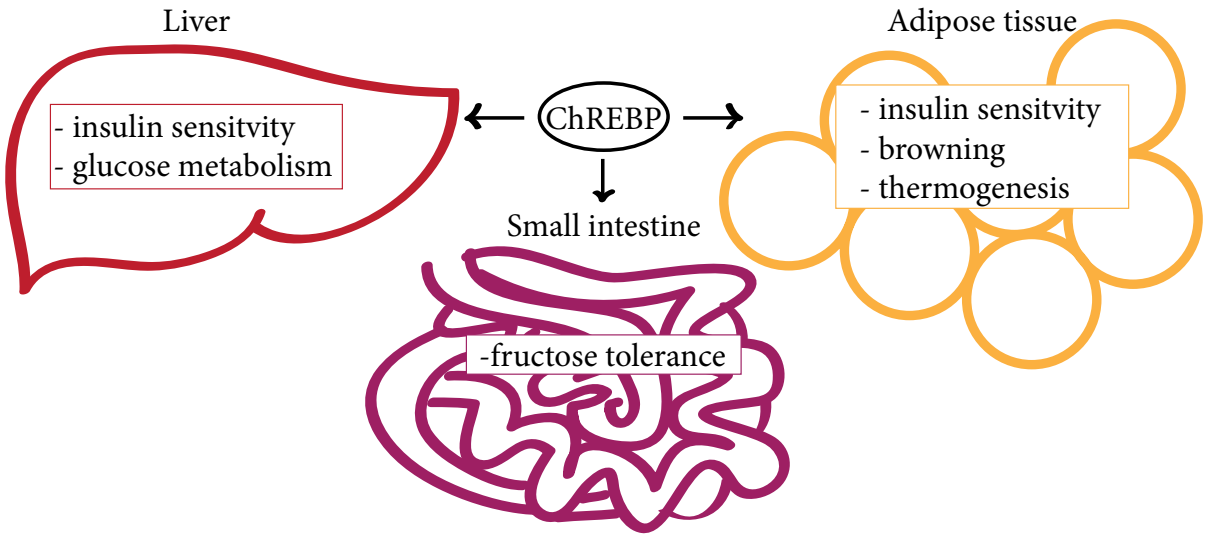


Figure 1.5: **ChREBP promotes whole-body energy homeostasis**
ChREBP is central to insulin sensitivity and glucose metabolism in the liver. Adipose tissue ChREBP promotes insulin sensitivity, adipose browning and thermogenesis. ChREBP expression in the small intestine is required for fructose tolerance.

ChREBP plays a complex role in mammalian health. In its capacity as a central regulator of metabolic energy homeostasis, ChREBP acts simultaneously to both promote and protect from a variety of pathological states, including NAFLD and diabetes. While ChREBP knockout mice are insulin resistant with a low body fat content relative to wild type [72, 81], high-fat diet fed mice overexpressing hepatic ChREBP have both fatty liver and improved glucose tolerance [80]. These correlations also extend to humans. In patients with nonalcoholic steatohepatitis (a severe form of NAFLD), hepatic ChREBP mRNA levels correlated positively with liver adiposity and negatively with degree of insulin resistance, implying that even in the disease state ChREBP simultaneously promotes liver fat levels and insulin sensitivity [80].

These data imply that ChREBP activity uncouples the commonly associated conditions of body fat and diabetes. In this way, they serve to highlight a gap in our understanding of the aetiology of insulin resistance. In light of this, it becomes particularly relevant to note that metformin, the most commonly prescribed oral diabetes drug, appears to decrease ChREBP activity in patients [88, 89, 90].

1.4 ChREBP may play a role in appetite regulation

ChREBP has been indicated to play a role in appetite regulation. One study found that, when ChREBP-/- mice were crossed with leptin-deficient ob/ob mice, the ob/ob ChREBP-/- progeny ate approximately 25% less and weighed approximately 40% less than their ob/ob counterparts (ChREBP-/- mice both ate and weighed slightly more than WT mice). This correlated with altered hypothalamic mRNA levels of the orexic hormones, NPY and AgRP [91], implying that ChREBP may interact with other factors to alter eating behaviour. Additionally, ChREBP-/- mice fed the high-fat high-sucrose "western diet" ate significantly less than their wild type counterparts [92], although this decreased consumption is likely attributable to sucrose intolerance, rather than to a direct role of ChREBP in regulating food consumption volumes [72]. Chow-fed liver-specific ChREBP knockout mice showed dramatically increased plasma levels of classical orexic hormone ghrelin and significantly decreased levels of the anorexic hormone leptin compared to wild type controls in the fed state [81]. This putative role of ChREBP in appetite regulation appears to be conserved between the vertebrate and invertebrate lineages of the protein. When the Drosophila homologue was deleted from insulin-producing cells, the resulting flies ate more [93], implying an anorexic function of ChREBP in these tissues, consistent with what was seen in the liver-specific knockout mice. Cumulatively, these data indicate a model in which the role of ChREBP in appetite regulation dovetails with its overall role in metabolic homeostasis. In the context of an organism without correct satiety signalling (the ob/ob mouse), ChREBP may promote overeating as the necessary negative-feedback signalling is lacking. Conversely, in the context of the insulin-producing cells which negatively regulate food intake in the healthy fly, deletion of ChREBP cuts short this inhibitory loop resulting in increased food consumption.

There is a growing body of evidence that ChREBP may mediate food intake, although the degree of this influence and its mechanism remain to be elucidated. Interestingly, ChREBP target protein FGF21, a key hormone in sugar homeostasis, promotes insulin sensitivity and adipose browning as well decreased sugar intake, providing a possible mechanism for many of the ChREBP functions described above [94, 26].

1.5 ChREBP responds to various markers of energy state

1.5.1 Both regulators and products of cellular metabolic networks play a role in ChREBP regulation

The physiological effects detailed in previous sections can be attributed to the many metabolites which regulate ChREBP activity. Published work points to glycolytic and pentose phosphate pathway intermediates, such as G6P and xylulose 5-phosphate (X5P), as well as metabolically sensitive kinases and phosphatases in the regulation of ChREBP activity. The extent to which each of the relevant factors affect ChREBP function is not yet clear.

PKA and PP2A were the first direct regulators of ChREBP activity identified. Treatment of primary rat hepatocytes with PKA inhibitor H-89 was found to increase activity of a ChREBP-driven luciferase reporter, while treatment with PP2A inhibitor cantharidic acid decreased it. Correspondingly, treatment with H-89 increased nuclear localisation of GFP-ChREBP while cantharidic acid had the opposite effect in these cells. Mutagenic work then found that consensus PKA phosphorylation sites S196 and T666 together appear to regulate nuclear localisation, DNA binding ability and activation of ChREBP. Treatment with cAMP analogue db-cAMP abolished ChREBP DNA binding activity in rat livers [95]. Experiments in rat liver extract demonstrated that pentose phosphate pathway intermediate X5P specifically activates PP2A and that cellular X5P formation was followed chronologically by ChREBP nuclear import and transcription of the ChREBP target gene *lpk* [96]. The prevailing story became that of opposing regulation by X5P signalling satiation and cAMP driving the response to hunger. Under low sugar conditions, liver cAMP levels rise leading to PKA activation, ChREBP phosphorylation, and subsequent repression. Under high sugar conditions, X5P is synthesized, activating PP2A which dephosphorylates ChREBP and activates it (Figure 1.6a).

Further studies however, painted a more complex picture. The phosphoinhibitory mutation S196A was found to have no dramatic effect on protein activity in primary hepatocytes and cellular cAMP does not change in response to media glucose in these cells [97, 98]. Together, these data implied two things. First, although cAMP treatment, and subsequent PKA activation, may be sufficient for ChREBP repression in primary hepatocytes, it is not the primary mechanism of glucose-sensitive repression in these cells. Second, the absence of PKA phosphorylation at its characterised target sites is insufficient for ChREBP activity in the absence of an additional, unknown factor. Later research pointed to F26BP as a likely candidate for this potential activator in primary hepatocytes [38].

In the context of this cumulative research it is vital to note that none of these data preclude a function for cAMP in vivo. In fact, in the physiological context of the liver, where PKA signalling plays a vital role in the coordination of metabolic flux, it is likely that cAMP driven PKA activation is an important regulator of ChREBP. The identification of F26BP as a ChREBP activator dovetails with the narrative of cAMP as a physiological inhibitor of ChREBP and leads to a model in which PKA and PP2A regulate ChREBP on two levels. First, via the direct phosphorylation and dephosphorylation discussed above, and second with the regulation of PFKFB, which controls cellular levels of ChREBP activator F26BP (Figure 1.6b).

ChREBP regulation in β-cells is different from that of hepatocytes, in keeping with the differential metabolic signalling in these cells. While PKA inhibition or phosphoinhibitory mutations are sufficient for constitutive nuclear localisation (although not protein activation) of hepatic ChREBP, the same effect is not seen in cultured β-cells [95, 96, 69, 98]. Despite this PKA insensitivity, β-cell ChREBP shows glucose stimulated nuclear localisation [70]. This behaviour may be attributable to Ca^{2+} ions, the influx of which is stimulated by high glucose levels in these cells [99]. Calcium chelation or inhibition by treatment with EGTA, diazoxide, or nifedipine all resulted in glucose-independent cytosolic sequestration of ChREBP in both the Min6 mouse insulinoma cell line and

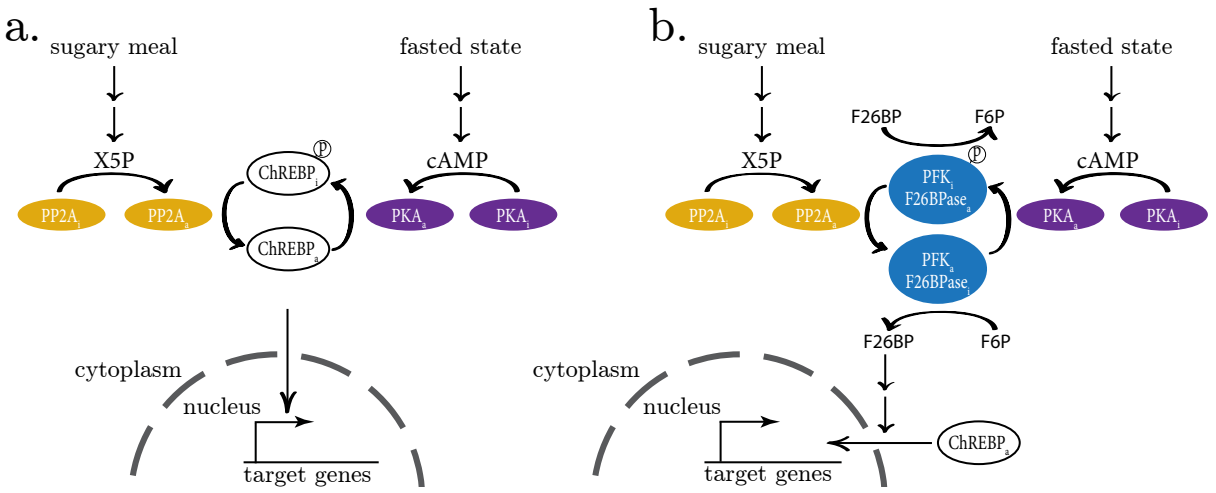


Figure 1.6: **ChREBP is regulated in response to markers of carbohydrate energy state** The reciprocal actions of PKA and PP2A regulate ChREBP both directly and indirectly. **a.** Hepatic ChREBP appears to be directly phosphorylated and dephosphorylated by PKA and PP2A, respectively. **b.** PKA and PP2A act on the bifunctional PFKFB enzyme, shown in blue. In the fed state, X5P activates PP2A which dephosphorylates PFKFB, promoting the kinase activity and inhibiting the phosphatase activity. PFKFB catalyses the phosphorylation of F6P to F26BP which promotes the activation of ChREBP. In the starved state, cAMP activates PKA which phosphorylates PFKFB, promoting its phosphatase activity and leading to the dephosphorylation of F26BP back to F6P. The physiological necessity of these pathways in the regulation of ChREBP activity remain under debate.

primary mouse β-cells. Overexpression of Sorcin, a Ca^{2+} binding protein and ChREBP interactor, reduced ChREBP nuclear localisation and target gene expression, whereas ChREBP was found to bind to its target gene promoters in a glucose-independent manner in Sorcin knockdown cells [100]. It is not however certain whether this nuclear localisation and target promoter binding in the absence of Sorcin leads to increased transcription. K^+-induced intracellular Ca^{2+} elevation was unable to stimulate transcription of ChREBP target gene ACC in the INS-1 rat β-cell line [101], implying that Ca^{2+}-elevation alone is insufficient for functional ChREBP activation. These data indicate that, in pancreatic β-cells, the glucose-dependent nuclear localisation of ChREBP occurs indirectly, via the Ca^{2+}-stimulated release of ChREBP cytosolic sequestration by Sorcin and that this release is sufficient for target promoter binding but not necessarily transcriptional activation.

G6P has also been implied to be responsible for ChREBP activation. Studies in the INS-1 derived 832/13 pancreatic β-cell line found that overexpression of glucokinase (the major hexokinase isoform found in β-cells) resulted in a leftward shift of the glucose response curve of ChREBP, correlating directly with the increased G6P levels found in these transfected cells. Conversely, overexpression of glucose-6 phosphate dehydrogenase, which converts G6P into pentose-phosphate pathway intermediates, resulted in a dose-dependent decrease in ChREBP activation at high glucose, simultaneously highlighting the importance of G6P and the minimal role of X5P in ChREBP activation in these cells [102]. These results were further supported by the opposing effects of hexokinase inhibitors and 2-deoxyglucose (2DG), which is processed by hexokinase to form 2-deoxyglucose-6-phosphate but is not further metabolised [102]. Consistent with these

findings were previous studies which found that 2DG but not mannoheptulose (a hexokinase inhibitor) were able to induce transcription of ChREBP targets *pklr* [103] and *acc* [101] in INS-1 cells.

Taken together, it appears that under low sugar conditions in pancreatic β-cells, ChREBP is sequestered in the cytoplasm by Sorcin. Upon the addition of media glucose, the cytoplasm is flooded with Ca^{2+}, driving the dissociation of Sorcin from ChREBP, allowing it to translocate to the nucleus and bind to target gene promoters. The increased extracellular glucose prompts a rise in intracellular G6P, which, through unknown mechanisms, activates ChREBP, allowing it to promote the transcription of target genes (Figure 1.7). G6P may also play a role in ChREBP activation in adipose tissue. In cultured adipocytes, 2DG but not 3-O-Methylglucose (a glucose analogue which is not processed by hexokinase) activates transcription of ChREBP target gene FAS, whose expression level correlates with cellular G6P levels [104].

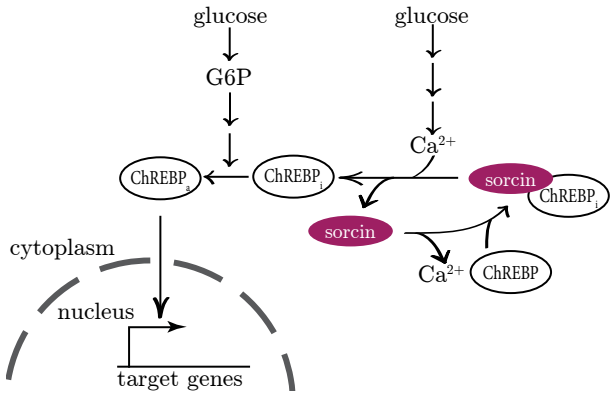


Figure 1.7: **ChREBP is regulated by G6P and Sorcin in pancreatic β-cells**
In low sugar conditions, Sorcin binds to ChREBP and sequesters it in the cytoplasm. When extracellular glucose levels are high, cytosolic Ca^{2+} levels increase, leading to release of ChREBP by Sorcin. High extracellular glucose levels also cause a spike in intracellular G6P which activates ChREBP.

Work in liver cells however, implies that G6P alone may not be sufficient for ChREBP activation. Transfection of primary hepatocytes with an F26BP-specific dephosphorylase (a kinase deficient version of the PFKFB enzyme, referred to as PFK-KD in Table 1.1) decreased ChREBP target gene expression and abrogated ChREBP binding at promoters without altering cellular G6P levels, while treatment with 2DG did not affect the expression of ChREBP target genes *g6pc* and *pklr* [38]. The F26BP data, along with the Ca^{2+} work in β-cells imply that, while G6P may be required, it is not necessarily sufficient for ChREBP activation. The published responses of ChREBP to mutations and changes in the discussed energy state markers are summarized in Table 1.1.

The results detailed above may tie into the observed effect of polyunsaturated fatty acids (PUFAs) on ChREBP activity. These molecules inhibit ChREBP, driving a reduction in ChREBP mRNA levels, nuclear localisation and the expression of ChREBP target genes [105, 14, 90, 106]. These repressive effects correspond with reductions in the activities of glucokinase and glucose-6 phosphate dehydrogenase and may occur indirectly, via modulation of the levels of glucose metabolites which regulate ChREBP. For example, linoleate treatment decreases levels of putative ChREBP activators G6P and X5P in primary hepatocytes [105]. Similarly, fatty acids inhibit PFK, in opposition to the action of F26BP [107, 36]. LXR, as a driver of PUFA synthesis, may also play a role in these effects [12].

Table 1.1: **ChREBP is regulated by diverse markers of energy state**

Type	Function	Condition tested	Model used	Reporter[a]	ChREBP response
Treatment	PKA inhibitor	H-89	primary hepatocytes	*lpk*	increased activity increased nuclear localisation [95]
	PP2A inhibitor	cantharidic acid	primary hepatocytes	*lpk*	decreased activity decreased nuclear localisation [95]
			832/13 β-cell line	*Gal4-UAS*	no effect on activity [69]
	cAMP analogue	db-cAMP	rat liver	-	decreased DNA binding [95]
		pCPT-cAMP	primary hepatocytes	*acc*	abolished activity [108, 97]
	PGI and G6PD inhibitor	2DG	832/13 β-cell line	*Gal4-UAS*	increased activity [102]
	Hexokinase inhibitor	D-mannoheptulose	832/13 β-cell line	*Gal4-UAS*	decreased activity [102]
	AMPK activator	metformin	INS1 β-cell line	*txnip*	decreased nuclear localisation decreased promoter occupancy decreased activity in high glucose [90]
		AICAR	primary hepatocytes	-	no effect on nuclear localisation [105]
	fatty acid	palmitate	INS1 β-cell line	*txnip*	decreased nuclear localisation decreased promoter occupancy decreased activity in high glucose [90]
	ketone bodies	acetoacetate	primary hepatocytes	*lpk*	decreased nuclear localisation in high glucose decreased activity [109]
		β-hydroxybutryate	primary hepatocytes	*lpk*	decreased nuclear localisation in high glucose decreased activity [109]

Table 1.1- *Continued from previous page*

Type	Function	Condition tested	Model used	Reporter[a]	ChREBP response
	Ca^{2+} chelator	EGTA	Min6 β-cell line	-	cytoplasmic under high glucose [100]
	PPARα agonist	WY14643	HB2 brown adipose cells	*chrebpβ*	decreased activity [15]
				fas	decreased activity [15]
			primary hepatocytes	*lpk*	decreased activity [14]
	LXR agonist	GW3965	Huh7 hepatocyte cell line	-	decreased promoter occupancy [32]
Transfection	F26BP dephosphorylase	PFK-KD	primary hepatocytes	-	decreased nuclear localisation decreased promoter binding [38]
	G6P dephosphorylase	G6Pase	832/13 β-cell line	*Gal4-UAS*	decreased activity [102]
	PPP enzyme	G6PD	832/13 β-cell line	*Gal4-UAS*	decreased activity [102]
		G6PD shRNA	832/13 β-cell line	*Gal4-UAS*	increased activity [102]
	Hexokinase isoform	Glucokinase	832/13 β-cell line	*Gal4-UAS*	increased activity [102]
	Ca^{2+} binding protein	Sorcin	Min6 β-cell line	-	decreased nuclear localisation [100]
		Sorcin si/shRNA	Min6 β-cell line primary β-cells	-	increased nuclear localisation in low glucose [100]
	nuclear receptor	PPARα	HB2 brown adipose cells	*fas*	decreased activity [15]
		LXR	Huh7 hepatocyte cell line	ChORE	increased activity [32]
Mutation	phosphomimic	S196D	primary hepatocytes	-	constitutively cytoplasmic [95]
		S140D S196D	primary hepatocytes	*lpk*	constitutively cytoplasmic no activity [98]
	phosphoinhibitory	S196A	primary hepatocytes	-	constitutively nuclear [95]

Table 1.1- *Continued from previous page*

Type	Function	Condition tested	Model used	Reporter[a]	ChREBP response
			832/13 β-cell line	*Gal4-UAS*	no effect on activity [110]
		S140A S196A	primary hepatocytes	*lpk*	constitutively nuclear no effect on activity [98]
		S196A T666A	832/13 β-cell line	*Gal4-UAS*	higher maximal activity no visible change in localisation [69]
			primary hepatocytes	-	increased nuclear localisation in low and high glucose [105]
		S140A S196A T666A	primary hepatocytes	*acc*	no major effect on overall activity partial rescue of cAMP inhibition [97]
		S196A S626A T666A	832/13 β-cell line	*lpk*	no major effect on overall activity no effect on forskolin inhibition [31]

[a] The reporter column refers to the promoter used in reporter assays, where relevant.

The differences in responses observed by different groups are likely due to a combination of factors including the tissue studied and the type of reporter used to study it. Reporter assays which use endogenous target gene promoters (e.g. the *lpk* promoter used by the Uyeda lab [95, 96, 109, 111]) are asking a fundamentally different question than those which use foreign promoters (e.g. the Gal4-UAS driven luciferase expression assayed by the Chan lab [69, 112, 102]). The former is probing the response of the ChREBP activation *pathway*, the activity of which is not always due to the activation state of the ChREBP protein exclusively. The latter is probing the activation state of the ChREBP protein specifically, at the potential expense of the relevant physiological context. For example, while treatment of hepatocytes with PUFAs or WY14643 (a PPARα agonist) has the same effect on expression of the *lpk* promoter-driven luciferase reporter construct, WY14643 appears to act directly by regulation of ChREBP, while *n*-3 Polyunsaturated Fatty Acids exert their effect via MLX, the obligate binding partner of ChREBP [14]. Similarly, the finding that cantharidic acid treatment decreased activity of a *lpk* promoter driven luciferase [95] but not a Gal4-UAS-driven construct [69] may indicate that the phosphorylation removed by PP2A doesn't directly regulate ChREBP activity *per se* so much as its ability to interact with other factors necessary for transcriptional activation in the absence of a nuclear localisation sequence (NLS) containing DNA binding tag (i.e. Gal4).

The above interpretation of results is additionally complicated by the fact that labs which use different types of reporter constructs often also look at this regulation in different tissues. As discussed above, although the overall function of ChREBP appears to be promotion of whole body energy homeostasis, this function is accomplished by different, tissue-specific actions. This tissue specificity seen on the organismal level is reflected by distinct signals of satiety and starvation on the biochemical level. While a glucose-rich meal induces hepatic PP2A activation [96], the opposite is true in pancreatic β-cells, where intermediates in the glucose metabolic network actively inhibit PP2A [113]. Further, while the pentose-phosphate pathway (which produces X5P) is a major route for liver glucose metabolites [114], it serves a comparatively minor role in the metabolism of pancreatic β-cells [115, 116]. Overall, the X5P-PP2A mechanism of ChREBP activation appears to hold metabolic relevance for liver but not pancreatic β-cells.

As with X5P and PP2A, the roles of cAMP and PKA in liver metabolism differ markedly from those in β-cell metabolism. In hepatocytes, cAMP is a marker of the low energy state, whereas cAMP production is actually stimulated by glucose in pancreatic islets. There is also evidence that the response of hepatocytes to cAMP is dependent on both duration of exposure [117] and cell cycle phase [118]. This temporal sensitivity is consistent with the finding that RNAPII occupancy at the promoter of ChREBP target gene *lpk* increases within the first few hours after glucose addition then returns to basal levels after 24 hours [14] and implies that the degree and amount of ChREBP reporter activity observed may be heavily dependent on the time after induction the assay is performed at. The cumulative body of work suggests that ChREBP is regulated by metabolites in a complex, tissue-specific manner.

1.5.2 The regulation of ChREBP by energy state markers is mediated by its interaction with 14-3-3

ChREBP activity is regulated in a complex manner by interaction with 14-3-3. The association was initially discovered in a yeast two-hybrid screen for ChREBP binding proteins which found that 14-3-3 associates with the N-terminal region of ChREBP, corresponding with what is now known as the GSM. Phosphatase treatment and the S196A/S568A/S626A/T666A quadruple mutation both abrogated pulldown, leading to the hypothesis that the ChREBP–14-3-3 interaction is phosphorylation-dependent [119]. This was later supported by isothermal titration calorimetry (ITC) experiments, which found that 14-3-3β and a ChREBP peptide containing residues 125-142 showed a K_d of 0.45 µM for the S140 phosphorylated peptide and no interaction for the dephosphorylated peptide (the other 14-3-3 isoforms also bound albeit with much lower affinities) [98]. This picture was complicated upon publication of the crystal structure of the α2-helix of ChREBP bound to 14-3-3. In this structure, ChREBP–14-3-3 binding is mediated by coordination of a free sulfate ion rather than a phosphorylated residue. In the physiological conditions of the live cell (as opposed to the 2.2 M ammonium sulfate solution used for crystallisation) it is very possible that this coordinating ion is a phosphate rather than a sulfate molecule. Thus, structural data indicate that the ChREBP–14-3-3 interaction occurs in a phosphorylation-independent manner and may instead be mediated by the coordination of a free phosphate ion [120] (Figure 1.8).

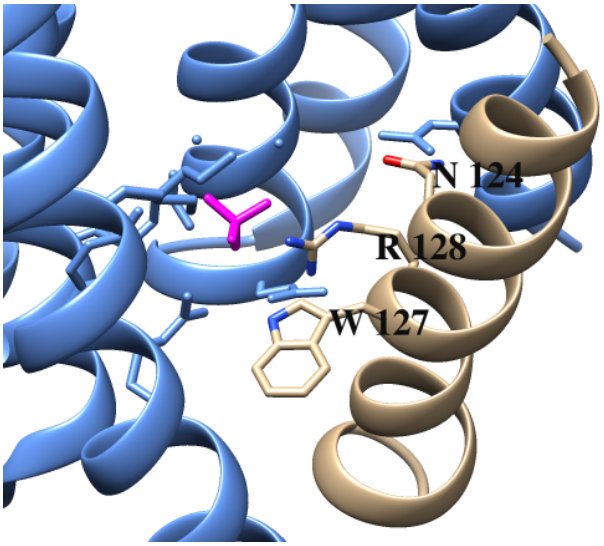


Figure 1.8: **Structure of the ChREBP α2-helix bound to 14-3-3β**
$ChREBP_{117-137}$, including the α2-helix is shown bound to 14-3-3β. 14-3-3β is shown in blue and the ChREBP peptide is in tan. The sulfate molecule coordinated between the two proteins is shown in magenta. This position may be occupied by a phosphate in vivo. Residues N124, W127 and R128 of ChREBP, which appear to form hydrogen bonds (not shown) with residues N226, R58 and E182, respectively of 14-3-3, are labelled. Image derived from Protein Data Bank (PDB) crystal structure 4GNT [120] and visualised using UCSF Chimera [121].

14-3-3 appears to play opposing roles in ChREBP activity, promoting cytosolic sequestration in low, and ChREBP activation in high energy conditions, respectively. Luciferase reporter assays in HEK293 cells found that 14-3-3 overexpression inhibited ChREBP activity in a dose-dependent manner [119]. Pulldown and ITC data, in conjunction with experiments in rat liver extracts indicate that the interaction between ChREBP and 14-3-3 is promoted by low energy signals, namely AMP and the ketone body β-hydroxybutyrate. These molecules bind directly to the ChREBP–14-3-3 heterodimer, rather than associating with the isolated proteins, and stabilize the complex. This binding appears to block ChREBP nuclear localisation by inhibiting its interaction with importinα [109, 111]. In the reverse direction, the phospho-mimic mutations S140D and

S196D appeared to cooperate to promote binding in the presence of nuclear export protein chromosomal maintenance 1 (CRM1) [98]. Together these data indicate that under low energy conditions in liver and HEK293 cells, 14-3-3 represses ChREBP activity by preventing its nuclear import and promoting nuclear export via the starvation signals provided by AMP, ketone bodies and PKA-dependent phosphorylation (Figure 1.9).

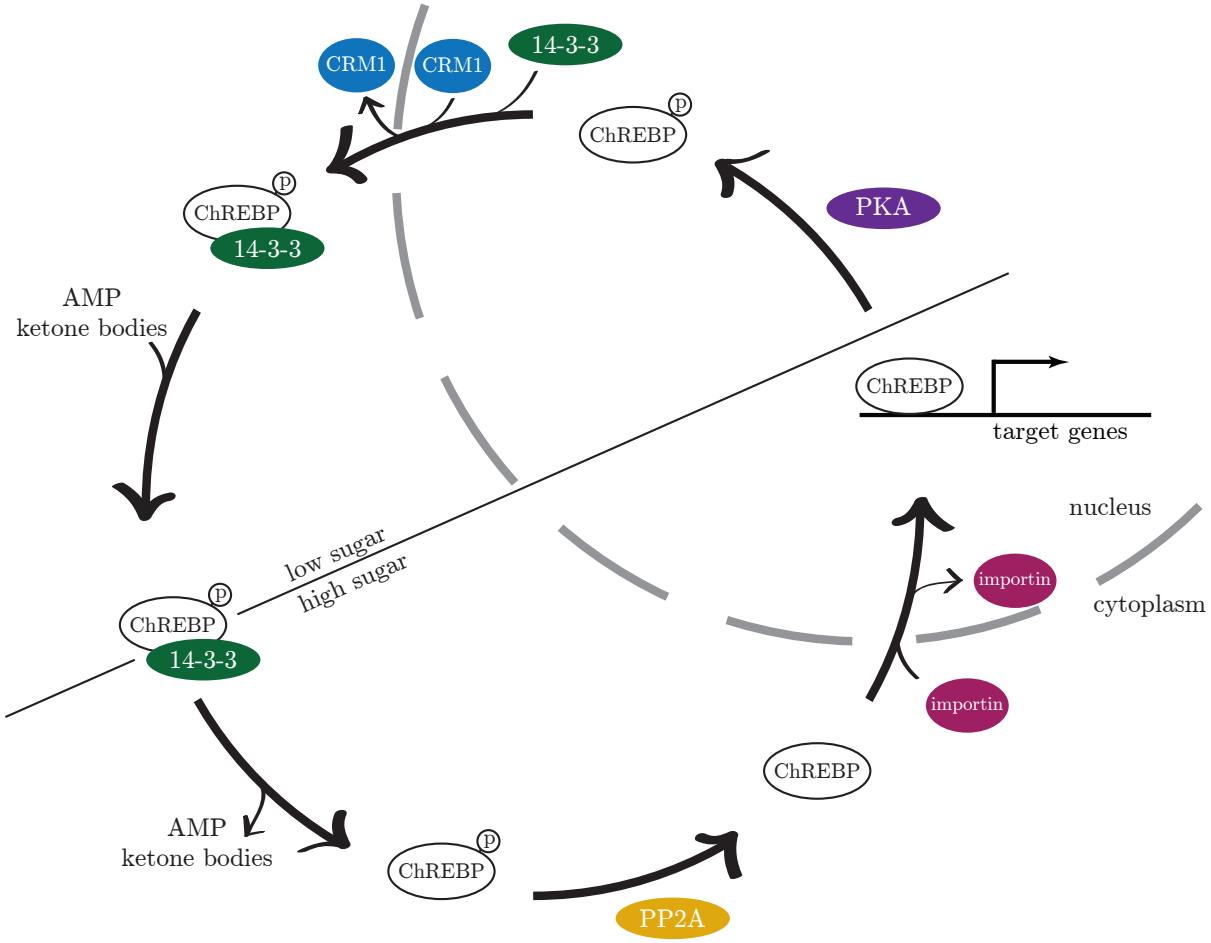


Figure 1.9: **14-3-3 mediates the interaction between ChREBP and metabolic energy signals**
In low sugar conditions, hepatic PKA is activated, phosphorylating ChREBP and increasing its affinity for 14-3-3 and the CRM1 nuclear exporter. The levels of AMP and ketone bodies rise and these metabolites associate with the ChREBP-14-3-3 complex, inhibiting its nuclear import. Increases in intracellular glucose metabolites lead to the activation of PP2A, which catalyses the dephosphorylation of ChREBP. Dephosphorylated ChREBP is able to dissociate from 14-3-3 and bind to importin which transports ChREBP into the nucleus where it can activate the transcription of target genes. A number of lines of evidence have cast doubt on this as the universal mechanism of ChREBP regulation however, as discussed in the text.

To complicate matters, pulldown data in 832/13 cells indicate that ChREBP has equal association with 14-3-3 under both low and high glucose conditions. Mutagenic studies found that the R128A mutation (in the crystal structure, R128 of ChREBP appears to directly hydrogen bond with E182 of 14-3-3β [120]) not only abolishes the interaction between ChREBP and 14-3-3, but also ChREBP transcriptional activation under high glucose conditions. Furthermore, overexpression of difopein, a competitive inhibitor of 14-3-3 proteins, results in a dose-dependent inhibition of

ChREBP activity under high glucose conditions. Additionally, 14-3-3 overexpression in these cells leads to a slight increase in ChREBP activation under high glucose conditions. Finally, despite its previously described role as a cytosolic sequestrant of ChREBP, 14-3-3 actually appears to have a higher relative nuclear abundance than ChREBP in 832/13 cells [112]. Thus, in contrast to the liver data, 14-3-3 appears to activate ChREBP in the 832/13 pancreatic β-cell line.

Cumulatively, the body of work surrounding the role of 14-3-3 in ChREBP activity indicates the following. First, that 14-3-3 is necessary for ChREBP activity. Second, 14-3-3 contributes to repression of ChREBP activity under low energy conditions. Finally, under certain, possibly tissue or cell-type-specific circumstances, 14-3-3 can promote maximal activation of ChREBP.

1.6 ChREBP is regulated by an N-terminal glucose-sensing module (GSM)

1.6.1 The GSM binds directly to G6P and AMP in vitro

ITC and microscale thermophoresis (MST) experiments in my own lab have shown that the purified GSM–14-3-3 complex binds directly to G6P in vitro with low µM affinity (Thomas Pysik, unpublished data). These data indicate that ChREBP may be, to our knowledge, the first known metazoan transcription factor to be directly activated by binding to a glucose metabolite.

G6P binding to the GSM–14-3-3 heterodimer was found to be competitive with AMP binding. These results directly imply a model of ChREBP activation in which, upon sugar feeding, cellular G6P levels rise and G6P is able to compete out AMP for GSM binding, allowing protein activation. Interestingly, size exclusion chromatography and analytical ultracentrifugation found that the GSM and 14-3-3 remain in complex in the presence of saturating concentrations of G6P (Thomas Pysik, unpublished data), which is consistent with previous work demonstrating that 14-3-3 is necessary for maximal activation of ChREBP [112].

Changes to the GSM–14-3-3 interface are sufficient to disrupt protein function in vitro without preventing GSM–14-3-3 association. Alanine mutation of tryptophan 127, a residue which appears to directly hydrogen bond with 14-3-3 in the crystal structure (Figure 1.8, [120]), is sufficient to abrogate both G6P and AMP binding of the GSM although it does not destroy the GSM–14-3-3 interaction. Similarly, alanine mutation of 14-3-3 residues serine 47, lysine 51 or arginine 58, located at the GSM–14-3-3 interface in the crystal structure resulted in weakened or abrogated binding of G6P and/or AMP without destroying the GSM–14-3-3 complex (Thomas Pysik, unpublished data). These results are consistent with the crystal structure showing that AMP binding occurs at the ChREBP–14-3-3 interface [111], and further imply that the integrity of this interface is required for G6P binding.

1.6.2 The GSM is composed of an inhibitory LID and an activatory GRACE

ChREBP activity and regulation can be seen as a function of the domain structure of the protein. ChREBP contains an N-terminal GSM and a C-terminal bHLH DNA-binding domain of the Myc/Max/Mad superfamily [69, 122]. To activate transcription, ChREBP heterodimerises with MLX, and binds to tandem E-box sequences known as Carbohydrate Response Elements (ChOREs) [123]. The GSM regulates this activation and is made up of a low glucose inhibitory domain (LID) and a glucose-response activation conserved element (GRACE). When fused to a DNA binding region, the GSM is sufficient to act as a glucose-responsive transcriptional activator in live cells [69]. The LID is a repressive module and ChREBPβ, which lacks all but the last 15 residues of the LID, is a glucose-insensitive constitutively active transcription factor [13]. The GRACE on the other hand has activatory properties and is necessary and sufficient for transcriptional activation when fused to a DNA-binding module [69, 124]. The LID and GRACE regulate each other in a manner mediated by conserved regions within the two [69].

1.6.3 The GSM contains six highly conserved regions known as MCRs

The glucose sensitivity of the GSM is dependent on a series of 6 highly conserved stretches known as Mondo conserved regions (MCRs) [68]. Deletion mutagenesis and luciferase assays in 832/13 cells imply that each of the MCRs 1-4, which are located in the LID, are necessary for repression of ChREBP under low sugar conditions, while MCR5 in the GRACE domain may be the object of this basal repression [1] [125] (Figure 1.10).

Figure 1.10: **The GSM is composed of six MCRs**
MCR2 contains a CRM1-dependent nuclear export sequence (NES), MCR3 contains a 14-3-3 binding site and MCR4 contains a bipartate NLS. The functions of each MCR are discussed further in the text. The boundaries used by various labs to define each MCR are detailed in Table 1.2.

Each MCR appears to play a distinct but complementary role in protein activity. Attempts by various labs to disentangle the functions of individual MCRs have produced a wealth of data implying complex interdependent regulation. Much of this data has apparently contradictory components likely due to a combination of two factors. First, MCR boundaries are not consistent across publications and, in the case of MCRs 2 and 3, actually overlap (Table 1.2). Second, different groups used either an endogenous promoter or a Gal4 reporter system which, as discussed previously, asks different questions of the system. Specifically, only two groups were responsible for thoroughly querying the role of MCRs in ChREBP. The Towle lab used a ChORE promoter

[1] In the upcoming section, portions of the LID domain will be described as "required for protein activity". Each occurrence of this phrase comes with the implicit acknowledgement that a full deletion of the LID is constitutively active and that the region in question is only necessary in the context of a construct containing other portions of the LID domain.

Table 1.2: **The MCR borders of ChREBP used by various labs and publications**

		MCR					
Group	Publication	1	2	3	4	5	6[a]
Chan	[112][b]	1-71	72-99	100-143	144-196	-	-
McFerrin & Atchley	[68]	43-60	80-99	100-149	153-192	271-305	242-271
Towle	[70]	45-58	82-112	118-135	170-195	275-290	-
	[125]	45-58	81-111	117-149	166-192	277-294	-
	[125][b]	1-58	86-114	116-136	141-197	277-294	-

[a] MCR6 is not included in the majority of these works as it was not identified until 2012 [68]. This late identification is why MCR6 actually lies between MCRs 4 and 5.

[b] The numbers in these rows indicate boundaries used in deletion experiments in the respective publications as opposed to those defined in a figure. This is why one publication can have two sets of boundaries. Note that insertion/rescue mutations performed by a given lab could follow the published figure boundaries, the deletion boundaries, or some combination of the two. When this distinction seems likely to be relevant the residues used are specifically noted in the text.

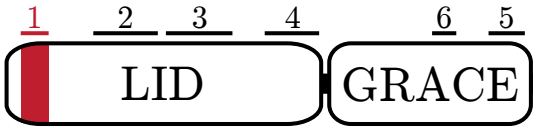
driven luciferase reporter to assay the ability of FLAG-tagged constructs to rescue the repression caused by a dominant negative form of ChREBP. The Chan group used a traditional Gal4-UAS driven luciferase reporter system which provided the constructs tested with an exogenous NLS.

MCR1 may regulate glucose-responsive nuclear localisation of ChREBP

The Towle lab found that deletion of any individual MCR from 1-4 renders full length ChREBP functionally inactive in 832/13 cells [125]. Interestingly, similar experiments by Chan group members in the same cell line found that deletion of any single MCR *except* for MCR1 abolished glucose responsiveness. In their study, deletion of MCR1 greatly reduced but did not abolish protein activity under high glucose conditions [112]. It is therefore possible that part of the function of MCR1 lies in repression of cytoplasmic sequestration under high glucose conditions and that this function was overridden by the nuclear localisation sequence of the Gal4 tag. The role of MCR1 in the regulation of subcellular localisation is supported by the fact that mutations in the MCR1 of human MondoA abolish 2DG-stimulated protein nuclear localisation in the L6 rat myoblast cell line [126].

Phosphorylation of MCR1 may play a role in this regulation. Mass spectrometry of primary hepatocytes identified S56 as a site of phosphorylation in live cells. The Gal4-tagged S56D phosphomimic construct shows increased luciferase reporter expression under both low and high glucose conditions in hepatocytes. When the same mutation is assayed for activation of a ChORE-driven reporter, increased high glucose but not low glucose activation is seen. The phosphoinhibitory S56A mutation behaved similarly to the wild type construct in the context of both a ChORE-driven and a Gal4-UAS driven reporter, underscoring the idea that, while the presence of this phosphorylation has an activatory effect, it is not required for protein activity. This is consistent with a model in which MCR1 phosphorylation contributes to maximal activation of ChREBP but is itself insufficient to override the cytoplasmic retention maintained by MCR1

in low sugar conditions in hepatocytes [108].



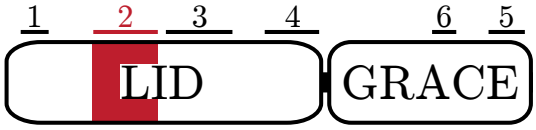
MCR1
- possible role in glucose-responsive nuclear localisation
- S56 phosphorylation increases maximal activity
- not necessary for glucose-responsive activity

Figure 1.11: **MCR1 may regulate glucose-responsive nuclear localisation of ChREBP**
The known and putative functions of MCR1 are shown. MCR1 is unique among MCRs 1-4 in not being required for glucose-responsive protein activity.

When ChREBP regulation is investigated through the lens of this hypothesis, it is important to note that nuclear localisation is necessary but not sufficient for protein activity. This is evidenced by experiments with CRM1 nuclear export inhibitor leptomycin B (LMB) and the L89A mutation of the ChREBP nuclear export sequence (NES). Both LMB treatment and the L89A mutation induce constitutive nuclear localisation but neither results in altered protein activity [70]. The known and putative functions of MCR1 are summarised in Figure 1.11.

MCR2 contains a nuclear export sequence and represses ChREBP activity in a nuclear export independent manner

MCR2 contains a functional NES between residues 86 and 95 and is required for export of ChREBP from the nucleus [119]. Deletion of this region abolishes protein activity despite constitutive nuclear localisation [98]. Specific mutations in the NES lead to constitutive nuclear localisation and have varying effects on protein activity in 832/13 cells. L89A, as mentioned previously, shows no effect on ChREBP reporter activity, while the F90A mutant shows more than double the activity of the wild type protein under high glucose conditions [70]. Similarly, the L86A/L93A double mutant has dramatically increased nuclear localisation relative to the wild type protein but shows no activity in a luciferase reporter assay [112]. Further highlighting the importance of the NES-independent role of MCR2 in overall ChREBP activation is the fact that a FLAG tagged Δ85-197 mutant has only minorly decreased activity relative to a ΔLID construct, while a Gal4-tagged Δ72-196 version shows a dramatic reduction in activity. This implies that the region 72-85, just upstream of the NES and defined as the N-terminus of MCR2 by the Chan lab (Table 1.2) is required for ChREBP activity in a manner not likely to be mediated by nuclear localisation [112, 125]. Together, these data indicate that MCR2: 1. acts in ChREBP nuclear export, 2. plays a repressive role in ChREBP activity and 3. that these two roles are functionally separable. The known functions of MCR2 are summarised in Figure 1.12.



MCR2
- functional NES
- NES-distinct repressive function

Figure 1.12: **MCR2 contains a nuclear export sequence and represses ChREBP activity in a functionally separable manner**
The known functions of MCR2 are shown.

MCR3 binds 14-3-3 and represses ChREBP activity by distinct mechanisms

MCR3 is necessary and sufficient for the interaction between ChREBP and 14-3-3 [112, 120]. Although MCR3 appears to be required by ChREBP for activation, it also appears to play a negative regulatory role in ChREBP function and a crystal structure of the MCR3-14-3-3-interface with AMP bound has been published [111]. Attaching MCR3 to the constitutively active LID deletion mutant of ChREBP results in a marked decrease in construct activity [112]. Interestingly, MCR3 is able to compete out the activity of a Gal4-ChREBP reporter construct in a dose-dependent manner in trans. This property is retained even in mutants of MCR3 which can no longer bind 14-3-3 [112, 125]. These data are consistent with the following two ideas, either MCR3 binds factors necessary for ChREBP activation, and/or MCR3 interacts directly and repressively with elements in the activatory GRACE domain. The fact that the GRACE domain by itself functions as a strong transcriptional activator indicates that, if the former hypothesis is true, it is only relevant in the context of a protein containing some fraction of the LID domain where the transactivatory activity of the GRACE is dramatically inhibited (in high glucose conditions) or abolished (in low glucose conditions). Consistent with the activator binding hypothesis is the fact that 14-3-3 binding is required for ChREBP activation and loss thereof reduces protein activity under high glucose conditions [112], as discussed in the previous section. It is therefore possible that MCR3, by way of its interaction with 14-3-3 and additional factors, acts to simultaneously *release* repression and promote activation of ChREBP under high glucose conditions. The known functions of MCR3 are summarised in Figure 1.13.

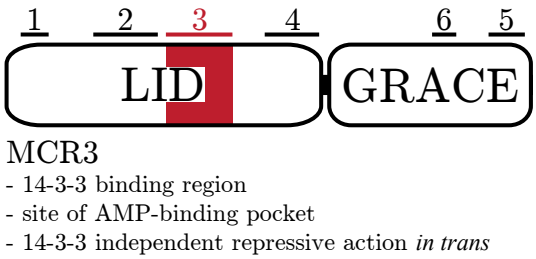


Figure 1.13: **MCR3 binds 14-3-3 and represses ChREBP activity by distinct mechanisms**
The known functions of MCR3 are shown. MCR3 is unique among the MCRs for its ability to repress ChREBP activity in trans.

MCR4 is required for both nuclear localisation and low glucose repression

To bring the functions of MCR4 into focus, we return to the two orthologous single MCR deletion experiments from the Towle and Chan labs. The Towle group, who used a FLAG-tagged construct and a ChORE containing luciferase reporter gene, found that a deletion mutant of any single MCR1-4 was sufficient to abolish protein activity. The Chan lab meanwhile found that, with their Gal4-tagged construct the ΔMCR4 mutant exhibited media glucose-insensitive activity. They went on to test a variety of multi-MCR deletion mutants and found that all those without MCR4 also lacked glucose-responsiveness. The above data highlight the apparent simultaneous functions of MCR4, it contains a functional NLS (Figure 1.14) and thus appears to be required for activity in the full-length protein; however, when the need for an NLS is removed, its parallel role as a repressor under low sugar conditions is highlighted. Hydrogen/deuterium exchange mass spectrometry (HDX-MS) data, generated in a collaboration between my lab and the Rand lab at

the University of Copenhagen showed that G6P binding by the purified GSM leads to altered solvent accessibility of the NLS, including residues 168-176 seen in the crystal structure in Figure 1.14, [127] (Thomas Pysik in collaboration with Kaspar Rand and Johanna Veiga, unpublished data), consistent with a role of this region in glucose-responsiveness of the protein.

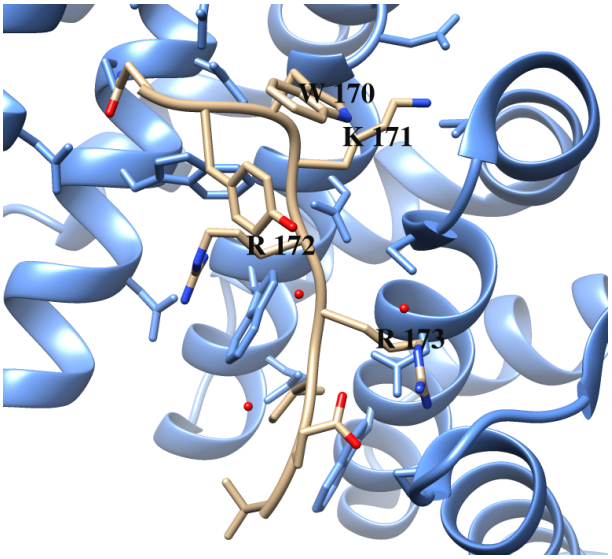


Figure 1.14: **Structure of importinα bound to a peptide of the MCR4 NLS** $ChREBP_{168-176}$ is shown bound to importinα. Importinα is shown in blue and the ChREBP peptide is in tan. The universally conserved W170 as well as residues from the conserved KRR171-3 basic patch have been labelled. Image derived from PDB crystal structure 6MJL, unpublished [127]. Structure visualised using UCSF Chimera [121].

Although MCR4 is clearly required for glucose sensitivity, its presence is not sufficient for glucose-responsive activity. Constructs lacking MCRs 1-3 show glucose-insensitive luciferase reporter activity in both cultured hepatocytes and 832/13 cells [112, 108]. It is also relevant to note that caChREBP, a constitutively active LID-deletion form of ChREBP used in many studies (including but not limited to [124, 128]), lacks the NLS of the MCR4 but is nuclear localised, implying that, in the absence of the LID, the MCR4 NLS is not required for nuclear localisation. The known functions of MCR4 are summarised in Figure 1.15.

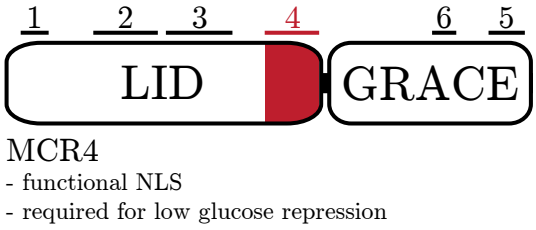


Figure 1.15: **MCR4 is required for both nuclear localisation and low glucose repression**
The known functions of MCR4 are shown. In addition to having an NLS, MCR4 is necessary but not sufficient for low glucose repression of ChREBP.

MCRs 1-4 in the LID act cooperatively to regulate ChREBP activity and localisation

Combinatorial deletion of MCRs within the LID domain serves to highlight their roles as regulatory elements necessary for protein glucose responsiveness and activity. Removal of MCRs 1-2, 2-3 and 1-3 all result in a construct with dramatically reduced or abolished activity [112, 125]. However, while the ΔMCR3-4 construct showed minimal ability to rescue repression by dominant negative ChREBP, the equivalent construct showed sizeable glucose-insensitive activity in the context of a Gal4 reporter. These data again highlight the importance of the additional Gal4 tag to allow disentangling regulation which occurs via subcellular localisation from that which is the result of

alternative mechanisms. The data from the ΔMCR3-4 mutant highlight the role of MCR4 in low glucose repression. In the context of the other combinatorial deletion mutants, it also showcases the unique role of MCR3 as a repressive module whose presence is nonetheless necessary for protein activity in an MCR4 containing construct. The ΔMCR4 mutant alone exhibits low levels of glucose-insensitive activity but when MCR3 is also deleted, the resulting construct is almost as active as a full LID deletion mutant [112]. Looked at another way, when MCR3 is tagged to the LID deletion construct, the resulting protein exhibits a dramatic reduction in activity relative to the LID deletion construct. This MCR3-ΔLID construct behaves indistinguishably from an MCR2-MCR3-ΔLID construct, indicating that the MCR2 is not responsible for significant repression in the absence of MCRs 1 and 4. Although MCR3-ΔLID and MCR2-MCR3-ΔLID behave the same, the addition of MCR4 to this construct (in other terms a ΔMCR1 construct) is sufficient to both return glucose responsiveness and dramatically decrease construct activity. Although it cannot be excluded that this return of glucose responsiveness is due to a sufficient portion of the LID domain returning to render the region functional, rather than a specific role of MCR4 per se, it is notable that MCR4 addition also decreases maximal activation of the construct. This, combined with the MCR4 data discussed in the previous section imply a specific function of this region in protein repression. Together, these data demonstrate the combinatorial role of each individual conserved region and suggest a model in which each MCR1-4 works together to shape the GSM behaviour seen in the whole protein.

MCRs 5 and 6 may regulate ChREBP activation

The GRACE is a strong transcriptional activation domain which, in the absence of the LID is sufficient to turn ChREBP into a constitutive activator [124]. Despite the overall role of the GRACE as a transcriptional activator, mutagenic experiments in MCR5 were consistent with its function as a repressive regulatory element in the context of the whole protein. The N278A mutant shows increased high sugar activation and the triple alanine mutation of MCR5 residues 275-277 or 289-291 results in proteins with increased transcriptional activity under both low and high glucose conditions. In fact, increased transcriptional activity was found for a variety of alanine mutations throughout MCR5. The activity changes were accompanied by increased nuclear occupancy in these mutants, consistent with the hypothesis that GRACE works together with the LID domain to regulate ChREBP glucose-responsiveness [70].

MCR5 has been proposed to cooperate with MCRs in the LID to mediate glucose-sensitive ChREBP repression. The S56D, H51A, N278A triple mutant showed approximately two-fold increased basal activity relative to the S56D, H51A MCR1 double mutant. A further mutation in the MCR2 NES created a S56D, H51A, F90A, N278A quadruple mutant with constitutive, glucose-insensitive activity [125]. Similarly, the addition of the MCR2 F90A mutation to the 275-277 or 289-291 triple alanine mutations discussed above is sufficient to drive an over two-fold increase in basal ChREBP activity relative to F90A or the respective triple mutants alone [70]. Taken together, these data demonstrate functional cooperation between MCR5 and MCRs 1 and 2, although the molecular mechanism of this cooperation remains to be elucidated.

MCR6 is less well characterised than MCR5. The totality of published data on MCR6 comes from two bioinformatic studies. Sequence alignment led to the hypothesis that MCR6 was responsible for both G6P binding and interaction with activatory factors such as p300. Specifically, it was found that MCR6 shares a motif with the G6P binding sites of proteins GPI and Gfat1 as well as a motif known to be bound by transcriptional activators such as p300 [68, 59]. These hypotheses have not yet been experimentally validated in published literature, however Nina Heppner in my lab demonstrated that ChREBP$_{39-251}$, which lacks the predicted G6P binding motif, can bind G6P with a K_d on par with that of the full-length GSM (Masters Thesis [129]), implying that this region is not instrumental in G6P binding.

The region directly downstream of the GSM promotes full protein activation

Although the GSM is necessary and sufficient for glucose-metabolite regulated activation of ChREBP, there is reasonable evidence that it is not the only part of the protein which functions in this role. The GSM deletion construct Gal4-ChREBP$_{299-864}$ shows activity in reporter assays and the activity of the isolated GRACE domain is increased by the addition of its downstream residues [69].

Early work in MondoA support these ChREBP data. Studies found that MondoA$_{322-445}$, which corresponds to the region directly downstream of the GRACE in ChREBP (overlapping the last six residues of this region), is a transcriptional activation domain [130]. Any potential protein regulation which may occur in this downstream region is not well characterised and warrants further study.

The mechanism of the functional interaction between the LID and GRACE domains remains uncertain

It is possible that the regulation seen between the LID and GRACE occurs via a domain-domain interaction. Consistent with this is the orientation-dependence of this regulation, the functional cooperativity seen between MCRs 1 and 2 and MCR5 and the fact that the MCR1-4 construct ChREBP$_{45-197}$ is able to compete out the activity of a ChREBP reporter gene in a dose-dependent manner in trans [69, 70, 125]. Additionally supporting these data is the evidence that, in the absence of the LID, GRACE acts as a constitutive activator yet, in the context of full-length ChREBP, mutations to MCR5 actually increase protein activity, consistent with the idea of direct-repressive inter-domain contacts.

As of yet however, there is no conclusive evidence for a direct domain-domain interaction. The LID-GRACE orientation-specificity of the glucose-response could be the result of misfolding of the reversed protein or of the LID and GRACE acting together to coordinate an unknown interacting factor. The mutagenic data from MCR5 may also be consistent with this indirect interaction. Similarly, the ability of the MCR1-4 construct to compete out the activity of a ChREBP reporter may be caused by competition for an activating factor. When GRACE-VP16 and Gal4-LID

were co-expressed in 832/13 cells, they were not able to activate reporter gene transcription. Interaction was also not seen in a pulldown [69]. Thus, although a direct interaction between LID and GRACE has been hypothesised [125], whether such an interaction occurs is not yet known.

1.7 Open questions regarding the mechanism of ChREBP activation

The mechanism of activation of ChREBP is understood in broad terms. Sugar metabolites lead to release of repression of the GRACE by the LID resulting in the transcriptional activation of target genes. However, the biochemical specifics of each step are controversial at best and not yet physically tested at worst. In order to understand the mechanism of ChREBP activation, we must first characterise the mechanism by which LID and GRACE regulate each other and protein activity. Is regulation the result of a direct LID-GRACE interaction, or is it mediated by an intermediary factor? How does G6P binding affect this inter-domain regulation? Furthermore, it is necessary to determine what factor or factors are responsible for this activation, as well as whether these factor(s) are conserved across tissue types and between species. With regard to the latter goal, there have been certain steps forward. Data from my own lab, showing that the purified GSM binds G6P, combined with experiments in the 832/13 pancreatic β-cell line indicate that direct G6P binding is required for ChREBP activation (Thomas Pysik, unpublished data), [102]. However, work in pancreatic β-cells and hepatocytes implies that, while G6P may be necessary, it is not sufficient for ChREBP activation [38, 100]. The question remains, to what extent is G6P responsible for the activation of ChREBP? Finally, we must disentangle the observed metabolic regulation of ChREBP subcellular localisation from that of protein activity.

In order to define the mechanism of ChREBP activation, I first performed Co-immunoprecipitation (CoIP) and fluorescence two-hybrid (F2H) assays, as well as fluorescence resonance energy transfer-fluorescence lifetime imaging (FRET-FLIM) experiments to determine whether the LID and GRACE interact and how this putative interaction is affected by the presence of G6P and media glucose conditions in vitro and in live cells, respectively. I complemented these experiments with targeted mutagenic studies on conserved residues in MCRs 3, 4 and 6. In these, I performed luciferase reporter assays and fluorescence microscopy experiments to dissect the role of these regions in protein activity and subcellular localisation and to disentangle the relationship between these two properties.

Chapter 2

Aims of the Project

The overarching aim of this project is to characterise the mechanism of activation of the transcription factor ChREBP.

2.1 First Aim

Determine whether the LID and GRACE interact and, if so, characterise the regulation of this interaction.

Since its initial characterisation, ChREBP has been known to be activated by environmental sugar intake in vivo. The detailed mechanism of this activation remains uncertain, however it is known to involve a release of intramolecular repression of the GRACE by the LID. Various lines of evidence highlight the possibility that this intramolecular regulation occurs via a direct interaction between LID and GRACE. Meanwhile, the specific factor or factors involved in triggering this activation remain controversial. Live cell data from other groups and in vitro data from my own lab show that G6P is necessary for ChREBP activation in cell culture and that the purified GSM of ChREBP is capable of directly binding G6P. I hypothesised that the LID and GRACE domains interact and that this interaction is disturbed by the direct binding of G6P, causing protein activation. To test this hypothesis, I performed CoIP, F2H and FRET-FLIM experiments.

2.2 Second Aim

Characterise the relationship between ChREBP activation and subcellular localisation.

Limited proteolysis and HDX-MS work performed in a collaboration between my lab and the Rand lab at the University of Copenhagen demonstrated that AMP binding and G6P binding induce distinct conformational changes in MCR3 or MCRs 3, 4 and 6 of the GSM, respectively. Notably, the perturbation seen in MCR6 is adjacent to a nonsynonymous missense SNP associated

with lowered circulating triglyceride levels. The putative role of G6P as a ChREBP activator, combined with the ability to change the solvent accessibility of these regions, indicated that characterisation of MCRs, 3, 4 and 6 may be key in dissecting the regulation of ChREBP activation and nuclear localisation. To characterise the relationship between ChREBP activation and subcellular localisation, I cloned point mutants of conserved residues within these regions and performed luciferase reporter assays combined with confocal, fluorescence recovery after photobleaching (FRAP) and photoactivation microscopic experiments.

Chapter 3

The LID and GRACE domain interact in a media glucose-responsive manner

3.1 Introduction

The activation of the GSM has been characterised as a release of intramolecular repression of the GRACE by the LID [69]. Various lines of evidence point to a domain-domain interaction being central to the activity of ChREBP, including the fact that the MCR1-4 construct is able to compete out the activity of a ChREBP reporter gene in a dose-dependent manner in trans [125]. However, a direct interaction has not been demonstrated [125, 69].

If ChREBP repression were caused by the association of LID and GRACE, ChREBP activation would require release of this interaction. Live-cell experiments indicate that G6P is necessary for ChREBP activation [102], and experiments from my lab have demonstrated that the GSM binds G6P directly (Thomas Pysik, unpublished data). It follows that G6P binding may drive the derepression of the GRACE by the LID. I hypothesised that the LID and GRACE domains interact and that this interaction is regulated by G6P or other glucose-responsive factors in vivo. To test this hypothesis, I performed CoIP, F2H and FRET-FLIM experiments on the domains of the human protein.

3.2 CoIP data indicate that the LID and GRACE interact

I performed pulldown assays using the LID and GRACE domains of human ChREBP. In these, I transfected U2OS 2B2 cells stably expressing GFP (negative control) or GFP-LID with Cherry, GRACE-FLAG-Cherry, or Cherry-GRACE. I harvested and lysed the transfected cells, and pulled down GFP-containing protein in the lysate using an anti-GFP nanobody [131] conjugated to agarose beads (Figure 3.1a). GFP-LID pulled down both GRACE-FLAG-Cherry

and Cherry-GRACE with similar efficiency. Critically, GFP alone was unable to pull down Cherry, GRACE-FLAG-Cherry, or Cherry-GRACE. Addition of 100 µM G6P to the IP buffer had no effect on the interaction detected (Figure 3.1b, c). It should be noted that although the detected interaction of the separately expressed domains is weak, in the physiological context of the full-length protein, these domains are in constitutive proximity and a weak interaction may likely be relevant for binding.

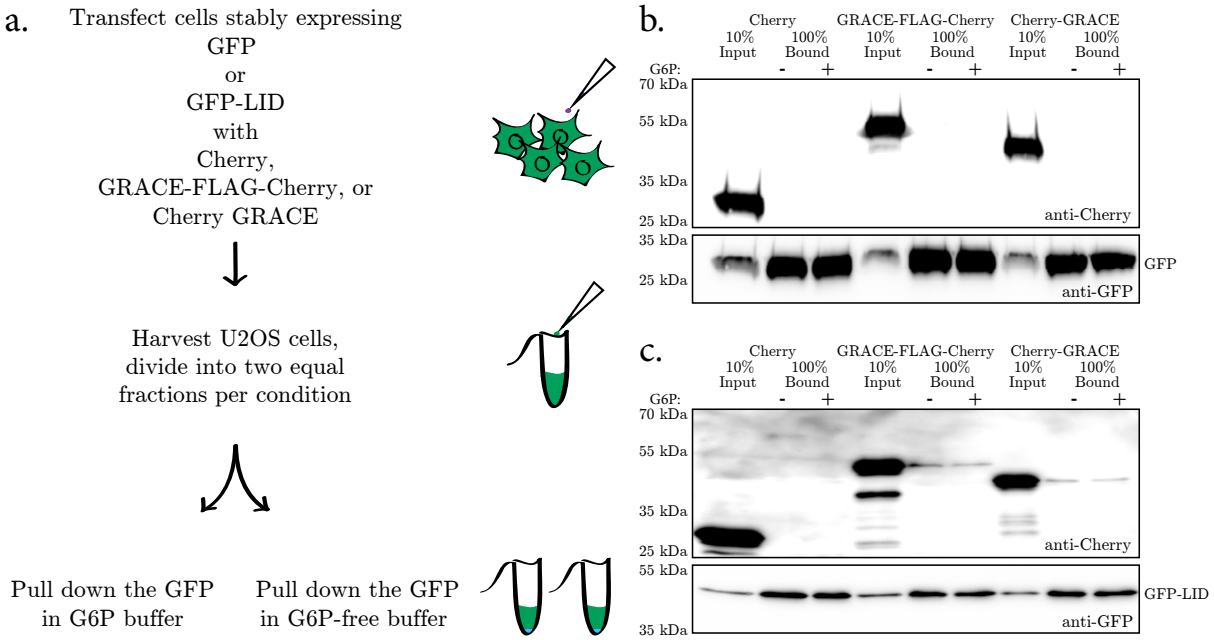


Figure 3.1: **Pulldowns demonstrate that the LID and GRACE domains interact**
a. U2OS 2B2 cells stably expressing GFP or GFP-LID were transfected with the indicated Cherry-tagged constructs in 25 mM glucose media. Two days after transfection cells were harvested. The lysates were split into two equal fractions and pulled down in the presence or absence of 100 µM G6P. **b.** GFP alone does not pull down Cherry, GRACE-FLAG-Cherry, or Cherry-GRACE. **c.** GFP-LID pulls down both GRACE-FLAG-Cherry and Cherry-GRACE. 10% input and 100% bound are shown for all samples. The molecular weights of the constructs used are as follows: GFP (29 kDa), GFP-LID (46 kDa), Cherry (29 kDa), GRACE-FLAG-Cherry (40 kDa), Cherry-GRACE (39 kDa). The LID was defined as ChREBP_{37-192}. The GRACE was defined as $\text{ChREBP}_{197-298}$. Residue numbers are based off of Uniprot ID Q9NP71. Plasmid nucleotide sequences were derived from NCBI Reference Sequence: NM_032951.2.

These data demonstrate that the LID and GRACE domains interact weakly but do not support the hypothesis that this interaction is disrupted by G6P. However, as the association was only weakly detected, it remains possible that any changes in binding strength are outside of the dynamic range of the assay.

I then performed F2H assays to determine whether a change in LID–GRACE association occurs upon activatory conditions in cultured cells.

3.3 F2H experiments highlight potential glucose-responsive interactions within ChREBP

Having verified the biochemical interaction between LID and GRACE domains in CoIP experiments, I sought to determine whether this interaction can be detected in cells. To study this, I used an F2H assay [132], in which a bait protein is localised to a specific subcellular location and interaction is determined by colocalisation of prey protein to the bait locus (Figure 3.2a, b). GFP-tagged ChREBP, LID, GRACE, or 14-3-3 (a positive control), were co-transfected with Cherry-LacI either alone, or tagged with full-length ChREBP, the LID domain, the GRACE domain, or 14-3-3. These proteins were all expressed without major degradation (Figure 3.2c). The cells were grown in low glucose media and transfected, then glucose or an equivalent volume of water was added. One hour after glucose addition, the cells were fixed (Figure 3.2d). Microscopy was performed with the aid of Master student Giorgi Beroshvili. I wrote a Jython plugin for ImageJ to allow streamlined, unbiased quantification of the images taken. The code for this plugin is included in the Appendix 7.4.1 and the results of these experiments are detailed in the following section. All of the F2H experiments in the following section were performed simultaneously, and values are therefore directly comparable. For ease of discussion however, they are presented separately.

3.3.1 ChREBP colocalises with itself and with 14-3-3 in a glucose-responsive manner

In order to better characterise the regulatory interaction between LID and GRACE, it is important to understand whether one is observing an intramolecular binding or whether the physiologically relevant interaction actually occurs in trans between individual molecules of ChREBP. To determine whether an intermolecular interaction may play a role in the regulation of ChREBP activity, I performed an F2H assay with ChREBP as the prey protein and either Cherry-LacI (negative control), Cherry-LacI-ChREBP, or Cherry-LacI-14-3-3 (positive control) as bait.

What became apparent upon analysis of the representative images was that the negative control showed higher colocalisation, while the positive controls showed lower colocalisation at the array than ideal (based on relative intensity of GFP at the array versus background). Further complications arose when both the GFP and Cherry tagged proteins exhibited glucose-responsive changes both in nuclear localisation and in array brightness. For this reason, standard methods of F2H analysis were at risk of not giving an honest depiction of results. For example, if the bait concentration at the array increases in low glucose media and the relative colocalisation of the prey protein stays the same, quantification of prey at the array alone or normalised to expression level would show a media glucose-dependent change in colocalisation, i.e. a false positive. For the same reasons, false negatives were also possible when analysing the prey signal in isolation. To counteract this, I analysed all images using two complementary quantification techniques subject

to non-overlapping error types, percent distribution, and colocalisation index. For the percent distribution, all of the cells quantified were categorised as either colocalised or not colocalised based on whether the GFP (prey) intensity at the array was greater or less than 1.5 times the median nucleoplasmic prey intensity, respectively. The colocalisation index in contrast, shown in Equation 3.1,

$$\frac{GFParray/GFPnucleoplasm}{Cherryarray/Cherrynucleoplasm} \tag{3.1}$$

takes Cherry (bait) expression and localisation into account by calculating the index of the prey and bait intensities relative to nuclear background. This colocalisation index was further normalised to the negative control.

In a situation of glucose-insensitive colocalisation, where the bait concentration at the array is glucose-sensitive, the percent distribution would give a false positive while the colocalisation index for this situation would give a true negative. Conversely, in a situation of no colocalisation but glucose-sensitive bait concentration at the array, the colocalisation index would yield a false positive while the percent distribution would give a true negative. In this way, these two methods of quantifying my F2H data allow the filtering of false positives. In the upcoming sections, a media glucose-dependent change is only considered to be a true positive if a change is seen in both the colocalisation index and the percent distribution.

In the following figures, the percent distribution and colocalisation index graphs, as well as representative images for each condition are shown. In the colocalisation index graphs, p-values for the difference between low and high glucose are displayed when $p < 0.05$. As can be seen in Figure 3.3, ChREBP shows media glucose-responsive colocalisation to both ChREBP and 14-3-3, which decreases with increasing glucose. This glucose-responsiveness is clearly shown in the representative images, where ChREBP-ChREBP and ChREBP-14-3-3 colocalisation is visible in low glucose, but not in high glucose conditions. However, as can be seen in the percent distribution graphs, most arrays lack colocalisation, even in low glucose conditions and the magnitude of the observed change is small. As shown in the colocalisation index graph, despite the small absolute change in percent distribution, the difference in localisation between low and high glucose is significant when adjusted for bait intensity ($p = 0.0343$ and $p = 0.000835$ for the GFP-ChREBP Cherry-LacI-ChREBP and GFP-ChREBP Cherry-LacI-14-3-3 pairings, respectively). The colocalisation index values of the ChREBP-14-3-3 pairing are less than 1 in both low and high glucose because of the high array to nucleoplasm intensity ratio of the Cherry-LacI-14-3-3 construct relative to that of the Cherry-LacI negative control. Also important to note is the fact that the percent colocalisation of the negative control GFP-ChREBP Cherry-LacI pairing increases slightly with increased glucose. This increase is likely due to increased GFP-ChREBP nuclear localisation in high glucose leading to increased background aggregation at the array.

These data are consistent with previously published work demonstrating that the ChREBP-14-3-3 interaction is strengthened by low energy signals [109, 111]. Furthermore, from these results we can conclude that, in the context of the F2H assay, ChREBP interacts in trans with other ChREBP molecules. Whether these results have physiological relevance remains to be demonstrated.

3.3.2 F2H results indicate that the LID–GRACE interaction is glucose-responsive

To directly determine whether there is a glucose-responsive change in the LID–GRACE interaction in live cells, I performed F2H assays with GFP-LID as the prey protein and Cherry-LacI, Cherry-LacI-LID, Cherry-LacI-GRACE and Cherry-LacI-14-3-3 as bait. As can be seen in Figure 3.4, GFP-LID colocalises in a media glucose-responsive manner with Cherry-LacI-GRACE. This is visible in the representative images, where colocalisation can be seen in low but not in high glucose. As shown in the percent distribution graph, in low sugar conditions, GFP-LID was localised at the Cherry-LacI-GRACE array in 78% of cells, whereas only 58% of cells showed array localisation in high sugar conditions. This change in colocalisation is also evident in the colocalisation index which shows a significant drop in LID–GRACE colocalisation upon increased media glucose ($p = 0.00788$). These results are consistent with my hypothesis of a direct LID–GRACE interaction in basal conditions which is affected by the amount of sugar in growth media and therefore likely by the intracellular levels of glucose metabolites.

The LID–LID and LID–14-3-3 pairings were also tested. The former showed a slight but not significant decrease in colocalisation index with increased media glucose while the opposite trend can be seen in percent distribution (Figure 3.4). This difference in result is caused by a slight increase in the Cherry-LacI-LID array to nucleoplasm ratio in high glucose, driving a corresponding decrease in colocalisation index, while the equivalent GFP array to nucleoplasm ratio is unchanged. I conclude that if there is a physiologically relevant LID–LID interaction, it does not appear to be media glucose-regulated. This conclusion is complicated by data presented in Figures 3.5 and 3.6 which indicate that the Cherry-LacI-LID construct is not capable of glucose-responsive binding at the array.

In a phenotype similar to that of the LID–LID cells, the LID–14-3-3 pairing shows a decrease in colocalisation index with increased media glucose, but no change in percent distribution. This is likely due to the high affinity that LID and 14-3-3 have for each other. This high affinity is reflected in the almost 100% LID–14-3-3 colocalisation in both low and high sugar seen in the percent distribution graphs in Figure 3.4a, and evident in the representative images in Figure 3.4b. If the threshold for GFP colocalisation is doubled from 1.5 to 3 times median nucleoplasmic intensity, percent colocalisation decreases from 55% to 45% between low and high glucose. What is perhaps most interesting about the LID–14-3-3 pairing however, is its dramatically higher colocalisation level relative to ChREBP–14-3-3. This increase of LID colocalisation with 14-3-3 relative to that of ChREBP is consistent with a model in which either: the full-length protein sterically hinders the optimal LID–14-3-3 interaction, or where regions of the full-length protein

directly compete with 14-3-3 for binding to the LID domain. The former explanation may be linked to the stoichiometry of the protein-protein interaction. 14-3-3 forms a homodimer [133] and it is possible that the LID domain is capable of interacting in a 2:2 ratio with each respective Cherry-LacI-14-3-3 dimer, while the much larger full length ChREBP interacts with a 1:2 ratio. This proposed stoichiometry is supported by the observed 1:2 ChREBP:14-3-3 ratio of the purified GSM (Thomas Pysik, unpublished data). Altered stoichiometry between LID–14-3-3 and ChREBP–14-3-3 would drive increased prey intensity at the array for the LID–14-3-3 pairing relative to that of ChREBP–14-3-3.

Although the GFP-LID Cherry-LacI-GRACE pairing showed media glucose responsive colocalisation, such a clear result was not visible for the GFP-GRACE Cherry-LacI-LID pairing (Figure 3.5). While a decrease in the percent distribution of GFP-GRACE Cherry-LacI-LID was visible between low and high glucose, only a small (within error) decrease in the colocalisation index was observed under the same conditions. Thus, while the percent distribution was consistent with media glucose lowering LID–GRACE interactions, (consistent with what was observed for the orthogonal pairing in Figure 3.4a), the colocalisation index was not. Both the GFP and the Cherry array to nucleoplasm ratios of the GFP-GRACE Cherry-LacI-LID pairing decrease in high glucose. This decrease is the result of an increase in nuclear localisation in both of these channels not mirrored by an equivalent change in array intensity. One notable difference in these data relative to the GFP-LID Cherry-LacI-GRACE results in Figure 3.4 is that the overall percent colocalisation is lower. I postulate that this decrease reflects the fact that Cherry-LacI-LID is not as capable of glucose-sensitive binding to the co-transfected prey proteins as Cherry-LacI-GRACE is, possibly because molecular crowding at the array inhibits it forming the necessary conformation. The idea that array-localised Cherry-LacI-LID is incapable of forming the same conformations as GFP-LID is consistent with the GFP-14-3-3 data, discussed below.

3.3.3 Positive controls for glucose-responsive colocalisation imply that LacO-array tagging may decrease the glucose-responsiveness of Cherry-LacI-LID interactions

Although both the GFP-ChREBP Cherry-LacI-14-3-3 and the GFP-LID Cherry-LacI-14-3-3 pairs show media glucose-responsive localisation (Figures 3.3 and 3.4), the same cannot be clearly stated for the reverse pairs (Figure 3.6). A slight decrease is seen between low and high glucose for GFP-14-3-3 and Cherry-LacI-ChREBP, however this change is not significant. For the GFP-14-3-3 and Cherry-LacI-LID pair, no change in percent distribution is seen and, when the threshold is doubled from 1.5 to 3 times median nucleoplasmic intensity, there is actually a slight *increase* between low and high glucose (from 23% to 28%). As noted in the previous section, the same threshold change for the GFP-LID Cherry-LacI-14-3-3 pair highlights a media glucose-induced decrease in colocalisation and showed approximately twice as much colocalisation as the GFP-14-3-3 Cherry-LacI-LID pair. These data are consistent with the idea put forward in the previous section that Cherry-LacI-LID at the LacO array *may* be taking on certain conformations different from those relevant to its protein-protein interactions in vivo. As the LID domain is at the

N-terminus of the protein, directly adjacent to the Cherry-LacI tag, it is possible that a similar phenomenon is also observed in the Cherry-LacI-ChREBP protein, where the LID domain of this construct may be incapable of forming its media glucose-sensitive conformation. If this is true, it would imply that the media glucose-sensitive ChREBP-ChREBP interaction seen in Figure 3.3 may be mediated by LID-independent associations.

3.4 A new F2H assay format indicates that MCR6 is critical for the LID–GRACE interaction

The results of these F2H assays were consistent with the LID–GRACE interaction being media glucose sensitive. As MCR6 contains the predicted G6P binding region and transcriptional activation domain [68], I hypothesised that it may play a role in the regulation of the LID–GRACE interaction. To this end, I generated an MCR6 deletion mutant. Residues 249-260 were removed from this construct, including both the predicted G6P binding motif and the predicted transcriptional activation domain (Figure 3.7a) [68].

As the magnitudes of the differences detected in the previous assay were often small, I changed experimental design to increase assay sensitivity and reduce biological variation. In an attempt to address the former, I changed the order of media treatment from the "negative" approach above, where I measured the reduction in colocalisation upon glucose addition, to a "positive" approach measuring increased colocalisation upon dilution with glucose-free media. Additionally, to reduce biological variation, I generated cell lines stably expressing GFP-LID. A schematic of this new, positive F2H, is provided in Figure 3.7b.

3.4.1 "Positive" F2H assays indicate that MCR6 is critical for the LID–GRACE interaction

To investigate the role of MCR6 in the LID–GRACE interaction, I transiently transfected stable GFP-LID cell lines with Cherry-LacI, Cherry-LacI-GRACE and Cherry-LacI-GRACEΔMCR6 (Figure 3.7b) and performed positive F2H assays on them. The transfection rate for Cherry-LacI-GRACE and Cherry-LacI-GRACEΔMCR6 in this cell line was much higher than in the parental cell line used above, but the expression level was much lower. Accordingly, the bait intensity at the array was also decreased. To accommodate for this, the colocalisation percentage threshold was set to 1.2 times the nucleoplasmic background rather than the 1.5 times used above. The results are presented in Figure 3.8. An increase in LID–GRACE colocalisation can be seen in decreased media glucose, corresponding with the decrease in LID–GRACE colocalisation seen with increased media glucose in Figure 3.4 above. Surprisingly, the deletion of MCR6 returned the percent distribution to within the range of the negative control in both low and high glucose conditions. This abrogation of colocalisation implies that MCR6 is necessary for the LID–GRACE

interaction. These results are particularly interesting in the context of a region predicted to be both necessary for G6P binding and to contain a transcriptional activation domain [68].

3.5 FRET-FLIM confirms that the GSM conformation is glucose-responsive in live cells

Having demonstrated that the LID and GRACE domains interact and that this interaction appears to be regulated by media glucose conditions, I hypothesised that the LID and GRACE domains in the intact GSM open up upon increased media glucose in live cells. To test this hypothesis, I generated GFP-GSM and GFP-GSM-Cherry FRET constructs and assayed them using fluorescence lifetime microscopy in live cells. To ensure maximal glucose-responsiveness in this assay, I used the highly glucose-responsive INS1-derived 832/13 pancreatic β-cell line [134] (a kind gift from Professor Christopher Newgard at Duke University) used by both the Chan and Towle labs to characterise ChREBP [69, 112, 102, 70, 125].

The experimental outline is shown in Figure 3.9a. In brief, cells were transfected with the indicated constructs, the media was changed and FRET-FLIM of the live cells was begun at least 1 hour after media change. FLIM relies on the fact that a FRET interaction causes a shortened fluorescence lifetime of the donor protein. In my experimental setup, I used time-correlated single-photon counting (TCSPC) to record the fluorescence decay curve of the donor GFP in GFP-GSM and GFP-GSM-Cherry transfected cells under low and high glucose conditions. From these decay curves, one or two fluorescence lifetimes were extracted, as detected. A cartoon of how the conformation of the GSM would affect the GFP lifetime of each construct is shown in Figure 3.9b. All constructs tested were expressed without major degradation, as can be seen in Figure 3.9c. Figure 3.9d shows representative images from all four conditions demonstrating that the subcellular localisation and overall cell health appear unchanged between the negative control and the test condition. Having ensured that the constructs were expressed and that the cells appeared healthy under all conditions tested, I quantified the data using the amplitude-weighted lifetime.

In the case of the two-lifetime decay curves observed for GFP-GSM-Cherry, the amplitude-weighted lifetime shows the average lifetime of the donor molecule for a given condition, adjusted for the number of photons occupying the quenched (shorter lifetime) and unquenched (longer lifetime) states, respectively. For the one-lifetime decay curve of the GFP-GSM negative control, the amplitude-weighted lifetime is equal to the unadjusted lifetime of that molecule. The lifetime for the GFP-GSM negative control was 2.23 ns in both low and high glucose (2.23 95% CI 2.21-2.24 ns in low glucose and 2.23 95% CI 2.20 - 2.26 ns in high glucose).

Figure 3.9e shows that a greater fraction of the GSM is open in high glucose than in low glucose conditions. The increased amplitude-weighted lifetime seen upon increased glucose (1.86 95% CI 1.85-1.87 ns in low glucose and 1.91 95% CI 1.90 - 1.92 ns in high glucose) is consistent with my hypothesis that the LID and GRACE domains in the intact GSM dissociate upon increased media

glucose in live cells and is consistent with my F2H results showing that the LID and GRACE colocalisation at the array decreases upon increased media glucose (Figure 3.4).

3.6 Discussion and future directions

In this chapter I have shown that the LID and GRACE domains interact, that this interaction appears to be regulated by media glucose and that this regulation corresponds to an increased opening of the GSM in live cells under high glucose conditions. These results provide a potential mechanistic explanation for the functional activation observed previously in high glucose conditions. Additionally, the results in this chapter imply that, even in activatory conditions, the majority of the GSM remains in the closed, presumably inactive state. Finally, they indicate that GSM opening may be independent of G6P binding. CoIPs did not detect a change in LID–GRACE interaction in the presence of G6P and the GSM without MCRs 5 and 6 is still able to bind G6P in vitro with a similar K_d relative to the full length construct (Nina Heppner Master's Thesis [129]), whereas my F2H results showed that MCR6 deletion is sufficient to abrogate LID–GRACE colocalisation. Taken together, my results imply a MCR6-dependent mechanism of media glucose-driven GSM opening.

My F2H results imply a decrease in the association between ChREBP and 14-3-3 with increased media glucose, consistent with previously published work showing that the ChREBP 14-3-3 interaction is promoted by low energy signals [109, 111]. It is uncertain how these results fit with the published requirement of 14-3-3 for ChREBP activity [112]. It is possible that a secondary 14-3-3 binding site is involved, as put forward by the Uyeda group [135], or that the stimulated activation seen by the Chan lab upon increased 14-3-3 is the result of increased overall stabilisation of ChREBP peptides, rather than a direct role in protein activation.

Further validation of these data is required to confirm these initial indications. Firstly, FRET-FLIM in live cells with the MCR6 deletion mutant would assay whether the loss of LID–GRACE interaction seen by F2H assay corresponds to a conformational change in the GSM as a whole. Second, in vitro FRET-FLIM of the purified GSM and GSMΔMCR6 would assay whether the G6P independent LID–GRACE association seen by CoIP is matched by a lack of detectable change in the GSM as a whole and whether the phenotype seen in the MCR6 deletion mutant in live cells is due to an inherent property of the peptide itself, or due to a change in protein-protein interactions.

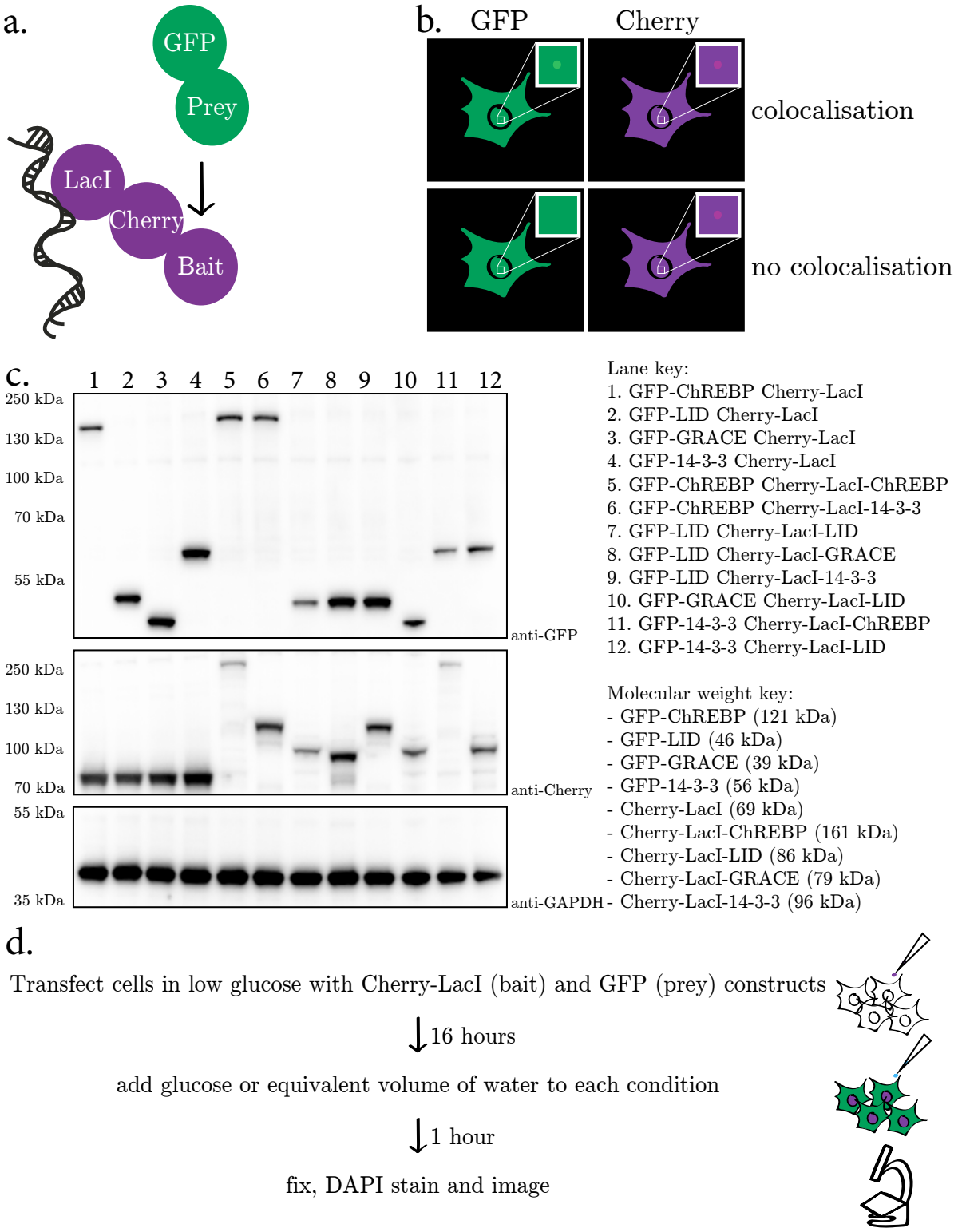


Figure 3.2: **Overview of F2H assays performed**
a. In an F2H assay the bait protein is localised to a defined subcellular location and the colocalisation of the prey protein is quantified. In these assays, a U2OS cell line containing 256 copies of a LacO array was used to target the bait protein to a single spot in the nucleus. If the prey protein is found at this location at a higher concentration than the surrounding nucleoplasm, the two proteins colocalise. **b.** An idealized cartoon of what a negative (no colocalisation) and a positive (colocalisation) result look like. **c.** Western blots showing that the F2H constructs are all expressed at the correct size. **d.** Experimental outline for the negative F2H detailed in the following section. In short, cells were transfected in low glucose, then changed to high glucose or maintained at low glucose prior to fixation and imaging.

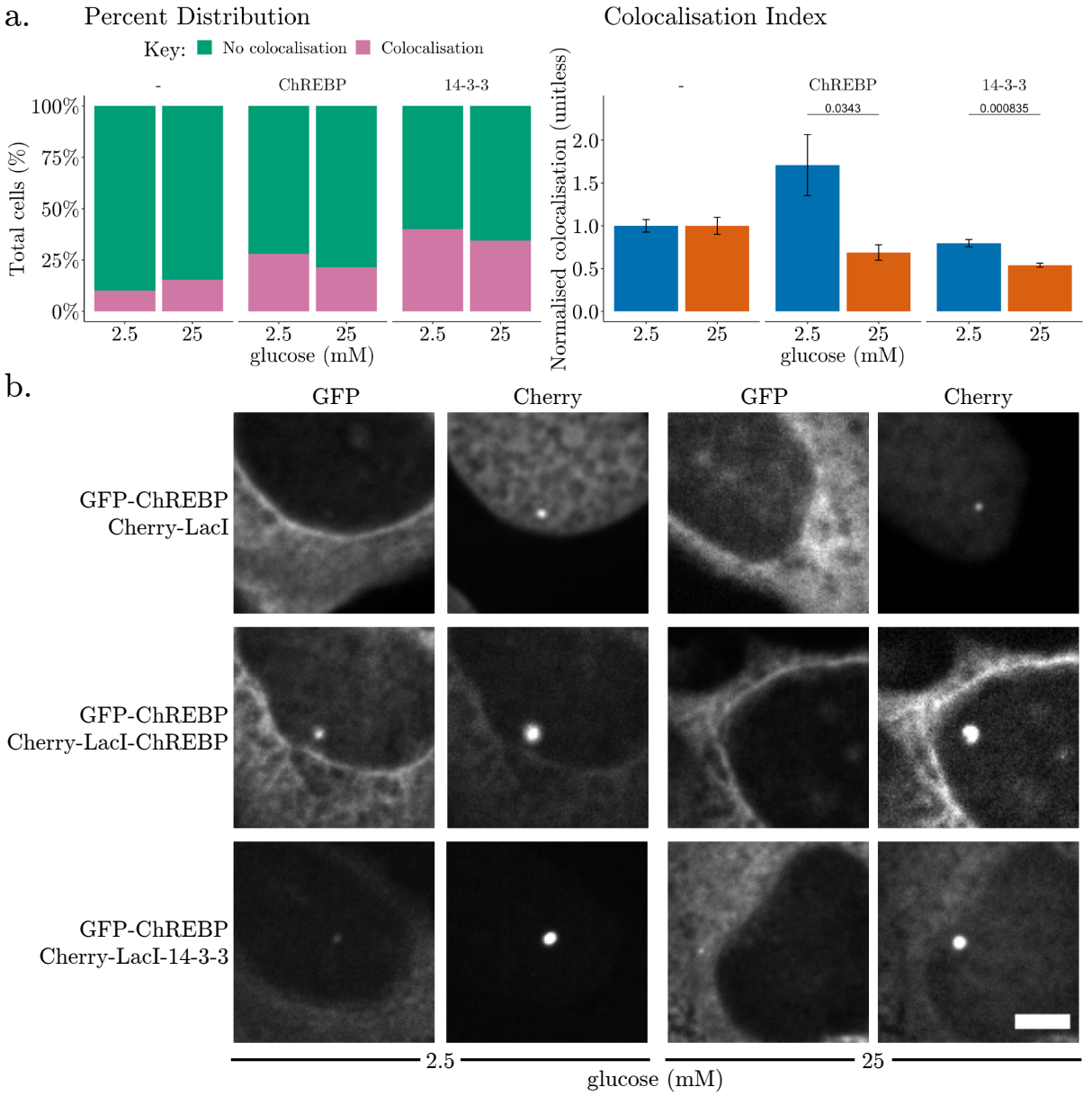


Figure 3.3: **ChREBP and 14-3-3 bait proteins appear to interact in a glucose-responsive manner with ChREBP as prey**
a. The graph on the left shows the total percent distribution of all cells in all 6 biological replicates as either array localised (pink) or nucleoplasm localised (green). The graph on the right shows the colocalisation index ($\frac{GFParray/GFPnucleoplasm}{Cherryarray/Cherrynucleoplasm}$) of each condition with the negative control set to 1. Error bars show SEM for 6 biological replicates. p-values were calculated using the two-tailed t-test and are shown when $p < 0.05$. The total number of cells for each condition represented in the percent distribution graphs are as follows, where numbers indicate the cell counts in low and high glucose, respectively: Cherry-LacI 169, 143; Cherry-LacI-ChREBP 93, 98; Cherry-LacI-14-3-3 160, 194. **b.** Representative images of each condition. Scale bar is 5 µm.

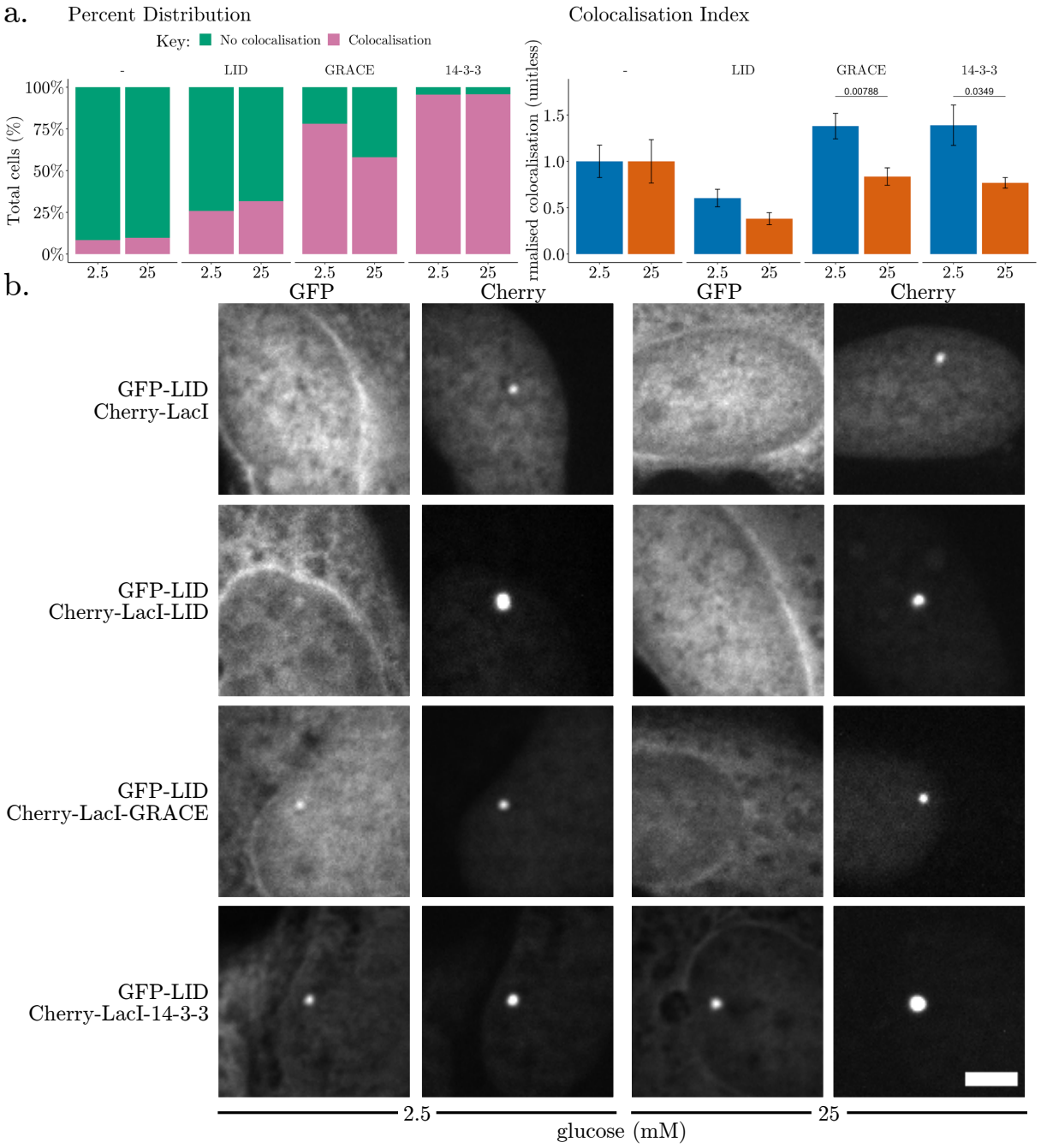


Figure 3.4: **GFP-LID interacts in a glucose-responsive manner with Cherry-LacI-GRACE**

There is not any clear glucose-responsive change in colocalisation between low and high glucose for GFP-LID with Cherry-LacI-LID. The GFP-LID Cherry-LacI-GRACE transfected cells show glucose-responsive changes in colocalisation visible in both their percent distribution, and their colocalisation index. The GFP-LID Cherry-LacI-14-3-3 transfected cells showed strong and consistent colocalisation in both low and high glucose media. **a.** The graph on the left shows the total percent distribution of all cells in all 6 biological replicates as either array localised (pink) or nucleoplasm localised (green). The graph on the right shows the colocalisation index ($\frac{GFParray/GFPnucleoplasm}{Cherryarray/Cherrynucleoplasm}$) of each condition with the negative control set to 1. Error bars show SEM for 6 biological replicates. p-values were calculated using the two-tailed t-test and are shown when $p < 0.05$. The total number of cells for each condition represented in the percent distribution graphs are as follows, where numbers indicate the cell counts in low and high glucose, respectively: Cherry-LacI 83, 92; Cherry-LacI-LID 182, 164; Cherry-LacI-GRACE 178, 150; Cherry-LacI-14-3-3 201, 165. **b.** Representative images of each condition. Scale bar is 5 µm.

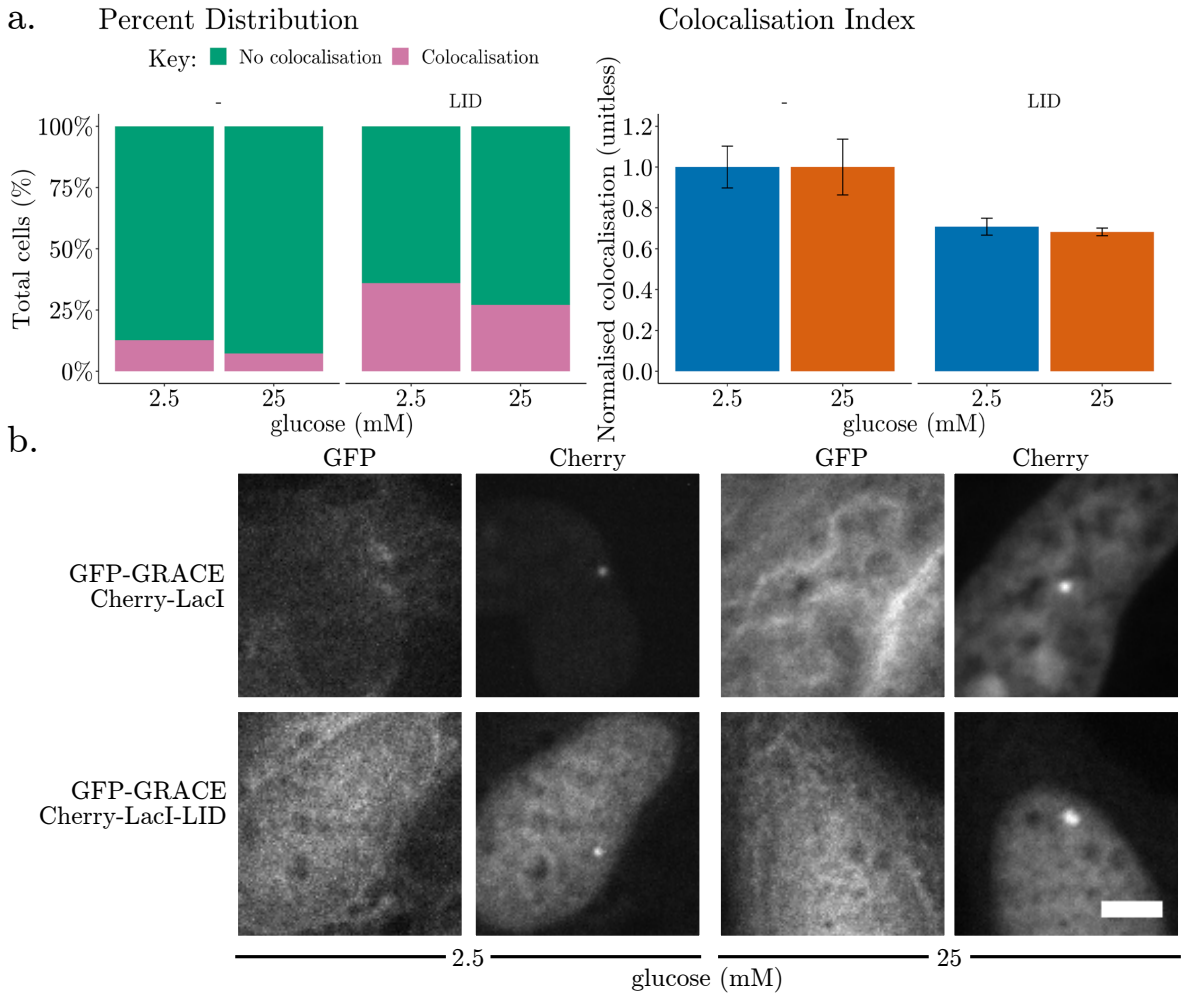


Figure 3.5: **No unambiguous glucose-responsive change is seen between GFP-GRACE and Cherry-LacI-LID**

a. The graph on the left shows the total percent distribution of all cells in all 6 biological replicates as either array localised (pink) or nucleoplasm localised (green). The graph on the right shows the colocalisation index ($\frac{GFParray/GFPnucleoplasm}{Cherryarray/Cherrynucleoplasm}$) of each condition with the negative control set to 1. Error bars show SEM for 6 biological replicates. No p-values of less than 0.05 were found between low and high glucose for the given constructs (two-tailed t-test). The total number of cells for each condition represented in the percent distribution graphs are as follows, where numbers indicate the cell counts in low and high glucose, respectively: Cherry-LacI 134, 111; Cherry-LacI-LID 161, 192. **b.** Representative images of each condition. Scale bar is 5 µm.

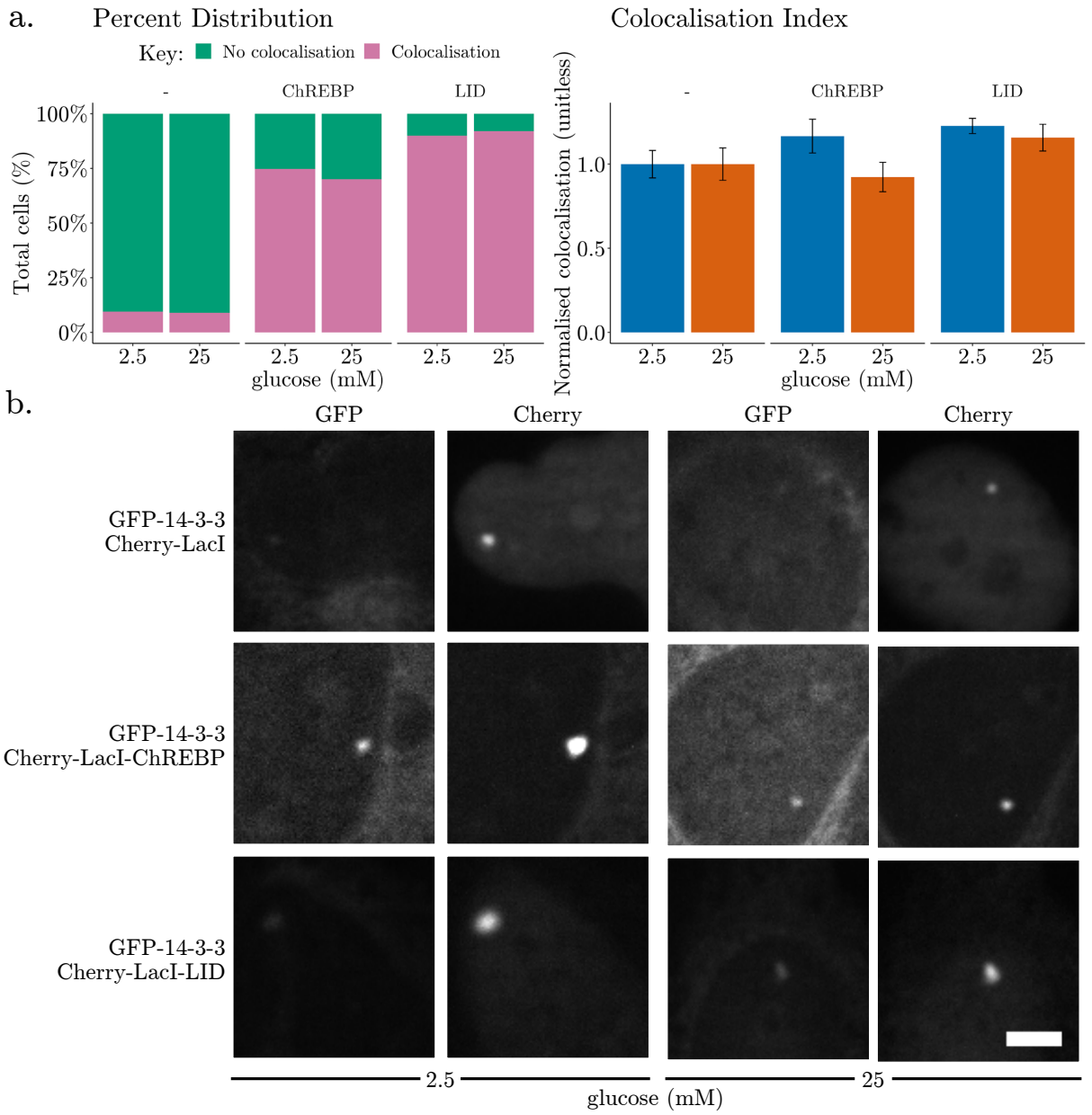


Figure 3.6: **No significant change in colocalisation is seen between GFP-14-3-3 and Cherry-LacI-ChREBP or Cherry-LacI-LID, respectively**
A slight but non-significant decrease in colocalisation is seen between GFP-14-3-3 and Cherry-LacI-ChREBP with increased media glucose. No unambiguous change in colocalisation between GFP-14-3-3 and Cherry-LacI-LID is observed. **a.** The graph on the left shows the total percent distribution of all cells in all 6 biological replicates as either array localised (pink) or nucleoplasm localised (green). The graph on the right shows the colocalisation index ($\frac{GFP array/GFP nucleoplasm}{Cherry array/Cherry nucleoplasm}$) of each condition with the negative control set to 1. Error bars show SEM for 6 biological replicates. No p-values of less than 0.05 were found between low and high glucose for the given constructs (two-tailed t-test). The total number of cells for each condition represented in the percent distribution graphs are as follows, where numbers indicate the cell counts in low and high glucose, respectively: Cherry-LacI 169, 145; Cherry-LacI-ChREBP 170, 127; Cherry-LacI-LID 177, 201. **b.** Representative images of each condition. Scale bar is 5 µm.

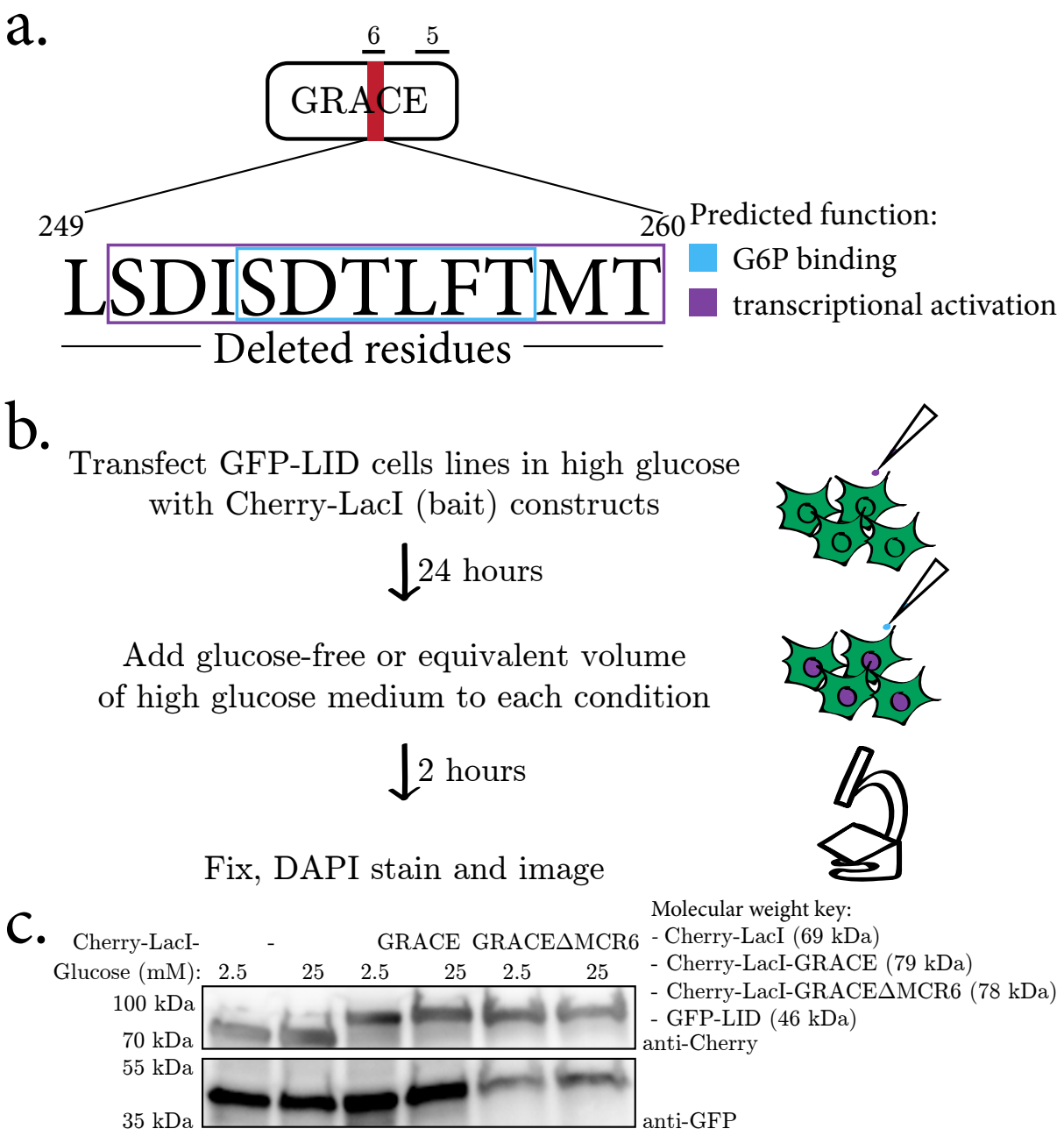


Figure 3.7: **Overview of "positive" F2H assay performed**
a. A diagram showing the location and predicted functions of the deleted MCR6 residues 249-260. **b.** Experimental outline of the positive F2H detailed in Section 3.4. **c.** The constructs transfected into the GFP-LID cell lines are stably expressed at all glucose levels assayed.

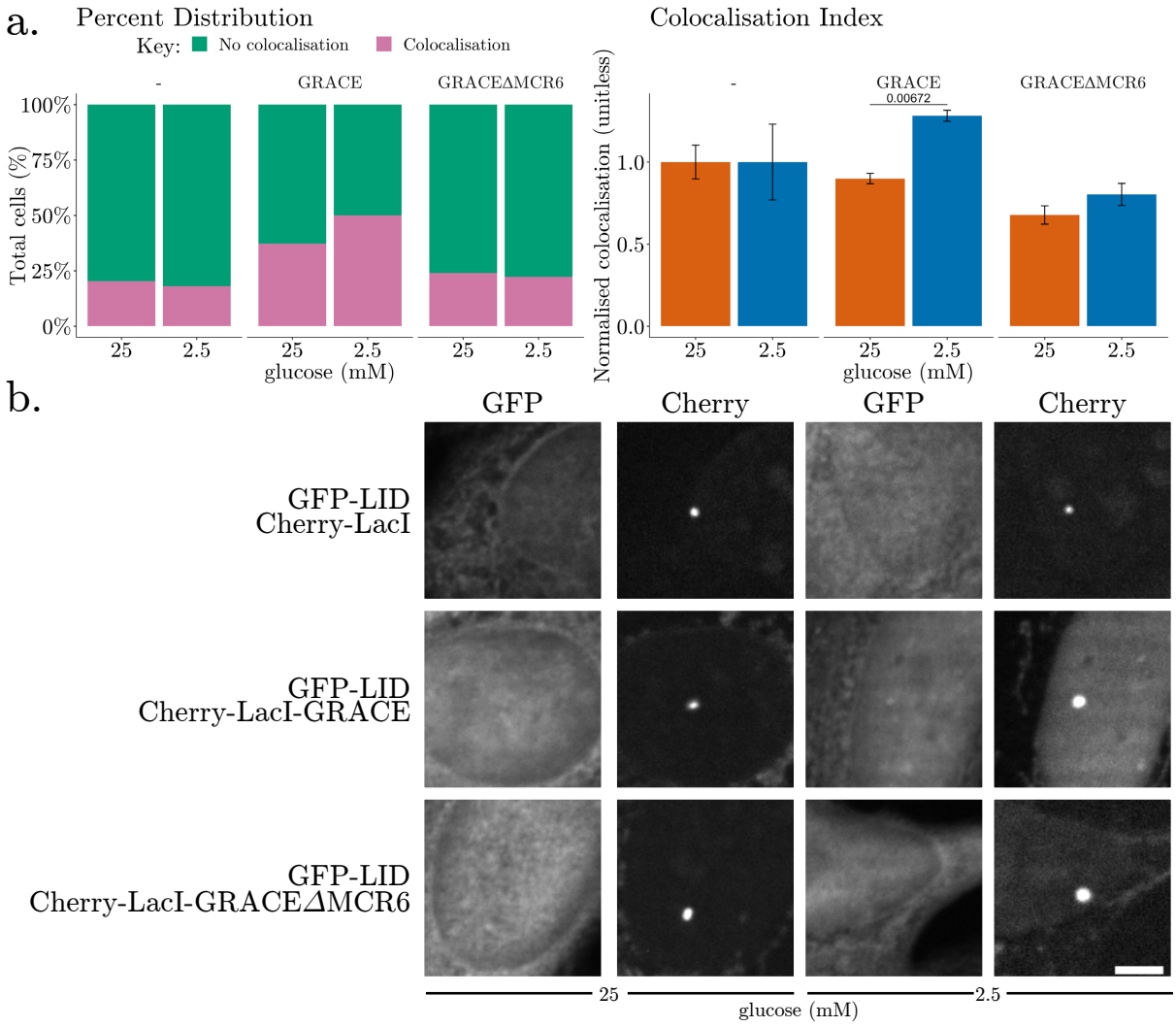


Figure 3.8: **MCR6 deletion prevents LID–GRACE colocalisation**
The Cherry-LacI-GRACE transfected cells show an increase in colocalisation upon decreased glucose. The Cherry-LacI-GRACEΔMCR6 transfected cells show an equivalent percent colocalisation to the Cherry-LacI negative control. **a.** The graph on the left shows the total percent distribution of all cells in all 3 biological replicates as either array localised (pink) or nucleoplasm localised (green). The graph on the right shows the colocalisation index ($\frac{GFParray/GFPnucleoplasm}{Cherryarray/Cherrynucleoplasm}$) of each condition with the negative control set to 1. Error bars show SEM for 3 biological replicates. p-values were calculated using the two-tailed t-test and are shown when $p < 0.05$. The total number of cells for each condition represented in the percent distribution graphs are as follows, where numbers indicate the cell counts in high and low glucose, respectively: Cherry-LacI 324, 183; Cherry-LacI-GRACE 43, 28; Cherry-LacI-GRACEΔMCR6 225, 202. **b.** Representative images of each condition. Scale bar is 5 µm.

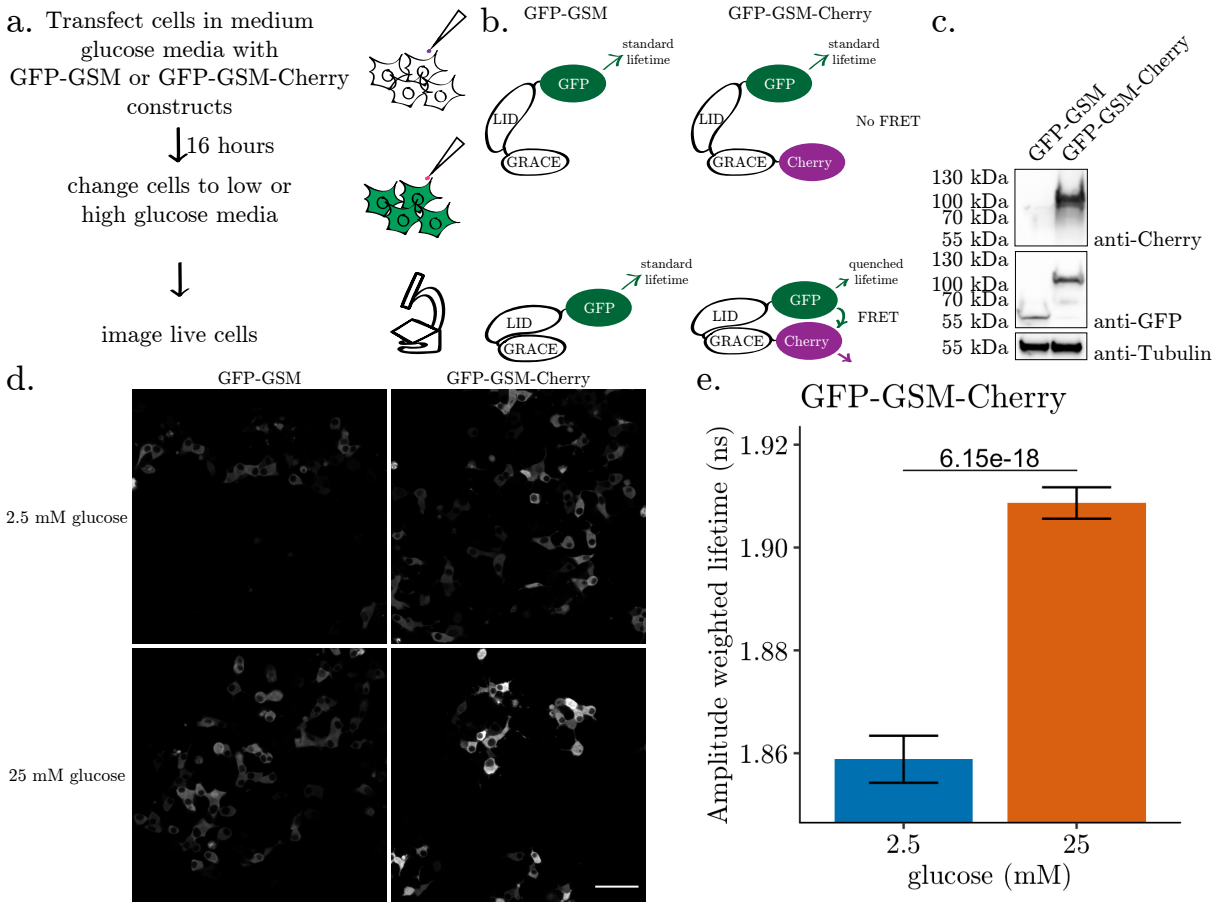


Figure 3.9: **FRET-FLIM microscopy shows a greater fraction of the GSM is in an open state in high glucose media**
a. Experimental outline of FRET-FLIM assay. Imaging was begun at least 1 hour after media change. **b.** Cartoon of the effect that an open or closed conformation would have on the detected lifetime of GFP in a given construct. **c.** The constructs used in this assay were both expressed without major degradation. GFP-GSM and GFP-GSM-Cherry have predicted molecular weights of 58 kDa and 85 kDa, respectively. **d.** An example field of view is shown for each condition assayed. No clear difference in cell health or subcellular localisation can be seen between the negative control and the test condition. Scale bar is 50 µm. **e.** Amplitude-weighted lifetime for the GFP-GSM-Cherry construct. Error bars are SEM for all cells imaged across 3 separate experiments on 3 separate days. p-values were calculated using the two-tailed t-test. The total number of cells for each condition are as follows, where numbers indicate the cell counts in low and high glucose, respectively: GFP-GSM 234, 240; GFP-GSM-Cherry 223, 245.

Chapter 4

The role of MCRs 3, 4 and 6 in the regulation of ChREBP

4.1 Introduction

HDX-MS showed that the GSM of mouse ChREBP has altered solvent accessibility after AMP and G6P binding (Thomas Pysik in collaboration with Kaspar Rand and Johanna Veiga, unpublished data). Both AMP and G6P binding drove changes in the solvent accessibility of the α2-helix of MCR3, while G6P additionally drove unique changes in the MCR3 region upstream of the α2-helix, as well as in the NLS of MCR4 and in MCR6. Limited proteolysis work by Thomas further identified that the C-terminus of the GSM, including MCR6, is selectively exposed in the presence of G6P but not AMP, highlighting this region as a point of interest. The discovery that G6P binding drives specific changes in the solvent exposure of the NLS indicated a role for G6P binding in the direct regulation of ChREBP nuclear localisation. More broadly, the fact that MCRs 3, 4 and 6 are directly affected by G6P binding, coupled with previously published data indicating that G6P is *the* glucose metabolite required for ChREBP activation [102], implies a key role for these regions in the activation of ChREBP. To investigate the role media glucose has on the subcellular localisation of ChREBP, I used FRAP and photoactivation microscopic techniques to study the import and export of mouse ChREBP in real time. Additionally, to dissect the link between G6P binding, ChREBP subcellular localisation and protein activation, I complemented the live cell experiments above with luciferase reporter and subcellular localisation assays in MCR3, 4 and 6 point mutants of the mouse protein.

4.2 FRAP and photoactivation experiments show that nuclear ChREBP is rapidly exchanged

ChREBP shows increased nuclear localisation in high glucose conditions [69]. This sugar sensitive nuclear localisation has been linked to a release of AMP driven inhibition of ChREBP binding to importinα [111]. However, the discovery that G6P binding causes a change in the solvent exposure of the NLS indicates that high sugar signals may also play a role in this regulation. To characterise the action of media glucose on ChREBP in live cells, I, in collaboration with Dr. Imke Mandemaker, used live-cell imaging to study the rate of nuclear import in low and high glucose conditions.

4.2.1 FRAP shows that nuclear ChREBP is rapidly imported in both low and high glucose conditions

To measure the rate of ChREBP nuclear import, I performed FRAP experiments using full-length GFP-ChREBP in low and high glucose conditions. The experimental setup paralleled that in Figure 3.9a, although the cells were transfected with GFP-ChREBP rather than FRET constructs. Whole nuclei were bleached and the rate of return of GFP signal to the nucleus was measured (Figure 4.1a). Imaging was performed by Dr. Imke Mandemaker.

As both import and export are likely to occur in the steady-state cell under normal growth conditions, a nuclear export inhibitor was necessary to observe the role of nuclear import in isolation. To this end, cells were treated with LMB, which binds covalently to the CRM1 nuclear export protein and blocks its activity [136]. The results are shown in Figure 4.1b and c which show the recovery curves for the indicated conditions, normalised to the nuclear prebleach intensity and to the sum of nuclear and cytoplasmic prebleach intensities, respectively.

Modelling of these data indicate that the nuclear import of ChREBP is regulated. The total flux, J, of any given protein, X, across the nuclear membrane can be represented by Equation 4.1, where J^p represents flux due to passive diffusion according to Fick's law applied to porous membranes [137], and J^n and J^c represent active transport to and from the nucleus, respectively.

$$J = J^p + J^n + J^c \tag{4.1}$$

Assuming that: 1. the fluorescent intensity of protein X is directly proportional to its concentration, 2. the rate of nuclear import of protein X is constant, 3. the cytoplasmic fraction of X not bound to import machinery remains constant, 4. the rate of nuclear import of protein X is linearly related to the cytoplasmic concentration of X and 5. the rate of nuclear import of protein X is linearly related to the cytoplasmic fraction of X not bound to import machinery; fluorescent

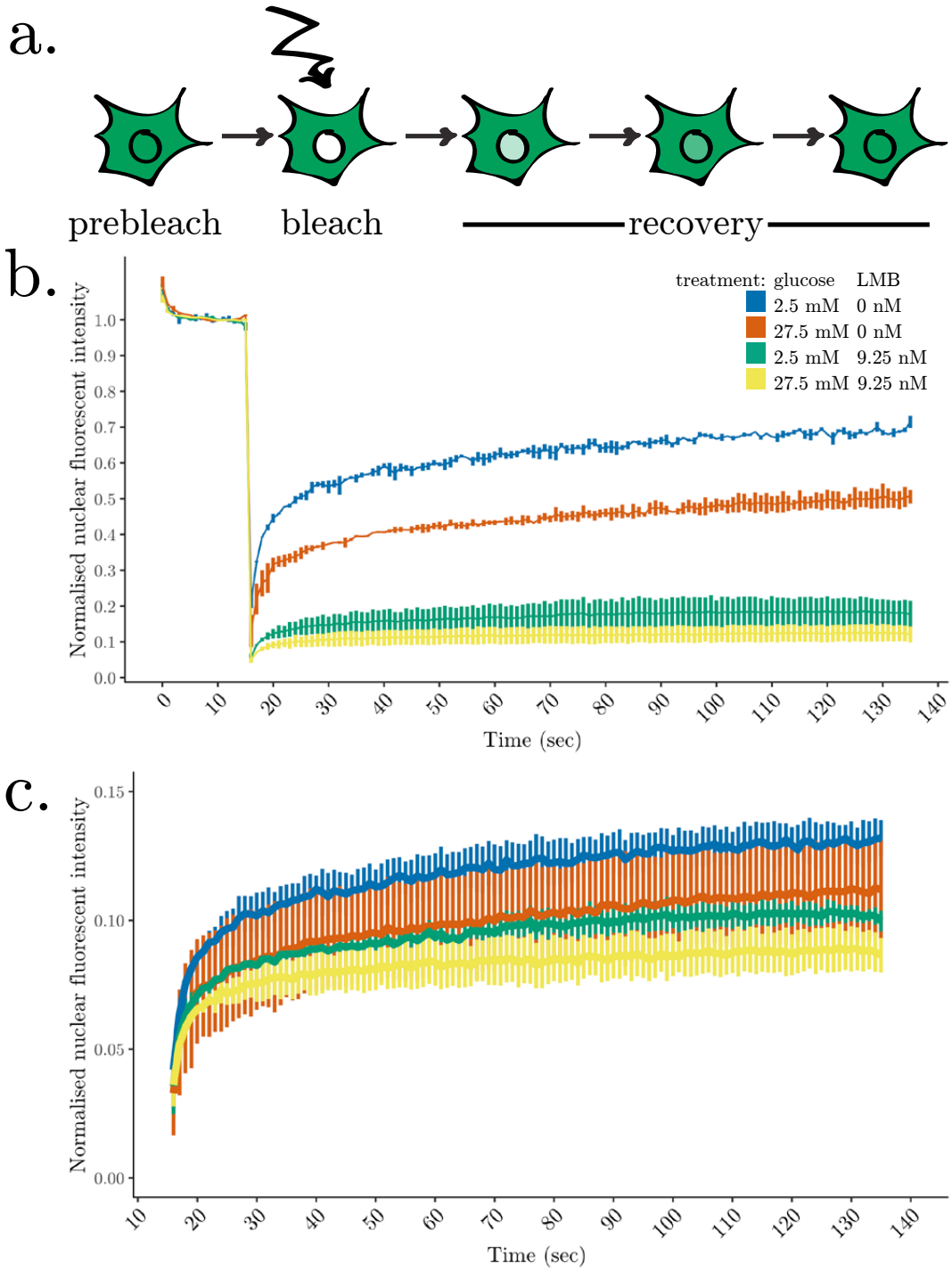


Figure 4.1: **Nuclear ChREBP is rapidly imported in both low and high glucose conditions**
a. A cartoon of the assay procedure. Fluorescent intensity is measured before and after bleaching. **b.** Nuclear fluorescent intensity normalised to nuclear prebleach intensity. **c.** Nuclear fluorescent intensity normalised to the sum of prebleach nuclear and cytosolic intensities. Prebleach intensities are cropped for clarity. Error bars = SEM. N = 2.

intensity in the nucleus after complete nuclear bleaching will follow the mono-exponential recovery curve shown in Equation 4.2, where F^0 and F^∞ represent the fluorescence at time of bleaching and the value the fluorescence asymptotically approaches during recovery, respectively, t represents

the time after bleaching and τ is a constant. Equation 4.2 is taken from the work of Cardarelli et. al, where it is derived in full [138].

$$F(t) = F^{\infty} + (F^{0} - F^{\infty})e^{\frac{-t}{\tau}} \tag{4.2}$$

The recovery curves from experiments with proteins such as GFP alone, GFP-NLS and JNK1 (whose nuclear import, like that of ChREBP, requires importinα) fit well to Equation 4.2, demonstrating the equation's relevance in vivo [138, 139, 140]. However, the data in Figure 4.1, do not fit the mono-exponential function shown in Equation 4.2 (see Figure 7.1 in the Appendix), implying that GFP-ChREBP experiences a type of regulation that causes one of the above assumptions to be untrue. From the similarity in the recovery curves Figure 4.1c, where nuclear intensity is normalised to total cellular prebleach intensity (as opposed to nuclear prebleach intensity alone), between LMB treated and untreated cells who have low and high cytoplasmic concentrations of ChREBP, respectively, it appears likely that assumption 4 (import rate is directly proportional to cytoplasmic concentration) may be untrue for GFP-ChREBP. The in vivo significance of this putative concentration independence remains to be determined.

Surprisingly, the results show that ChREBP is rapidly imported in both low and high glucose conditions. In fact, as can be seen from Figure 4.1b, over 50% of the steady-state nuclear fraction of ChREBP is exchanged every minute in low glucose cells. Similarly, over 40% of the steady-state nuclear fraction of ChREBP is exchanged every minute in high glucose cells. As the total amount of nuclear ChREBP is decreased in low glucose relative to high glucose conditions, this does not necessarily mean that the absolute rate of ChREBP import is higher in low glucose media. To dissociate the observed rate of recovery from differences in initial nuclear localisation, I normalised the recovery curves to the sum of the nuclear and cytoplasmic prebleach intensities. Surprisingly, the recovery of low glucose GFP-ChREBP remained slightly above that of high glucose GFP-ChREBP and this relative difference was emphasized after LMB-driven inhibition of nuclear export (Figure 4.1c), implying that the difference observed represents a real difference in the rate of nuclear import. Given the surprising nature of these data, it is relevant to note that in my hands ChREBP shows decreased nuclear localisation in low glucose (discussed below) and that similar FRAP experiments performed INS1 cells also observed rapid nuclear recovery of ChREBP in low and high glucose conditions [141].

To reconcile the import results with overall ChREBP localisation levels, I focused on possible oddities in the final percent recovery. The trace seen in Figure 4.1b plateaus at approximately 70% prebleach intensity in low glucose and approximately 50% prebleach intensity in high glucose. As a sizeable portion of fluorescent protein in the cell was bleached, these numbers would never return to 100%. However, it is possible that a portion of bleached nuclear ChREBP is comparatively immobile and remains in the nucleus for the duration of the assay, decreasing the observed percent recovery. This is not unlikely for a transcription factor under activatory conditions, and differential chromatin retention would explain how ChREBP appears to simultaneously have lower

steady-state nuclear localisation and higher nuclear import rates under low glucose conditions after LMB treatment, as seen in Figure 4.1c.

4.2.2 FRAP finds no major difference in ChREBP nucleoplasmic mobility between low and high glucose conditions

To assess whether the difference in subcellular localisation of ChREBP between low and high media glucose conditions was caused by nuclear retention, FRAP was performed on a linear portion of the nucleus and the rate of recovery measured, as shown in Figure 4.2a. Imaging was performed by Dr. Imke Mandemaker. As evident in Figure 4.2b, no clear difference can be seen between the conditions tested. It appears unlikely that ChREBP subcellular localisation differences between low and high glucose are caused by altered nuclear retention.

4.2.3 Photoactivation of nuclear ChREBP is unable to detect a marked difference in the rate of nuclear export between low and high glucose conditions

As it appeared that the difference in subcellular localisation of ChREBP between low and high media glucose conditions was caused by neither alterations in nuclear import, nor chromatin retention, it seemed likely that nuclear export was the causative agent. To test whether ChREBP nuclear export rates differed between low and high glucose conditions, I tagged full-length ChREBP with a photoactivatable GFP. Photoactivation experiments were then performed in which nuclear ChREBP was activated and the GFP intensity in the nucleus was measured (Figure 4.3a). Imaging was performed by Dr. Imke Mandemaker. The results are shown in Figure 4.3b with the pre- and post-activation intensities set to 0 and 1, respectively. These results fit well with the data on nuclear import, showing that approximately 50% of nuclear ChREBP is exported from the nucleus every minute. Although the high glucose export curve seen in Figure 4.3 may indicate a slower export rate relative to initial nuclear occupancy, this marginal difference is likely to disappear when the higher starting nuclear occupancy of high glucose ChREBP is taken into account. Unfortunately, the rates of ChREBP nuclear import and export are so rapid that we were unable to capture the linear range where export can be measured without background interference from import.

These data do not exclude the possibility that glucose mediated differences in ChREBP nuclear localisation are caused by differential export rates. However, in combination with the fact that GFP-ChREBP nuclear intensities recovered to the same fraction of total cellular intensity in both low and high glucose in the FRAP experiment above (Figure 4.1c), these data indicate that the difference in subcellular localisation is caused by a phenomenon which is not observable in the conditions or timescale in which this assay was performed. It remains possible that the observed difference in ChREBP subcellular localisation between low and high glucose is the result of a

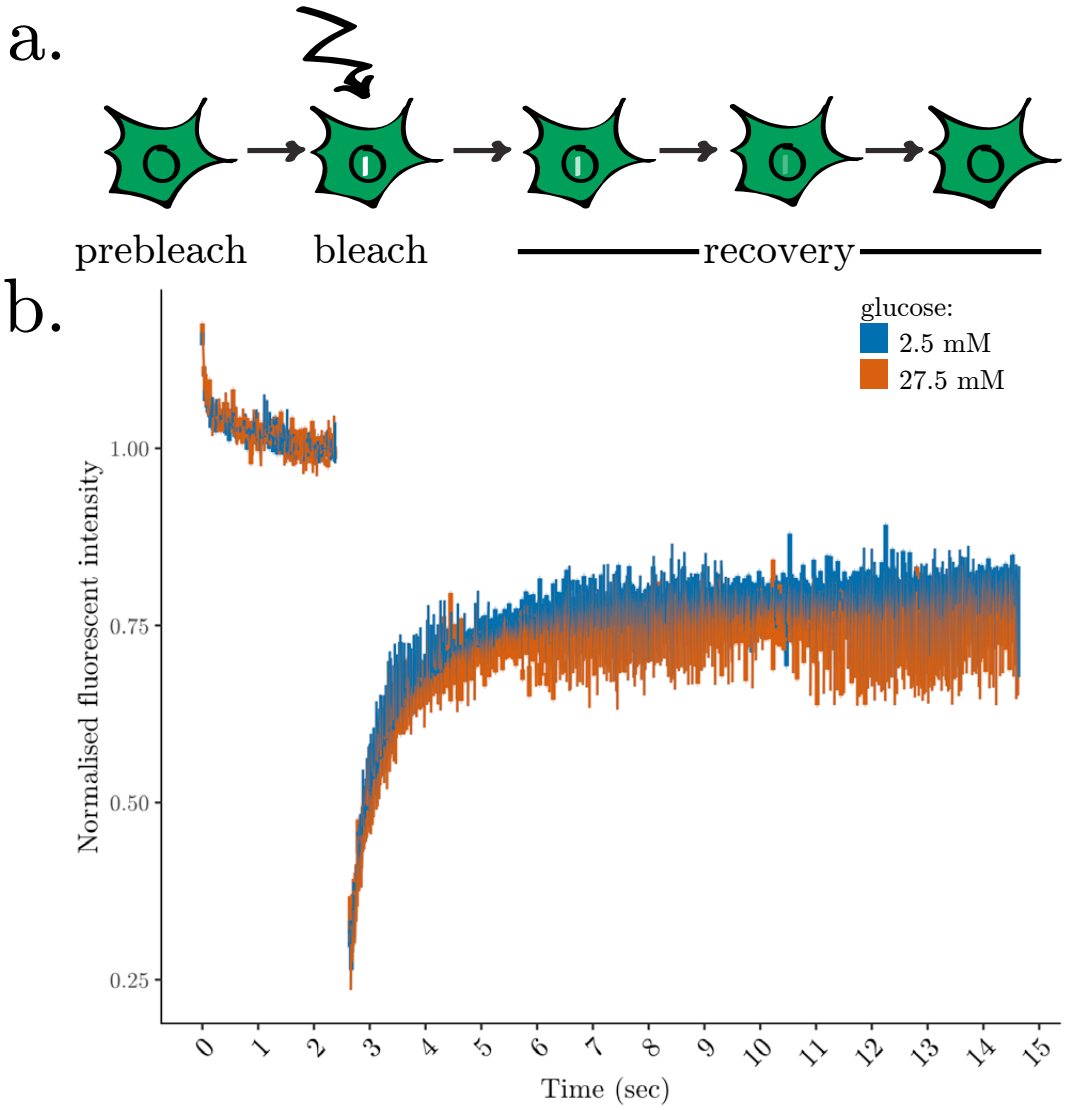


Figure 4.2: **No clear difference in ChREBP nuclear mobility was found between low and high glucose conditions**
a. A cartoon of the assay procedure. Fluorescent intensity is measured before and after bleaching. **b.** Fluorescent intensity in the bleached area normalised to the prebleach intensity in that area. Error bars = SEM of individual cells. N = 1 with 12 low glucose cells and 11 high glucose cells in total. Owing to the high cell-cell variation and lack of visible difference between the conditions tested this assay was performed only once.

rapid acute change in the nuclear import to export ratio upon altered glucose levels. This would be consistent with the Ca^{2+}-driven model of ChREBP activation put forward by the Leclerc group [100] and discussed in Section 1.5.1. However, studies in 832/13 cells found a difference in nuclear import rates at steady-state glucose levels observed at 15-minute intervals after LMB addition [70], demonstrating that media glucose driven differences in nuclear import rates are present in the steady-state cell. It is therefore likely that the mechanism which drives differential subcellular localisation of ChREBP occurs in the steady-state but was not observable with the time frame of the assay, either because the assay is too slow (to capture differential export rates of photoactivated pPAGFP-ChREBP before any measurable import occurs, for example) or because

the sugar regulated kinetics occur on a slower timescale than what was measured here.

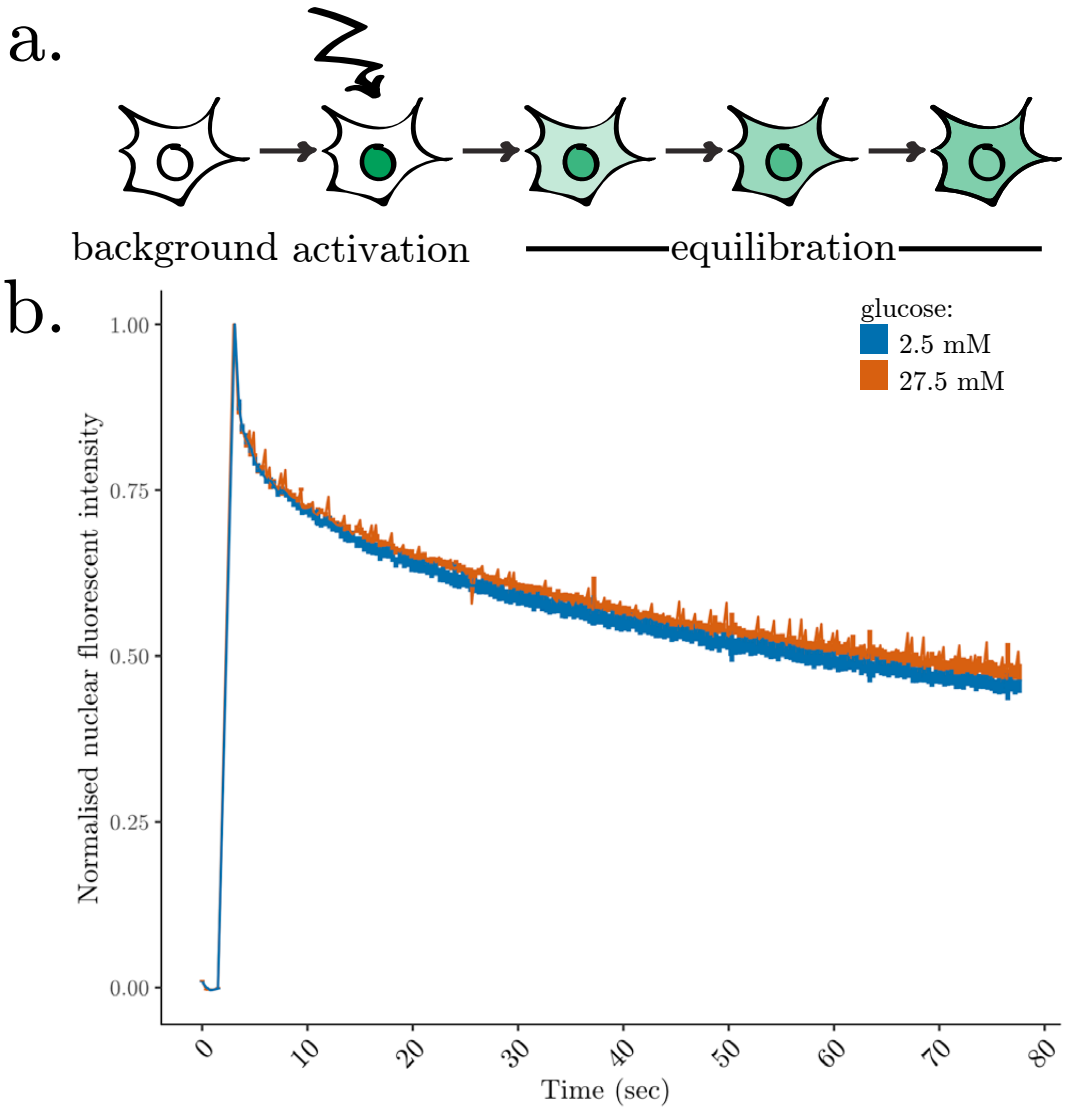


Figure 4.3: **No clear difference in ChREBP nuclear export was found between low and high glucose conditions**
a. A cartoon of the assay procedure. Fluorescent intensity is measured before and after photoactivation. **b.** Fluorescent intensity in the nucleus with the pre- and post-activation intensities set to 0 and 1, respectively. Error bars = SEM of individual cells. N = 1 with 14 low and 15 high glucose cells in total. Owing to our inability to capture the linear range where the rate of ChREBP nuclear export can be directly measured, this assay was only repeated once.

4.3 The role of MCRs 3, 4 and 6 in ChREBP activation and nuclear localisation

Having demonstrated that the nuclear fraction of ChREBP is rapidly exchanged, I sought to determine the role that the regions of interest from the HDX-MS data have on ChREBP activation and nuclear localisation. To study this, I generated targeted point mutations of conserved residues

in MCRs 3, 4 and 6 of the GSM and assayed the effect these had on glucose-responsive activity and nuclear localisation.

Figure 4.4a depicts the relative locations of the MCRs within the GSM and their known unique functions. MCR3 is required for 14-3-3 interaction and contains the binding site for AMP. As discussed in the introduction, it appears to have dual functions, being simultaneously necessary for GSM repression and activity. Similarly, MCR4, which contains the NLS, appears to be required both for ChREBP nuclear localisation and low glucose repression. Finally, MCR6 has been bioinformatically predicted to directly bind G6P and to interact with transcriptional activators such as p300 [68], while my own data imply it is necessary for the interaction between the LID and GRACE (Figure 3.8).

Previous experiments done on these regions had largely been at the level of whole MCR deletion, or mutation of the 14-3-3 binding region. An understanding of function at sub-MCR resolution required a more targeted approach. I sought to achieve a higher resolution understanding of the functionality of each MCR by mutating highly conserved residues implicated by HDX-MS to be critical for G6P binding. The Q241H SNP, which is located immediately upstream of MCR6 was also of interest for its location and for its physiological association with lowered plasma triglyceride levels. Figure 4.4b highlights the conservation of MCRs 3,4 and 6 and points out the location of each of the conserved residues chosen for mutagenesis.

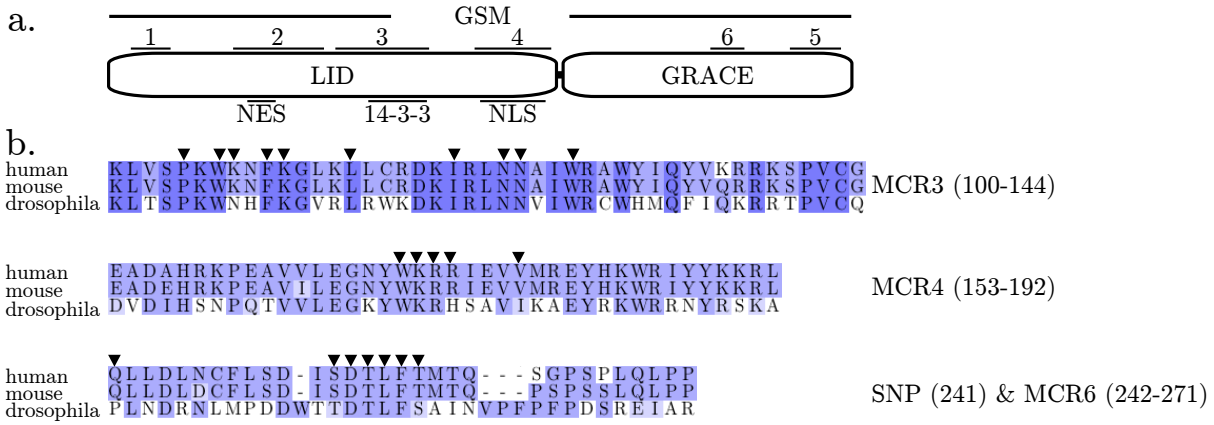


Figure 4.4: **MCRs 3, 4 and 6 are conserved between vertebrate and invertebrate lineages**
a. Diagram showing the location of MCRs along the GSM as well as known MCR functions. **b.** Sequence alignment showing the conservation of MCRs 3, 4 and 6, as well as the SNP between human, mouse and drosophila. Residues targeted for mutation are indicated by a black arrow. Residue backgrounds are filled according to their BLOSUM 62 score. Alignment was performed using MUSCLE [60] with default settings in Jalview 2.8.2 [61].

4.3.1 MCR3 mutations decrease or abolish GSM activity

To study the role the point mutations indicated in Figure 4.4b have in ChREBP activation, I assayed the ability of a Gal4-GSM construct to activate transcription of a Gal4-UAS-driven firefly luciferase (Fluc) gene (Figure 4.5a). The experimental setup is shown in Figure 4.5b.

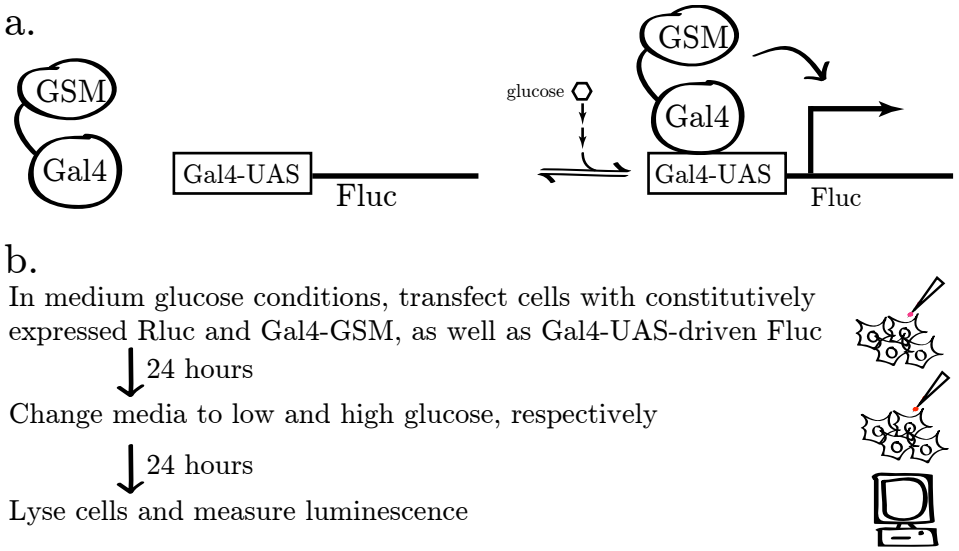


Figure 4.5: **Overview of luciferase assay of targeted MCR mutants**
a. Cartoon of the mechanism of the luciferase reporter assay with the wild type construct. **b.** experimental outline of luciferase reporter assay.

The Gal4-GSM MCR3 mutants were all expressed (Figure 4.6a), however the expression level of mutants in this region was variable, possibly due to the role of this region in protein stability [120], or cell health (discussed below). With the exception of the lysine mutations K107A and K110A, all MCR3 mutants tested had no glucose-responsive activity (Figure 4.6b, c). These results are consistent with previous data which found that MCR3 deletion or point mutations of the α2-helix abolish activity [70, 112, 125]. However, my results contrast with work from the Uyeda lab, which found that the W127A mutation has no effect on the glucose responsiveness of ChREBP [111]. The possibility that this discrepancy is due to lower relative expression of my W127A construct relative to WT cannot be excluded, however it is worth noting that other mutations with lower expression levels than WT, including W170L, KRR171-3A and F257A show clear reporter activity (see Figures 4.9 and 4.11). It is likely that this discrepancy, as with many other published contradictions in ChREBP activation, is due to use of distinct model systems and a different experimental setup. In contrast to my experiments in a pancreatic β-cell line, the Uyeda group used an *lpk* promoter-driven reporter assay in primary hepatocytes. It has previously been shown that ChREBP behaviour has tissue or cell-line-specific behavioural differences as discussed in Section 1.5.1 and detailed in Table 1.1. Further, by using a reporter driven by an endogenous promoter, Sato et al. probed the response of the ChREBP activation *pathway*, whereas my Gal4-UAS-driven reporter system was probing the activation state of the ChREBP GSM in comparative isolation. The evidence that this distinction is relevant is discussed in Section 1.5.1 of the introduction.

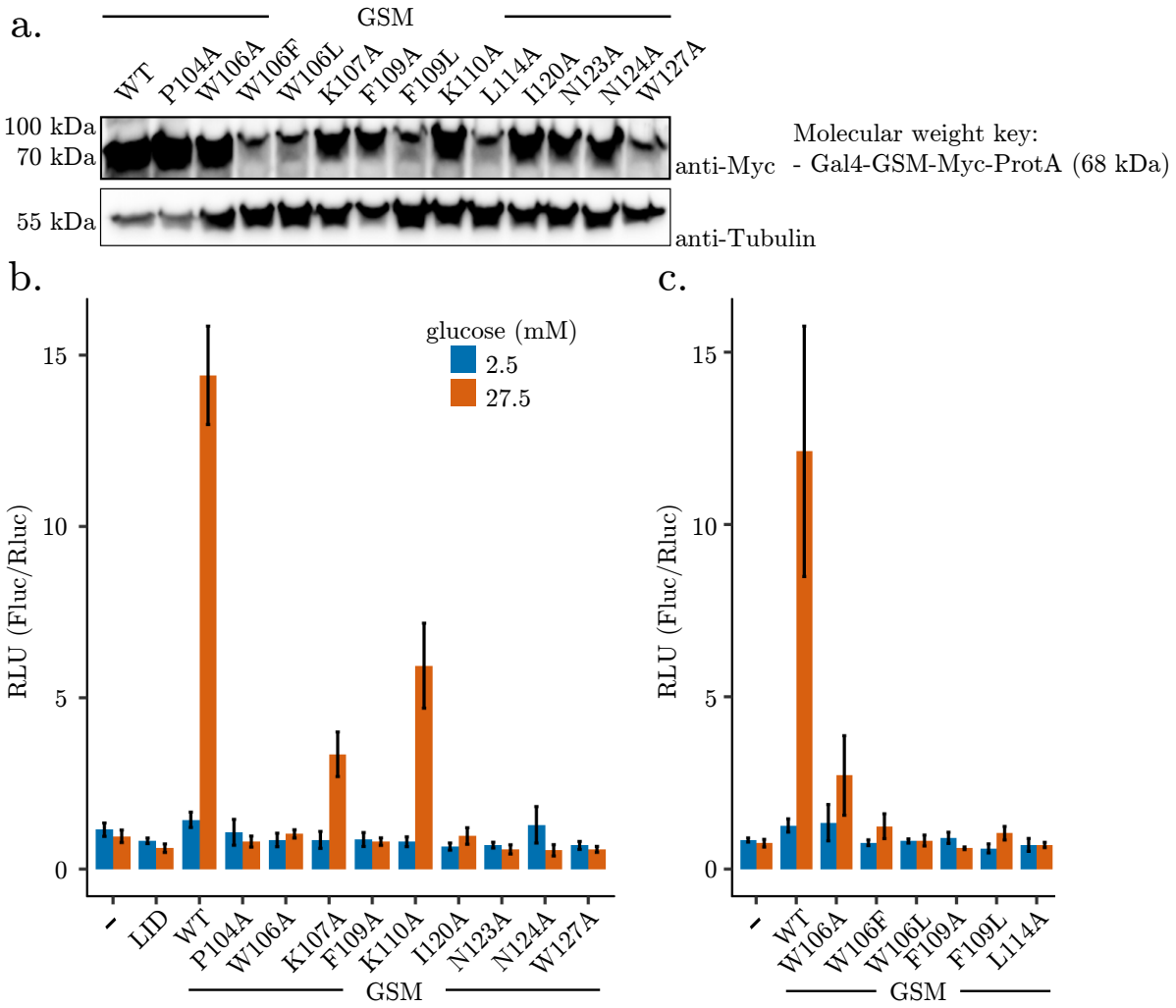


Figure 4.6: **Mutations to MCR3 reduce or abolish the activity of the GSM** **a.** The MCR3 point mutants are all expressed. **b.** Luminescence is expressed as relative light units (RLU), calculated as the Fluc signal normalised to the Rluc control. Of the residues altered, only lysine to alanine mutations still exhibited glucose-responsive activity. p = 0.025 between P104A and WT in high glucose. p = 0.025 between W106A and WT in high glucose. p = 0.020 between K107A and WT in high glucose. p = 0.025 between F109A and WT in high glucose. p = 0.045 between K110A and WT in high glucose. p = 0.024 between I120A and WT in high glucose. p = 0.025 between N123A and WT in high glucose. p = 0.024 between N124A and WT in high glucose. p = 0.025 between W127A and WT in high glucose. N = 4. **c.** More conservative mutations of W106 and F109 did not rescue GSM activity. L114A also abolished protein activity. N = 3. p-values = two-tailed t-test with hommel adjustment for multiple comparisons. Error bars = SEM.

4.3.2 MCR3 mutations show that G6P binding is distinct from glucose-responsive nuclear localisation

To study the role the residues highlighted in Figure 4.4b play in ChREBP subcellular localisation, I quantified the nuclear and cytoplasmic subcellular localisation of a GFP-ChREBP construct in 832/13 cells. The experimental setup is shown in Figure 4.7a. A Jython plugin was written for image quantification (source code given in the Appendix 7.4.2). Figure 4.7b shows representative images of wild type GFP-ChREBP. As can be seen in this panel, in my hands the proportion of

cells with nuclear localised GFP-ChREBP increased between low and high glucose conditions. Despite this increase, GFP-ChREBP remained cytoplasmically localised in most cells even in high glucose conditions. This is consistent with previously reported data on the subcellular localisation of ChREBP in 832/13 cells [69, 70].

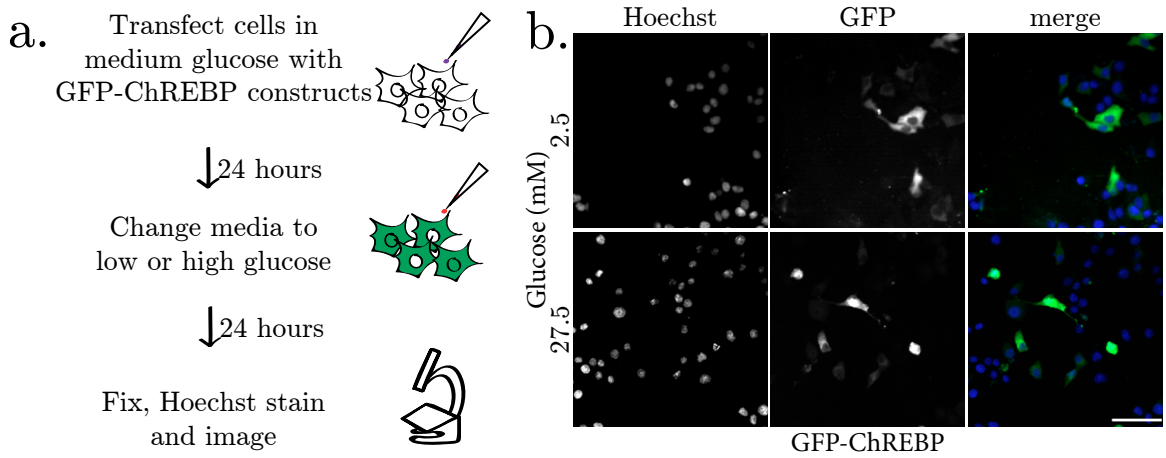


Figure 4.7: **Overview of subcellular localisation microscopy experiments performed a.** Experimental outline for the imaging of the subcellular localisation of GFP-ChREBP. **b.** Representative images of GFP-ChREBP under low and high glucose conditions, respectively. Scale bar = 50 µm.

MCR3 is necessary and sufficient for the association of ChREBP with 14-3-3 [98, 112, 120]. 14-3-3 interacts with the α2-helix of ChREBP; an association strengthened by low energy signals such as AMP, which binds to the interface between 14-3-3 and the α2-helix. The interaction with 14-3-3 and AMP is reported to directly inhibit ChREBP association with importinα [111]. Our HDX-MS data has demonstrated that both AMP and G6P binding drive solvent accessibility changes in the α2-helix, while G6P drives additional solvent accessibility changes in the upstream regions of MCR3. It is therefore of interest to determine the role of MCR3 residues in glucose-responsive subcellular localisation.

As with the luciferase constructs, GFP-ChREBP MCR3 mutants were expressed at varying levels (Figure 4.8a). This is likely at least in part due to the differential effect of these constructs on cell health as the weak bands in the western blot correlate with constructs with a preponderance of unhealthy cells under the microscope. Figure 4.8b and c show the quantified microscopy data for the MCR3 mutants. GFP, GFP-LID and GFP-GRACE have much higher nuclear localisation than GFP-ChREBP constructs and are excluded from the intensity ratio plots for clarity. With the exception of W127A, all mutations tested in MCR3 show clearly blunted or abolished glucose-responsive nuclear localisation. This correlates with the data from the luciferase reporter assay, where MCR3 mutations blunted or abolished glucose-responsive activation.

Surprisingly, despite its inability to bind G6P or promote transcriptional activation, the W127A mutant shows glucose-responsive nuclear localisation. This result shows that glucose-responsive nuclear localisation is a distinct step in ChREBP activation, separable from G6P binding and

protein activation. The behaviour of the other mutants and the potential significance of each is described below.

P104A, which was inactive in luciferase reporter assays, maintains stable levels of both cytosolic and nuclear ChREBP in low and high glucose. It is 84% and 85% cytosolic and 7% and 7% nuclear in low and high glucose, respectively. The non-significant change in intensity ratio seen between low and high glucose is the result of the non-significant decrease in fraction of "both" cells. Similarly, W106 mutants all appear to have reduced or abrogated glucose responsiveness, correlating with their lack of behaviour in the luciferase assays. Interestingly, W106A and W106F appear to have lower basal cytoplasmic localisation than WT, although this difference is not significant. If this were a real effect, it would imply that this residue may play a role in cytosolic sequestration of ChREBP. K107A, which is active in luciferase reporter assays, shows glucose-responsive localisation. Similarly, K110A shows reporter activity and appears to have media glucose induced nuclear localisation, although this is not significant. The F109 mutants, like the W106 mutants, appear to have abolished glucose-responsive nuclear localisation correlating with their lack of reporter activity. L114, which showed no reporter activity may still have glucose-responsive nuclear localisation although the change seen is within error.

Of the α2-helix mutants from the 14-3-3 binding region shown in the crystal structure, only W127A shows significant glucose-responsive nuclear localisation, although non-significant increases in the intensity ratio in high glucose can be seen for all. Interestingly, although N124 and W127 contact 14-3-3 directly in the crystal structure (Figure 1.8 and [120]), they behave differently in the nuclear localisation assay. Meanwhile, N123, which does not contact 14-3-3 in the crystal structure, exhibits an almost identical nuclear localisation pattern to N124 indicating that changes in 14-3-3 association may not necessarily be directly responsible for the localisation patterns seen here. It should be noted however, that 14-3-3 can bind purified GSM_W127A (Thomas Pysik, unpublished data) and it remains possible that these mutations each have a differential effect on 14-3-3 association.

4.3.3 MCR4 NLS mutations have varied effects on GSM activity

Canonical NLSs contain one or more basic patches required for binding to importin proteins. The bipartate NLS of ChREBP is located within MCR4 and is required for low glucose repression [108, 112]. The bipartate NLS is canonically composed of two basic patches separated by an approximately 10 bp linker. Although this function was initially assigned to the stretch between R158 and E175 [76], it is notable that the first basic patch in this region is not conserved relative to the surrounding sequence, while the downstream KR171-2 and KWR183-5 basic patches are unchanged between the human and drosophila proteins (Figure 4.4b). Indeed, mutagenic work done by the Uyeda lab in rodent ChREBP showed that ChREBP NLS requires basic patches in the region K159 to K190 for full functionality [135]. This group recently reiterated the functional relevance of this region when they uploaded the structure of N168 to V176, including the KR171-2

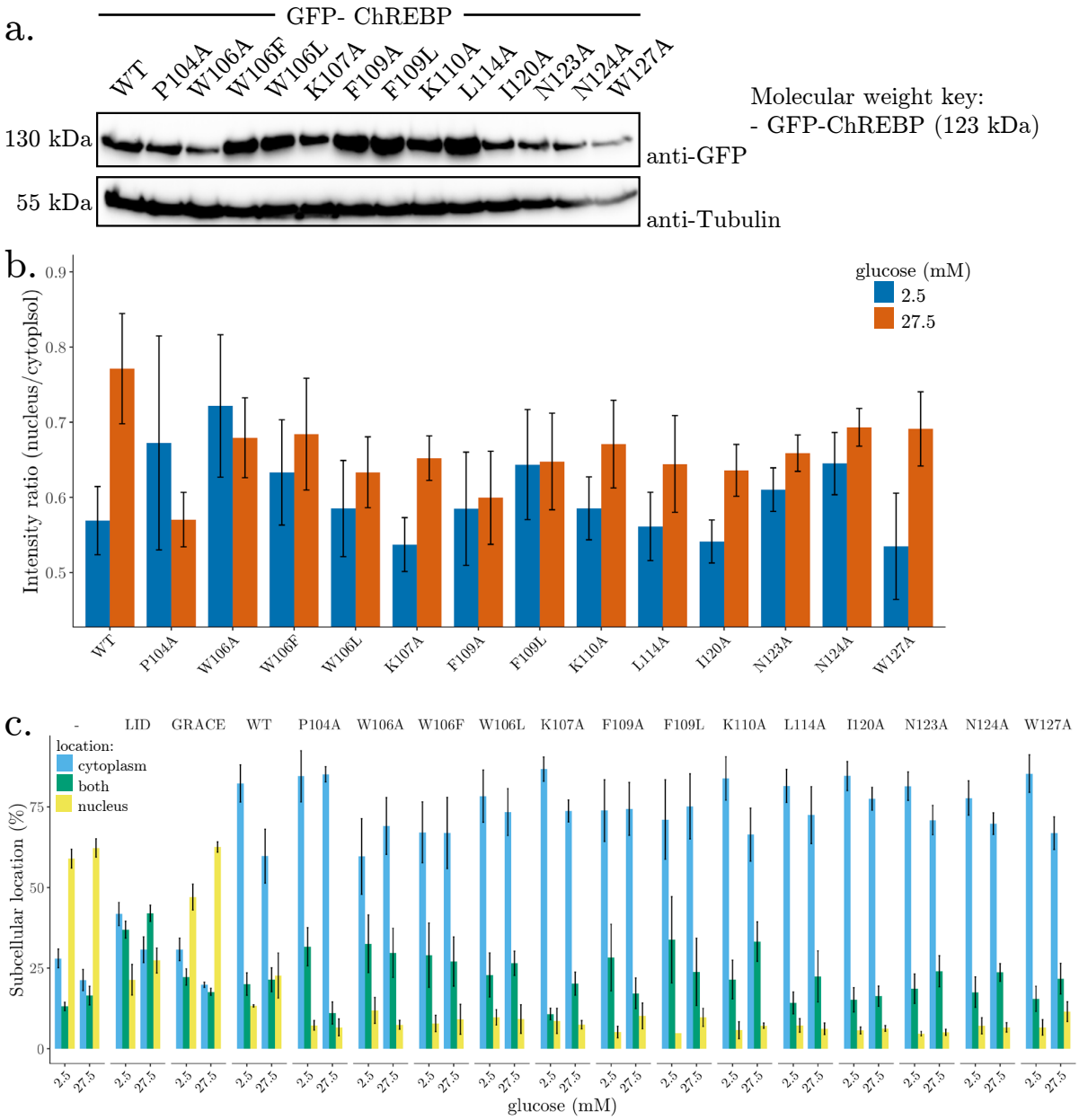


Figure 4.8: **MCR3 point mutations have varied effects on subcellular localisation** **a.** GFP-ChREBP constructs are expressed at the correct size. **b.** With the exception of W127A, all mutants shown exhibit reduced or abrogated glucose-responsive nuclear localisation. The ratio of nuclear to cytosolic GFP intensity for the indicated conditions is shown. $p = 0.04580$ for WT between low and high glucose. $p = 0.03750$ for K107A between low and high glucose. **c.** ChREBP localisation was classified as either nuclear, cytosolic, or both. The percent of cells with a given localisation is shown. Construct localisation was called based on whether the ratio of nuclear to cytosolic GFP intensity was < 0.75, between 0.75 and 1.25 inclusive, or > 1.25 and then assigned as cytosolic, both, or nuclear, respectively. $p = 0.0248$ for cytosolic GRACE between low and high glucose. $p = 0.00955$ for nuclear GRACE between low and high glucose. $p = 0.0379$ for cytosolic W127A between low and high glucose. $N = 6$. p-values = two-tailed t-test. Error bars = SEM. No error bar is shown for F109L in the nucleus as cells with nuclear localised GFP-ChREBP_F109L were only detected in 1 replicate.

basic patch mentioned above, bound to importinα onto the PDB (Figure 1.14 and [127] not yet

published).

As the HDX-MS data showed that solvent accessibility in the W170 to M178 region was specifically changed by G6P binding, I mutated conserved residues in this region to assess its role in protein activity. The Gal4$_{DBD}$-GSM construct is well suited to study the role of the NLS in protein activation as it contains its own functional NLS, thereby dissociating effects on protein localisation from those on protein activity. Mutations within this region appear to be better tolerated, at least with regard to construct expression level, than those in MCR3 (Figure 4.9a). Conservative mutation of the tryptophan immediately upstream of the KRR basic patch reduced but did not abolish reporter activity (W170L and W170F). In contrast, changing the highly conserved basic lysine 171 residue to alanine actually appeared to increase reporter activity in high glucose, although this increase was not significant after adjustment for multiple comparisons. Interestingly, while mutating valine 177 to alanine abolished reporter activity, the V177I "drosophila-like" mutant behaved like wild type (Figure 4.9b). Finally, my results demonstrate that the KRR 171-3 basic patch is not necessary for protein glucose responsiveness (Figure 4.9c).

4.3.4 MCR4 point mutations demonstrate that W170 and V177 are critical for glucose-responsive nuclear localisation, while K171 is not.

As solvent accessibility of the NLS is changed in response to G6P binding, it was relevant to determine the role of this region in the glucose-responsive subcellular localisation of ChREBP. To this end, I performed confocal microscopy on full-length GFP-ChREBP containing mutations in the MCR4 NLS. The GFP-ChREBP MCR4 mutants were all expressed at the correct size (Figure 4.10a) and distinct residues had distinct glucose-responsive subcellular localisation phenotypes.

W170 mutants show reduced activity in reporter assays and no major glucose-responsive change in subcellular localisation can be seen for either W170F or W170L (Figure 4.10b and c). Taken together with my luciferase results, these data imply that W170 is essential for glucose-responsive nuclear localisation but not protein activation.

Surprisingly, the K171A mutation, which removes the charge involved in importinα binding, shows a significant change in localisation between low and high glucose, while the more conservative K171R mutant does not. It is possible that any benefits achieved by charge maintenance in K171R are outweighed by steric clashes created by the larger arginine molecule. Most surprising is that the KRR171-3A construct, which represents a complete removal of the basic patch shown in the crystal structure [127] also showed significant changes in nuclear localisation between low and high glucose. Thus, this conserved basic patch is neither necessary for nuclear localisation, nor for glucose-responsiveness.

When the V177A mutant, which showed no reporter activity, is compared to the V177I "drosophila-like" mutant, which showed wild type equivalent glucose-responsive activity, it appears that the former may actually have higher glucose-responsive nuclear localisation than the latter, although

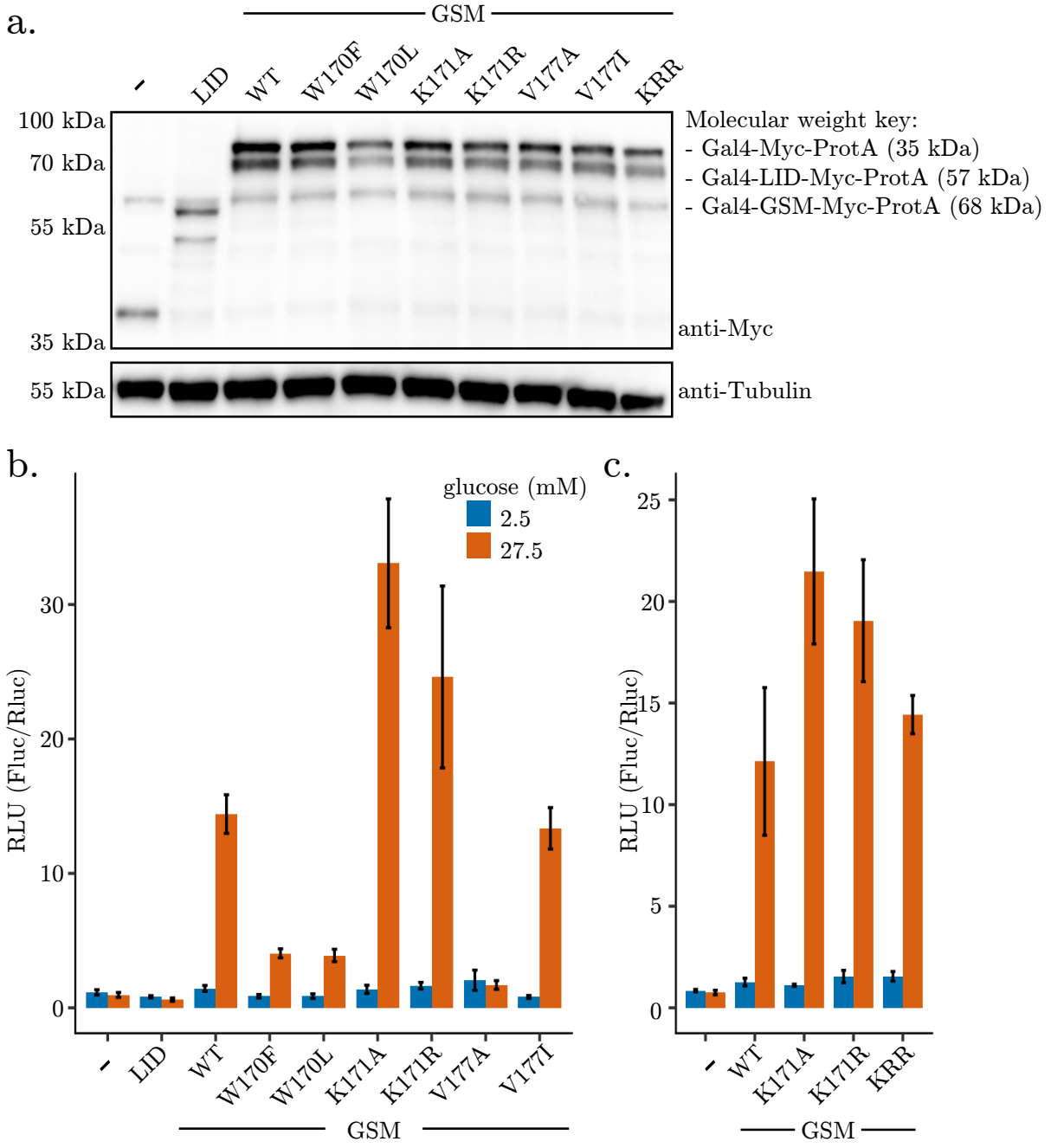


Figure 4.9: **Mutations to conserved residues in MCR4 have varying effects on GSM activity**
a. The MCR4 point mutants are all expressed. **b.** Mutations to W170 showed decreased activation in high glucose. K171 mutations showed a trend towards higher maximal activation of ChREBP although this was not significant after p-value adjustment. V177A abolished activity while V177I did not dramatically alter activation. p = 0.042 between W170F and WT in high glucose. p = 0.032 between W170L and WT in high glucose. p = 0.025 between V177A and WT in high glucose. N = 4. **c.** KRR in the graph legend refers to the KRR171-3A triple mutation. The charge of the KRR171-3 basic patch is not necessary for glucose-responsiveness of the Gal4-GSM construct. N = 3. p-values = two-tailed t-test with hommel adjustment for multiple comparisons. Error bars = SEM.

none of these changes are significant. Critically, the GRACE but not the MCR4-containing LID shows glucose-responsive nuclear localisation. This indicates that the glucose-responsive

regulation of nuclear localisation may occur in a region distal to the NLS.

These data, together with the W127A data from the previous section, imply that glucose-responsive nuclear localisation and glucose-responsive activation are two distinct steps in ChREBP activation and indicate that the GRACE domain may play a role in the glucose-responsive regulation of nuclear localisation. In summary, the MCR4 results show that the conserved regions surrounding the conserved KRR basic patch are actually more critical to protein function, both with regards to protein activation and subcellular localisation, than the patch itself.

4.3.5 Mutation of the SNP drives reduced GSM activity, while changes to adjacent MCR6 promote increased basal activity

MCR6 has been hypothesised to act as both the G6P binding motif of ChREBP, and a transcriptional activation domain involved in the binding of additional activatory proteins [68]. Nina Heppner in my lab demonstrated that $ChREBP_{39-251}$, which lacks the predicted G6P binding motif, can bind G6P with a K_d on par with that of the full-length GSM [129]. However, the question of whether or not MCR6 constitutes a transcriptional activation domain remains to be answered.

Additionally, the Q241H SNP immediately upstream of MCR6 has been linked to decreased plasma triglyceride levels in humans [73, 74]. The purified GSM containing this SNP binds G6P with an unchanged K_d relative to WT (Thomas Pysik, unpublished data) and the molecular mechanism of this phenotype is an open question.

To investigate the role that MCR6 and the Q241H SNP play in protein activity, I expressed the GSM containing the SNP or point mutants of the conserved MCR6 SDTLFT core. These mutations overall appeared better tolerated, with regards to construct expression level and protein activity compared to those in MCR3 (Figure 4.11). The SNP showed reduced activity in high glucose relative to wild type, although the possibility that this difference is the result of the slightly lower expression level seen in the western cannot be excluded. These results point to decreased ChREBP activity as the potential cause of the lowered plasma triglyceride levels observed in individuals with this SNP.

Surprisingly, given its predicted status as a transcriptional activation domain, all MCR6 point mutations tested, with the exception of S253A, tended toward increased basal and/or maximal activity. These results indicate that the DTLFT residues of MCR6 play a repressive role in the regulation of ChREBP activity. The quintuple mutation SDTLFT253-8A exhibited the same glucose responsiveness as the S253A single mutation, implying that S253 functionality is required for the release of basal inhibition caused by other mutants (Figure 4.11c), although a

DTLFT254-8A mutant would need to be tested to verify this hypothesis.

A glucose-response curve highlighted the increased basal activity of the DTLFT mutants. Each of these mutants appeared to show a leftward shift in their respective glucose response curves (Figure 4.12a, b). These data are consistent with a repressive role for the DTLFT residues of MCR6 in the activation of the GSM. Taken together with the evidence that MCR6 is necessary for the LID-GRACE interaction and that the GSM opens in high glucose (Figures 3.8 and 3.9), a model presents itself in which MCR6 mediates the LID-GRACE interaction in low glucose and, upon increased media glucose, the GSM binds G6P in an MCR6-independent manner. This G6P binding allows a change in interaction with an additional unknown factor. This changed interaction leads to GSM opening and protein activation. In this model, the alanine mutants of the DTLF region may render the GSM thermodynamically predisposed to the open conformation, lowering the threshold for protein activation.

4.3.6 Mutation of the SNP and MCR6 have varied effects on ChREBP subcellular localisation

To better understand the function of MCR6 and the Q241H SNP in the regulation of ChREBP, I performed subcellular localisation assays to learn how mutations in this region affect the glucose-responsive nuclear localisation of ChREBP. Surprisingly, despite the relatively uniform effect these mutations had on GSM activity, their effect on subcellular localisation was varied (Figure 4.13). Q241H showed a reduced glucose-responsive change in subcellular localisation relative to WT, paralleling the effect this SNP has on the activity of the Gal4 reporter. Similarly, S253A shows blunted nuclear localisation under high glucose, paralleling the effects this mutation has on protein activity. T258 however, has the same localisation phenotype as S253A, contrasting strongly with its reporter assay activity. D254A, L256A and SDTLFT253-8A showed significant changes in subcellular localisation between low and high glucose conditions. Surprisingly, the T255A mutation appears to completely abrogate glucose-responsive nuclear localisation. Taken together with the luciferase results, these data indicate that the conserved DTFLT core of MCR6 has an inhibitory function on the transcriptional activation competence of ChREBP, but plays a more complex role in the regulation of subcellular localisation.

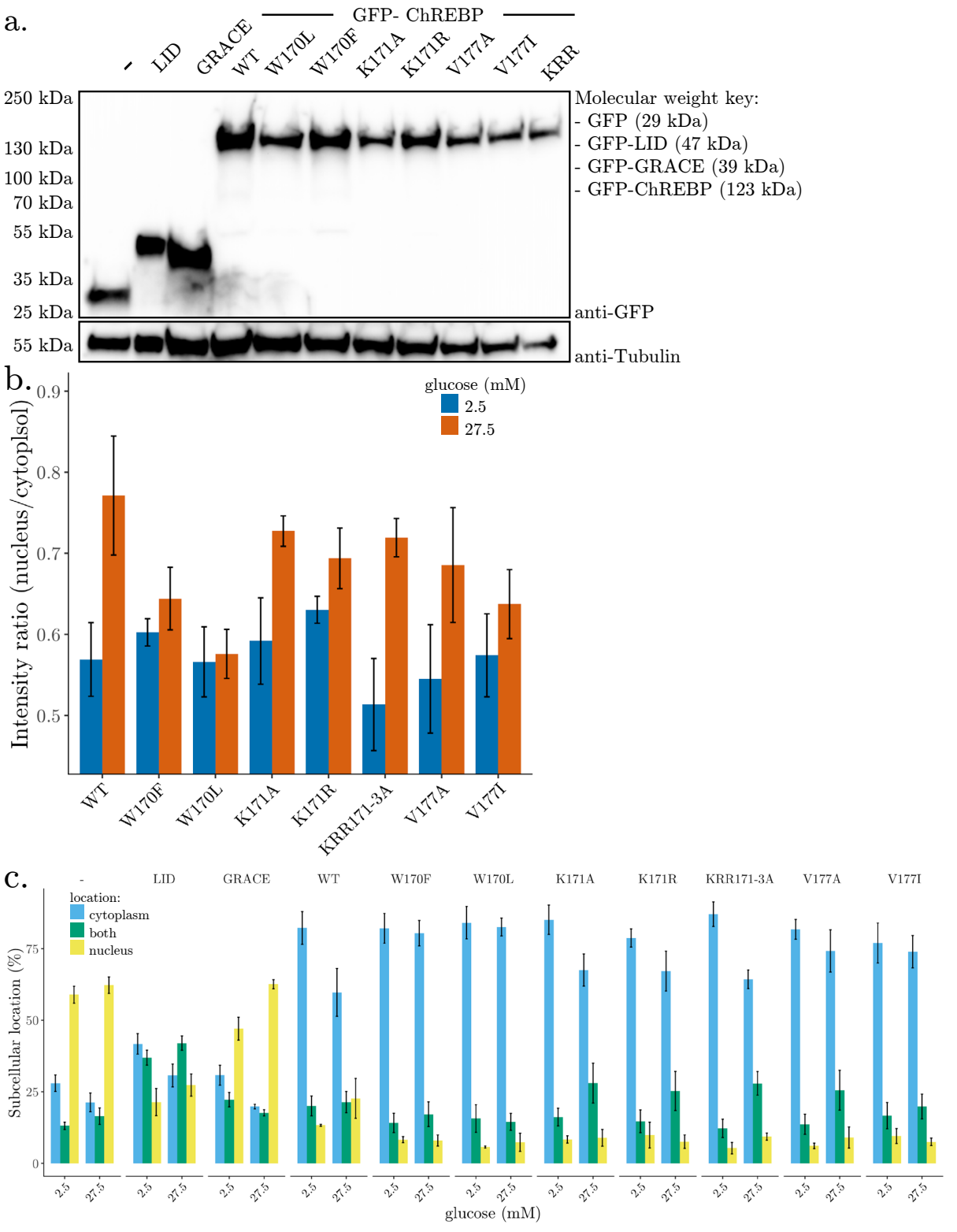


Figure 4.10: **MCR4 point mutations demonstrate W170 and V177 are critical for glucose-responsive nuclear localisation, while K171 is not**
a. GFP-tagged constructs are expressed at the correct size. **b.** The ratio of nuclear to cytosolic GFP intensity for the indicated conditions is shown. p = 0.04580 for WT between low and high glucose. p = 0.01830 for KRR171-3A between low and high glucose. **c.** ChREBP localisation was classified as either nuclear, cytosolic, or both. The percent of cells with a given localisation is shown. p = 0.0248 for cytosolic GRACE between low and high glucose. p = 0.00955 for nuclear GRACE between low and high glucose. p = 0.0436 for cytosolic K171A between low and high glucose. p = 0.0031 for cytosolic KRR171-3A between low and high glucose. p = 0.0169 for "both" localised KRR171-3A between low and high glucose. N = 6. p-values= two-tailed t-test. Error bars = SEM.

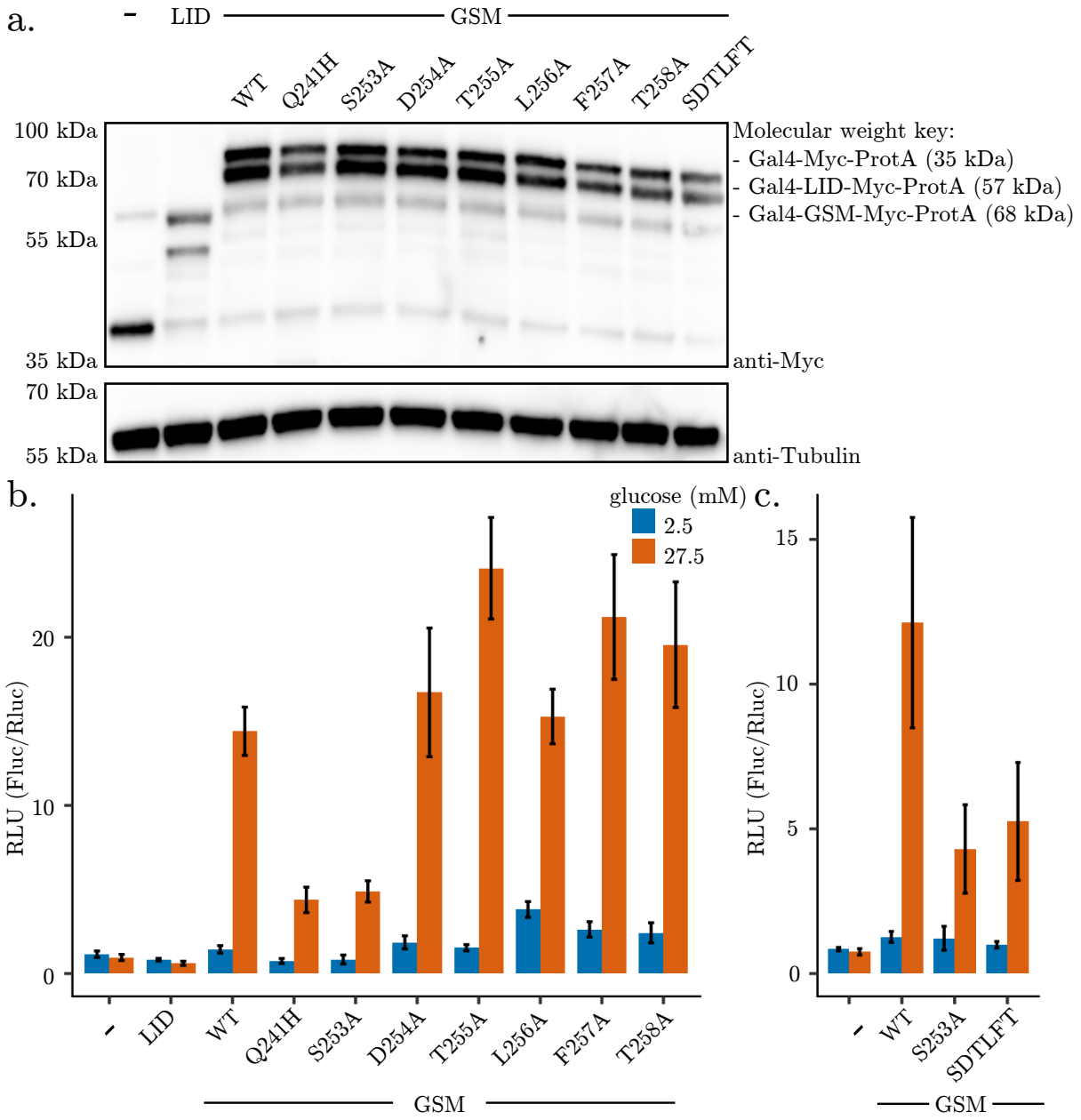


Figure 4.11: **While the Q241H SNP decreases GSM activity, mutations to the MCR6 core increase it**

a. The SNP and MCR6 point mutants used are all expressed. **b.** The Q241H SNP and the S253A mutation decrease maximal GSM activation. Mutations in the highly conserved DTLFT core of MCR6 lead to an increase in basal and/or maximal protein activity. p = 0.025 between Q241H and WT in high glucose. p = 0.034 between S253A and WT in high glucose. N = 4. **c.** SDTLFT in the graph legend refers to the SDTLFT253-8A quintuple mutation. This mutation leads to a decrease in glucose-responsiveness on par with that of S253A. N = 3. p-values = two-tailed t-test with hommel adjustment for multiple comparisons. Error bars = SEM.

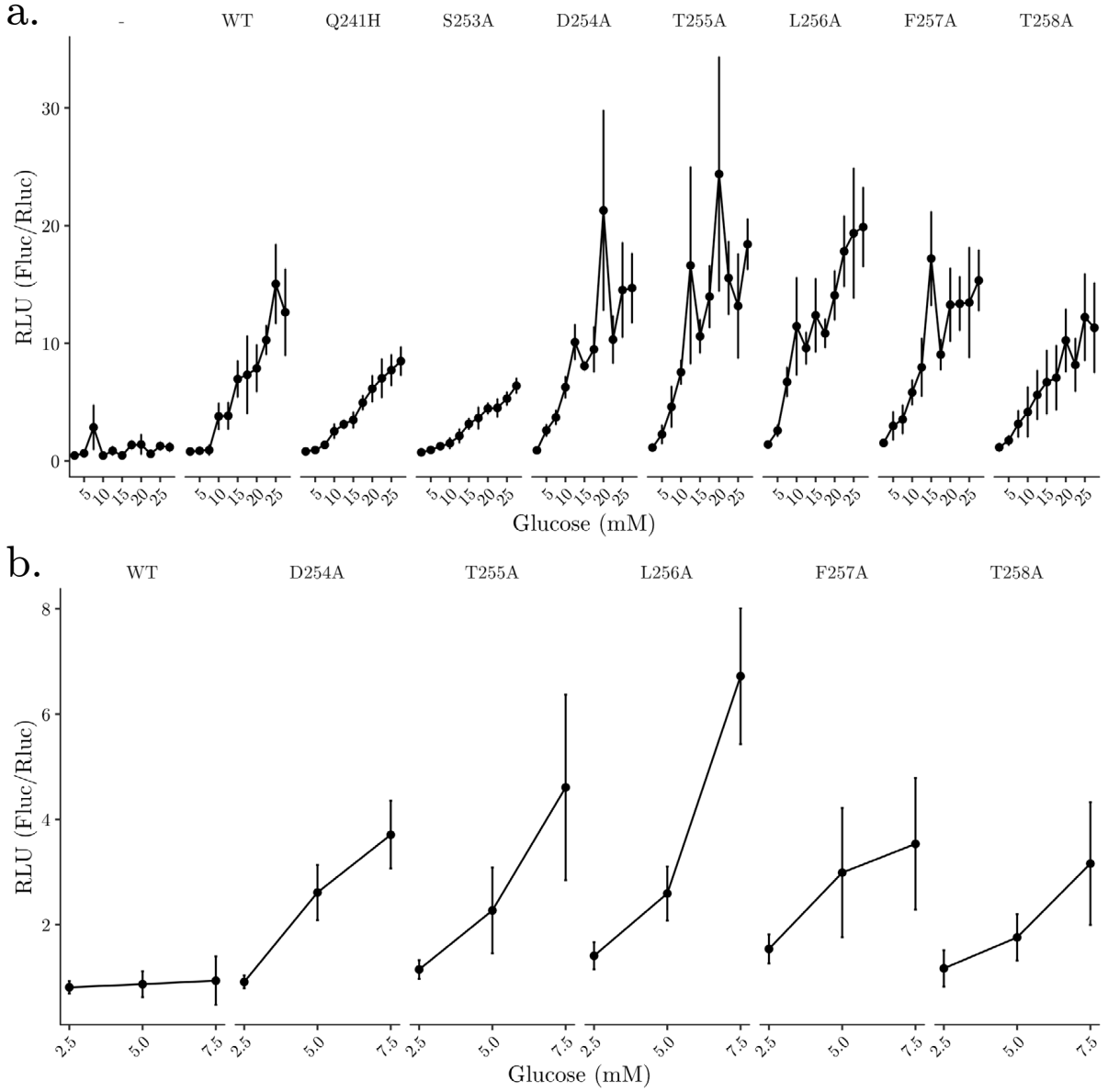


Figure 4.12: **Mutations to the MCR6 DTLFT core increase basal activity**
a. The response curve between 2.5 and 27.5 mM glucose for the Gal4 negative control, the wild type GSM, the Q241H SNP and alanine point mutants of MCR6 are shown. **b.** The same data as in a. with only the wild type and the DTLFT core mutants between 2.5 and 7.5 mM glucose selected. N = 4. Error bars = SEM.

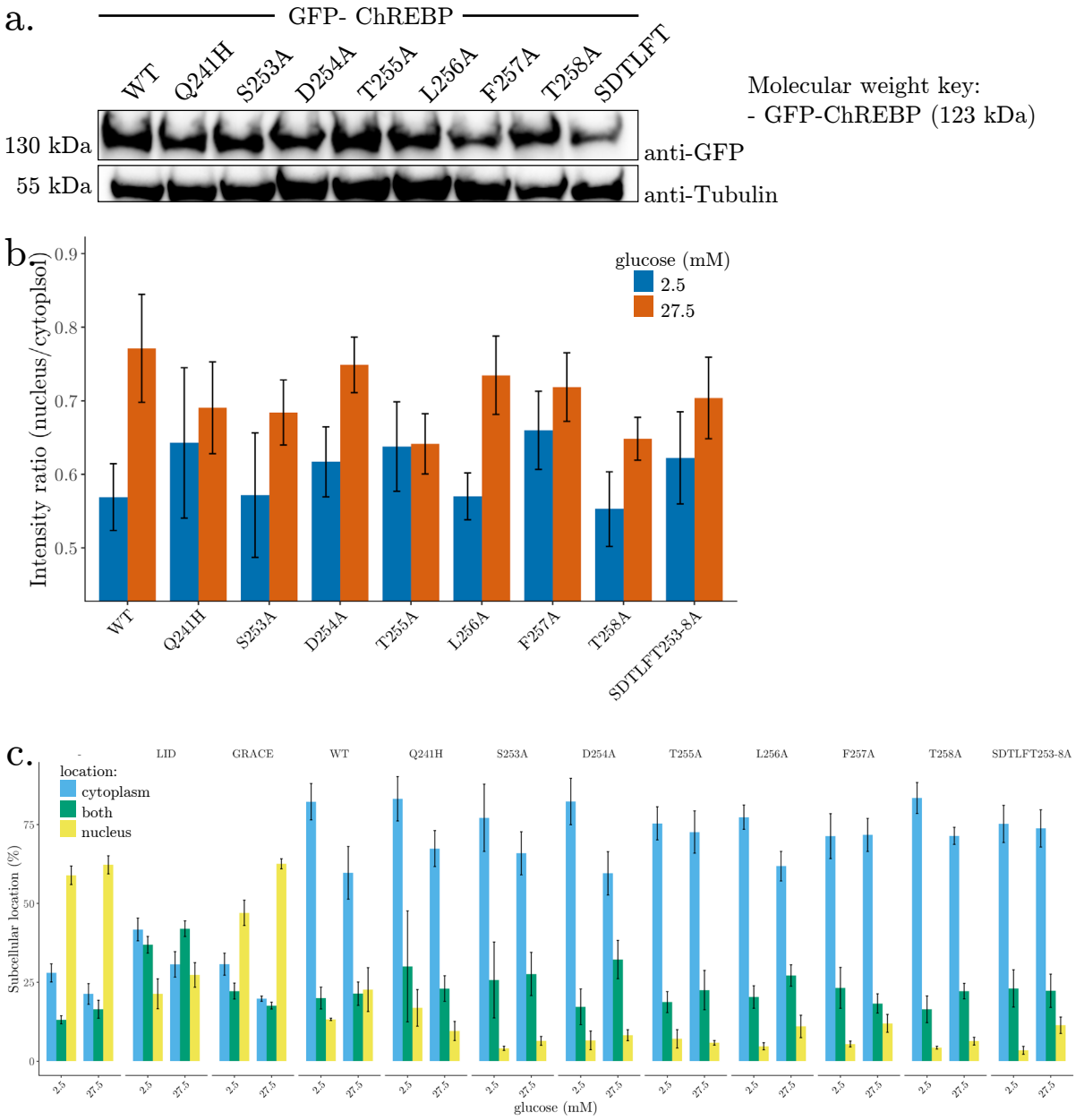


Figure 4.13: **Mutation of the SNP and MCR6 have varied effects on ChREBP subcellular localisation**

a. GFP-ChREBP constructs are expressed at the correct size **b.** The ratio of nuclear to cytosolic GFP intensity for the indicated conditions is shown. p = 0.04580 for WT between low and high glucose. p = 0.02860 for L256A between low and high glucose. **c.** ChREBP localisation was classified as either nuclear, cytosolic, or both. The percent of cells with a given localisation is shown. p = 0.0248 for cytosolic GRACE between low and high glucose. p = 0.00955 for nuclear GRACE between low and high glucose. p = 0.0471 for cytosolic D254A between low and high glucose. p = 0.0293 for cytosolic L256A between low and high glucose. p = 0.04880 for nuclear SDTLFT253-8A between low and high glucose. N = 6. Error bars = SEM.

4.4 A two-step model of ChREBP activation

In this chapter, FRAP and photoactivation techniques were used to demonstrate that ChREBP is rapidly imported into and exported from the nucleus under both low and high glucose conditions in live cells (Section 4.2). Furthermore, confocal microscopy work found that the W127A mutation, previously demonstrated to abrogate G6P binding in vitro (Thomas Pysik, unpublished data), does not abolish media glucose-responsive subcellular localisation (Section 4.3.1). Finally, in contrast to its predicted function as a transcriptional activation domain [68], my luciferase reporter experiments demonstrate that mutations to the DTLFT core of MCR6 actually increase GSM activity (Section 4.3.5).

Taken together, the results from this chapter have demonstrated that: ChREBP is rapidly imported and exported in steady-state low and high glucose conditions, G6P binding is not necessary for media glucose driven nuclear localisation and the MCR6 DTLFT core is required for basal inhibition of ChREBP activity. When the results from the previous chapter indicating that LID and GRACE domain interact in a media glucose-dependent, G6P independent manner, and that this interaction requires MCR6, are also considered, a model of ChREBP activation presents itself.

Upon changes in environmental glucose, intracellular G6P levels increase and outcompete AMP for binding to the GSM, as suggested by the observed competition between G6P and AMP binding in vitro (Thomas Pysik, unpublished data). G6P binding causes a conformational shift in the GSM, driving changes in the solvent accessibility of the NLS and MCR6 as evidenced by HDX-MS and limited proteolysis experiments (Thomas Pysik in collaboration with Kaspar Rand and Johanna Veiga, unpublished data). The GSM is now in an activation competent state and associates with an additional unknown activatory factor. This factor has media glucose regulated activity, and interaction with it promotes GSM opening and nuclear localisation in high glucose conditions. This is supported by the sugar sensitive nuclear localisation seen in the W127A mutant, despite that mutant's observed inability to bind G6P (Section 4.3.1). Once ChREBP has bound G6P *and* interacted with the activatory factor, it is capable of activating the transcription of target genes (Figure 4.14).

In this model, the W127A mutant is incapable of binding G6P but can interact with the activatory factor and thus shows glucose-responsive nuclear localisation but no luciferase reporter activity.

In this model, the MCR6 mutants promote the open conformation of the GSM, stabilising the interaction with the activatory factor and promoting increased basal activity. This is supported by the F2H results indicating that MCR6 is necessary for the LID-GRACE interaction (Section 3.4.1) and the luciferase data showing MCR6 mutations drive increased GSM reporter activity (Section 4.3.5).

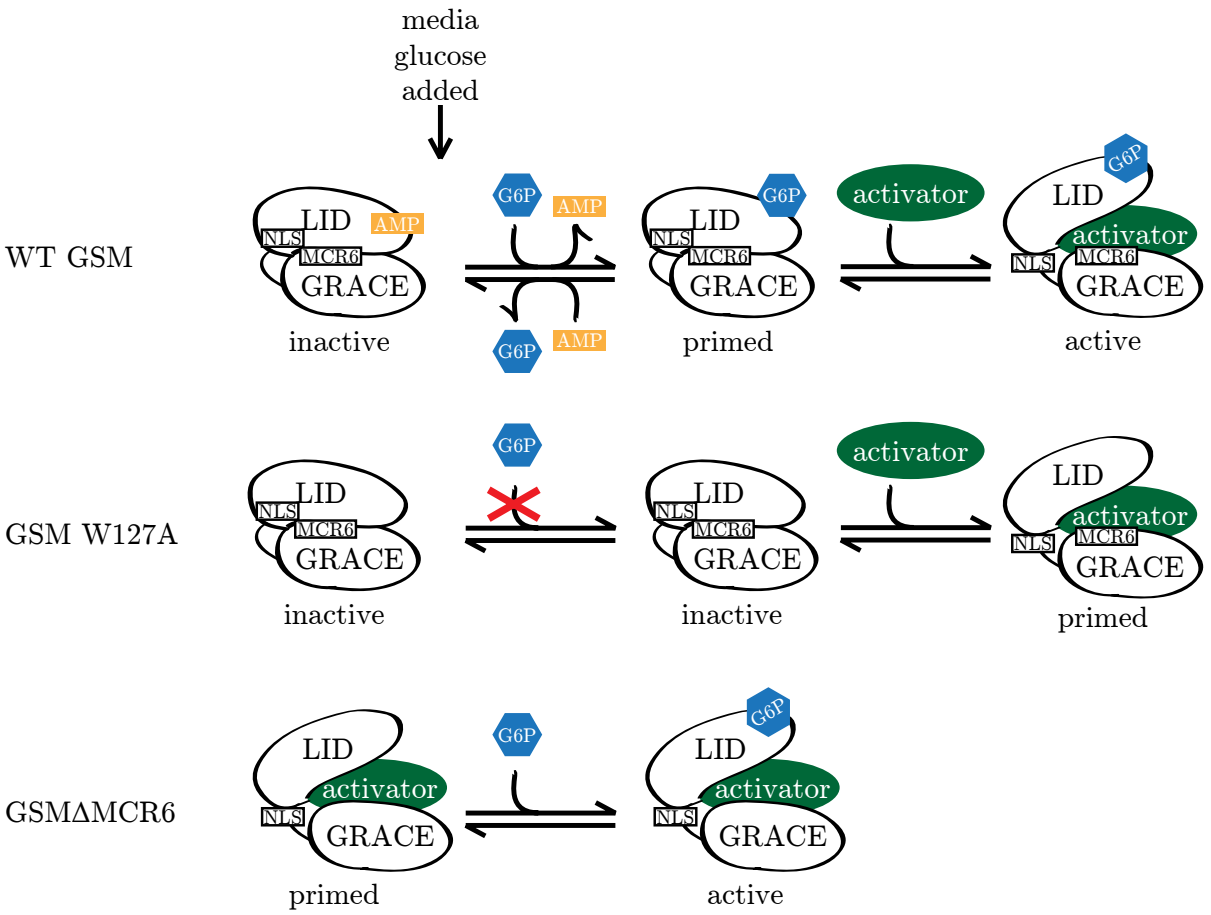


Figure 4.14: **A two-step model of ChREBP activation**
Increased environmental glucose promotes to an increase in intracellular G6P. This results in G6P binding by ChREBP. Further association with additional activatory factor(s) drives GSM opening and NLS exposure and completes activation. The W127A mutant is capable of interacting with the activator but not G6P and is therefore incapable of transcriptional activation. Mutations to MCR6 promote GSM opening and activator binding, reducing the threshold for GSM activation.

4.5 Discussion and future directions

The model proposed above leaves two key open questions. First, what is the nature or the activatory factor and how does it regulate ChREBP? Second, what interactions or intramolecular contacts allow MCR6 to maintain the GSM in the closed conformation? The answer to the first question may lie in one of the previously characterised regulators of ChREBP. For example, it is possible that a metabolite such as F26BP which promotes ChREBP activity in hepatocytes, may act as the unspecified activatory factor in the model above. It may be possible to pinpoint this factor using an in vitro transcription assay to identify the minimal components necessary for ChREBP activity. Based on the glucose insensitive activity of MCR4 mutants [112] and my mutation data from the KRR basic patch, I hypothesise that the answer to the second question lies in the NLS of MCR4. If this is the case, scanning alanine mutagenesis to identify mutations which phenocopy the increased basal activity of the DTLF mutants may be able to identify this interaction site.

In addition, more data are needed to understand how the FRAP and photoactivation results fit into my proposed model. When the FRAP data are compared with published nuclear import studies, the rate of GFP-ChREBP nuclear import is surprisingly fast. Specifically, with a recovery time of less than 30 seconds (determined by the point after the initial bleach at which the normalised intensity trace of the LMB treated cells reaches a plateau) the import rate measured is actually more rapid than the published rate of import of NLS tagged GFP [138] or the rate of passive diffusion of GFP across the nuclear envelope [139]. The rate of GFP-ChREBP nuclear import reported here also stands in stark contrast to the known rates of importinα-mediated import of other transcription factors such as JNK1 and STAT1 [140, 142]. Similarly, live cell studies of the rate of nuclear export of GFP-tagged CRM1 (the nuclear export protein responsible for the active export of ChREBP from the nucleus), show a slower rate of export than that observed here [143]. This is notable as the apparent rate of nuclear export in the photoactivation data is lower than the actual rate due to the confounding effect of simultaneous nuclear import. The microscopy data do not point to a mechanism for this rapid cycling or any way to directly link the rapid nuclear import to observed steady-state levels of nuclear occupancy. It is possible that the constant cycling observed in these experiments represents a mechanism to allow pancreatic β-cells to respond rapidly to changes in circulating glucose levels.

Chapter 5

Outlook

The mechanism of ChREBP activation remains controversial. The transcription factor appears be activated by glucose metabolites and inhibited by fats and low energy signals. This regulation may in turn be modified by post-translational modifications in response to the physiological energy state. The extent to which a given metabolite or post-translational modification regulates ChREBP activity may be itself controlled in a tissue or cell-type-specific manner.

The results presented in this thesis represent a step toward full understanding of ChREBP activation. These data demonstrate that the LID and GRACE do interact and that loss of this interaction corresponds to a media glucose-responsive opening of the GSM in live cells (Chapter 3). Surprisingly, they imply that this interaction is regulated by MCR6 and not G6P. These data are particularly of note in light of work done by Nina Heppner, who found that the removal of MCRs 5 and 6 from a purified GSM construct did not have a significant effect on the K_d of G6P binding [129]. Taken together, the fact that MCR6 is not necessary for G6P binding but does appear to be required for LID-GRACE association, and the fact that G6P does not appear to play a role in LID-GRACE association, imply that G6P binding and media glucose-driven GSM opening are distinct steps in ChREBP activation. Furthermore, the indication that MCR6 plays a role in the domain-domain interactions between the LID and GRACE provides a tantalising hint at the function of a previously uncharacterised region of the GSM.

This work further demonstrates that the steady-state nuclear fraction of ChREBP is exchanged on a timescale of minutes and that this import and export is largely independent of media glucose levels (Section 4.2). The biological relevance of this exchange is not yet known, however it is possible that this constant cycling provides a means to allow ChREBP to quickly respond to changes in metabolic energy levels. An understanding of how the rapid nuclear exchange of ChREBP observed by FRAP in both low and high glucose is functionally related to the overall increased nuclear localisation of ChREBP seen in high glucose would be valuable for our understanding of the role of subcellular localisation in ChREBP activity.

The results detailed in this work put into question the functional role of the highly conserved basic patch at the centre of the MCR4 NLS. Removal of charges in this patch, by either single

(K171A) or triple (KRR171-3A) alanine mutation had no detrimental effects on the activity of a GSM reporter construct. Similarly, both the single and the triple alanine mutant showed glucose-responsive nuclear localisation in the context of the full-length ChREBP construct (Section 4.3.3). These data are consistent with previous work showing that MCR4 is not necessary for nuclear localisation of ChREBP [125]. Furthermore, the tendency towards increased activation of the single mutant is consistent with the previously characterised role of MCR4 as a repressive module [112] and implies the possible dual functions of the NLS in MCR4 balancing the activatory function of nuclear localisation with an inhibitory effect on nuclear activity. Finally, the fact that the GRACE, which lacks the NLS, shows media glucose responsive nuclear localisation, while the LID, which contains the NLS, does not, may point to a mechanism in which a region distal to the NLS is responsible for the glucose-responsiveness of nuclear occupancy.

In addition to the implications of direct LID-GRACE association, rapid exchange of nuclear ChREBP, and apparent lack of functionality of the KRR basic patch in the glucose response of ChREBP in pancreatic β-cells, my data also point to a potential mechanism for the lowered plasma triglyceride levels observed in those with the Q241H rs3812316 SNP. My results show decreased maximal activation of the GSM containing the Q241H mutation and a reduction in the glucose-responsiveness of subcellular localisation of the full-length GFP-tagged SNP-containing protein (Section 4.3.5). The observed decrease in activity and glucose-responsive nuclear localisation of ChREBP in cell culture may directly translate to decreased ChREBP activity at the whole body level. As ChREBP is a major lipogenic driver in both liver and adipose tissue, a global decrease in ChREBP activity would likely correspond to a downregulation of lipogenic metabolism in both these tissues, resulting in a decrease in circulating triglycerides. Decreased ChREBP activity in individuals with the minor H-allele would be consistent with the previously hypothesized glucose sparing function of this SNP [74].

Immediately downstream of the SNP, the conserved DTLF core of MCR6 appears to have a repressive function on ChREBP activity. Single alanine mutations of any of these residues lead to increased activity in low glucose conditions (Section 4.3.5). Taken together with the F2H data in which MCR6 was necessary for LID-GRACE colocalisation, it seems possible that the MCR6 is responsible for mediating LID-GRACE contacts and that mutations to this region decrease the affinity of the two domains for each other, lowering the energetic threshold for protein activation.

Finally, the data put forward here show that media glucose-driven activity and subcellular localisation changes are functionally distinct stages of ChREBP regulation. This is evidenced by the W127A mutant, which shows no glucose-responsive luciferase reporter activity but maintains glucose-responsive nuclear localisation (Section 4.3.1). Taken together with in vitro data showing that the W127A mutation abolishes G6P binding by the purified GSM (Thomas Pysik, unpublished data), these data indicate that G6P binding and glucose-regulated subcellular localisation occur in functionally separable steps in pancreatic β-cells. If the regulation of nuclear localisation occurs in a manner independent of G6P binding by ChREBP, it must be mediated by a change in the interaction with an additional factor or factors. It is possible that this change in interaction is a release in inhibition of a repressive factor, such as the Sorcin mediated regulation proposed

previously [100], or interaction with an activatory factor, such as F26BP, which has been published to promote ChREBP activity in hepatocytes [38] and which binds directly to the purified GSM in vitro (Thomas Pysik, unpublished data).

In summary, my data support a model of ChREBP activation in which GSM repression is mediated by direct LID-GRACE contacts involving MCR6. In this model, G6P binding promotes protein activation but is insufficient to drive release of intramolecular contacts. Opening of the GSM occurs upon media glucose-driven changes in GSM interaction with additional unknown factor(s). These changes drive glucose-sensitive subcellular localisation of ChREBP in a manner independent of G6P binding. These results provide potential explanations for open questions arising from the published body of literature about ChREBP function and further our understanding of the regulation of a lipogenic transcription factor that plays a central role in metabolic health and disease.

These results raise questions about the role of protein-protein interactions in the regulation of ChREBP as well as the conservation of key sequences across species and between Mondo isoforms. It is very possible that the G6P binding independent changes in subcellular localisation observed are the result of altered protein-protein interactions. Candidate protein interactors include previously demonstrated regulators of ChREBP, such as the nuclear hormone receptor LXR, and the Ca^{2+}–binding protein Sorcin [32, 100]. An identification of the protein or proteins involved in the glucose-sensitive nuclear localisation of ChREBP is a key next step in elucidating the mechanism of regulation of ChREBP. Initial microscopic data from a cotransfection of ChREBP and LXR indicate that it is possible that LXR may be capable of promoting the nuclear localisation of ChREBP (Imke Mandemaker and myself, unpublished data), making it an interesting potential candidate for the activatory factor put forward in my model. Even in the context of a known activatory factor, questions would remain regarding the conservation of this mechanism of regulation across tissue types and between Mondo isoforms.

The residues mutated in this work are conserved between the two Mondo paralogues, MondoA and MondoB (ChREBP), opening the possibility that the proposed model may be relevant for both Mondo isoforms. However, work in pancreatic β-cells indicates that glucose-stimulated nuclear translocation occurs on a distinct timescale for the respective proteins [144]. Phosphorylation differences may explain some portion of this functional difference, as MondoA lacks many of the known phosphorylation sites of ChREBP [68]. Mutagenic addition of ChREBP phosphorylation sites to MondoA, combined with luciferase reporter assays and subcellular localisation studies would help characterise the extent to which phosphorylation is responsible for the differential behaviour of the two paralogues. More generally, identification of residues which differ between paralogues, as well as a comparison of the respective interactomes of each protein may provide clues to the regulation of these proteins. A better understanding of the differences in function of the two Mondo paralogues would be a critical step towards the characterisation of the metabolic response to sugar in tissues which express either or both proteins.

Finally, the generation of a structure of the G6P-bound GSM would be a central step towards understanding to what extent and how G6P-binding drives changes in the conformation of ChREBP.

Knowledge of exactly how G6P is bound would provide a first step towards understanding the precise mechanism by which G6P alters ChREBP behaviour. Furthermore, a structure has the potential to provide clues towards our understanding of the mechanism mediating functional differences between the two Mondo paralogues as well as how G6P binding affects protein-protein interactions of ChREBP. In summary, identification of the interacting factor responsible for ChREBP glucose-sensitive nuclear localisation, investigation of the mechanism of the functional differences between ChREBP and MondoA, and a structure of the G6P-bound GSM would be key next steps in our understanding of the regulation of this central metabolic transcription factor.

Chapter 6

Materials and Methods

A complete list of the reagents mentioned in the following sections is available as Table 7.1 in the Appendix.

6.1 Cell culture

Cells were grown in humidified incubators at 37 °C and 5% CO_2. All U2OS-derived cell lines were grown in DMEM (Sigma) supplemented with 100 U/ml penicillin-streptomycin solution (Thermo), 2 mM L-glutamine (Thermo), and 10% fetal bovine serum (Thermo). U2OS-derived cell lines were grown in 25 mM glucose conditions unless otherwise indicated. U2OS 2B2 cells were additionally cultured in the presence of 150 µg/ml hygromycin [132]. INS1-derived 832/13 cells [134] were cultured in RPMI (Thermo) supplemented with 10 mM HEPES (Thermo), 2 mM L-glutamine, 1 mM sodium-pyruvate (Thermo), 10% fetal calf serum, and 0.05 mM β-mercaptoethanol (Sigma) without the addition of antibiotics. These cells were grown in 11 mM glucose conditions unless otherwise indicated. Cells were passaged twice a week and checked regularly for mycoplasma. Cell counting was performed using the Vi-CELL XR (Beckman Coulter).

6.2 Transfection

U2OS-derived cell lines were plated at a density of 6×10^4 cells/ml on the indicated plates, 100 µl per well of a 96-well plate, the volume scaled with the per-well growth area of the plate. Cells were transfected the day after plating using X-tremeGENE HP DNA Transfection Reagent (Sigma) following manufacturer protocol. In brief, for each transfection, 2 µg plasmid DNA were diluted into 200 µl OptiMEM (Thermo). Four µl X-tremeGENE HP DNA Transfection Reagent was added to this and the solution was mixed thoroughly by pipetting. 30 minutes after mixing, the transfection solution was added dropwise to cells, 10 µl per well of a 96-well plate. After addition of reagent, plates were gently swirled before being returned to the incubator.

832/13 cells were plated at 2×10^5 cells/ml on the indicated plates, 200 µl per well of a 96-well plate. Cells were transfected the day following plating using Lipofectamine 2000 Transfection Reagent (Thermo) following manufacturer protocol. In brief, a DNA solution was made by diluting 200 ng DNA per 10 µl OptiMEM. In a separate container, a reagent solution was made by diluting 1 µl Lipofectamine 2000 Transfection Reagent per 10 µl OptiMEM. An equal volume of reagent solution was added to the DNA solution and the resulting transfection solution was mixed thoroughly by pipetting. Five minutes after mixing, the transfection solution was added dropwise to cells 10 µl per well of a 96-well plate. To ensure even distribution of transfection solution, plates were gently swirled before being returned to the incubator.

6.3 Cloning

The majority of the constructs used were cloned specifically for this work. The few exceptions are distinguished with a footnote in Table 6.1. Human ChREBP constructs were derived from the pMS-RQ-MLXIPL (Database ID CL3914) ordered custom from Invitrogen GeneArt Gene Synthesis (Thermo). 14-3-3 constructs were derived from pCMV6-Entry-14-3-3 (Database ID CL3679) purchased from Biocat (Catalog Number RC215940-OR). Full-length Mouse ChREBP constructs were derived from peGFP-N1-ChREBP (Database ID CL3266, made by Daria Zinne). pGampac-LID and pGampac-GSM were kind gifts from Professor Lawrence Chan and Professor Benny Hung-Junn Chang at the Baylor College of Medicine. Cloning was performed using either standard subcloning or inverse PCR mutagenesis techniques. All ligation steps were followed by a transformation into XL10 chemically competent cells.

Cloning of pmCherry-C1-LacI-GRACEΔMCR6 was performed by Pedro Garcia. Cloning of pGampac-GSM_P104A, pGampac-GSM_W106A, pGampac-GSM_K107A, pGampac-GSM_F109A and pGampac-GSM_K110A was performed by Julia Preisser. Cloning of pGampac-GSM_N124A, pGampac-GSM_W127A and pGampac-GSM_Q241H was performed with help from Giorgi Beroshvili. All other cloning was performed by myself.

6.3.1 Vectors

All fluorescent proteins used were expressed from C1 or N1 vectors (Clontech), as indicated. Luciferase assays were performed using pRL-TK (Promega), pG5luc (Promega), and pGampac [69] constructs. Human ChREBP constructs were derived from GenBank accession number NM_032951.2. Mouse ChREBP constructs were derived from GenBank accession number NM_021455.5. Human 14-3-3β constructs were derived from GenBank accession number NM_003404. Table 6.1 presents a comprehensive list of all vectors used in this work and details domain boundaries used for the LID, GRACE and GSM. The protocols used to generate these vectors are detailed in Section 6.3. Translated sequences of the open reading frames of pGampac-LID, pGampac-GSM and all non-point mutation vectors cloned are listed in Section 7.3 of the Appendix.

Table 6.1: **Details of constructs used in this work**

Construct	Organism	Database ID	Insert	Experiments
peGFP-C1[a]	-	CL0530	-	CoIP, subcellular localisation
pmCherry-C1[a]	-	CL4276	-	CoIP
peGFP-C1-ChREBP	human	CL3932	ChREBP	F2H
peGFP-C1-LID	human	CL3929	$\text{ChREBP}_{S37-L192}$	CoIP, F2H
peGFP-C1-GRACE	human	CL3930	$\text{ChREBP}_{R197-D298}$	F2H
peGFP-C1-14-3-3	human	CL3935	14-3-3	F2H
pmCherry-C1-LacI[a]	-	CL4236	-	F2H, photoactivation
pmCherry-C1-LacI-ChREBP	human	CL3949	ChREBP	F2H
pmCherry-C1-LacI-LID	human	CL3950	$\text{ChREBP}_{S37-L192}$	F2H
pmCherry-C1-LacI-GRACE	human	CL3951	$\text{ChREBP}_{R197-D298}$	F2H
pmCherry-C1-LacI-14-3-3	human	CL3952	14-3-3	F2H
pmCherry-C1-GRACE	human	CL4279	$\text{ChREBP}_{R197-D298}$	CoIP
pmCherry-C1-LacI-GRACEΔMCR6	human	CL4048	$\text{ChREBP}_{R197-D298\Delta249-260}$	F2H
N1-GRACE-FLAG-Cherry	human	CL4193	$\text{ChREBP}_{R197-D298}$	CoIP
peGFP-C1-GSM	human	CL3931	$\text{ChREBP}_{S37-D298}$	FRET-FLIM
peGFP-C1-GSM-Cherry	human	CL3933	$\text{ChREBP}_{S37-D298}$	FRET-FLIM
peGFP-C1-ChREBP	mouse	CL4181	ChREBP	FRAP, subcellular localisation
pPAGFP-C1-ChREBP	mouse	CL4257	ChREBP	photoactivation

Table 6.1- *Continued from previous page*

Construct	Organism	Database ID	Insert	Experiments
pRL-TK[a]	-	CL4281	-	luciferase
pG5luc[a]	-	CL4282	-	luciferase
pGampac	-	CL4172	-	luciferase
pGampac-LID[a]	mouse	CL4284	ChREBP$_{A2-L192}$	luciferase
pGampac-GSM[a]	mouse	CL4185	ChREBP$_{A2-E298}$	luciferase
pGampac-GSM_P104A	mouse	CL4156	ChREBP$_{A2-E298}$P104A	luciferase
pGampac-GSM_W106A	mouse	CL4157	ChREBP$_{A2-E298}$W106A	luciferase
pGampac-GSM_W106F	mouse	CL4190	ChREBP$_{A2-E298}$W106F	luciferase
pGampac-GSM_W106L	mouse	CL4194	ChREBP$_{A2-E298}$W106L	luciferase
pGampac-GSM_K107A	mouse	CL4158	ChREBP$_{A2-E298}$K107A	luciferase
pGampac-GSM_F109A	mouse	CL4159	ChREBP$_{A2-E298}$F109A	luciferase
pGampac-GSM_F109L	mouse	CL4195	ChREBP$_{A2-E298}$F109L	luciferase
pGampac-GSM_K110A	mouse	CL4160	ChREBP$_{A2-E298}$K110A	luciferase
pGampac-GSM_L114A	mouse	CL4191	ChREBP$_{A2-E298}$L114A	luciferase
pGampac-GSM_I120A	mouse	CL4161	ChREBP$_{A2-E298}$I120A	luciferase
pGampac-GSM_N123A	mouse	CL4162	ChREBP$_{A2-E298}$N123A	luciferase
pGampac-GSM_N124A	mouse	CL4069	ChREBP$_{A2-E298}$N124A	luciferase
pGampac-GSM_W127A	mouse	CL4070	ChREBP$_{A2-E298}$W127A	luciferase

Table 6.1- *Continued from previous page*

Construct	Organism	Database ID	Insert	Experiments
pGampac-GSM_W170F	mouse	CL4174	ChREBP$_{A2-E298}$W170F	luciferase
pGampac-GSM_W170L	mouse	CL4173	ChREBP$_{A2-E298}$W170L	luciferase
pGampac-GSM_K171A	mouse	CL4175	ChREBP$_{A2-E298}$K171A	luciferase
pGampac-GSM_K171R	mouse	CL4176	ChREBP$_{A2-E298}$K171R	luciferase
pGampac-GSM_KRR171-3A	mouse	CL4196	ChREBP$_{A2-E298}$KRR171-3A	luciferase
pGampac-GSM_V177A	mouse	CL4177	ChREBP$_{A2-E298}$V177A	luciferase
pGampac-GSM_V177I	mouse	CL4178	ChREBP$_{A2-E298}$V177I	luciferase
pGampac-GSM_Q241H	mouse	CL4071	ChREBP$_{A2-E298}$Q241H	luciferase
pGampac-GSM_S253A	mouse	CL4166	ChREBP$_{A2-E298}$S253A	luciferase
pGampac-GSM_D254A	mouse	CL4167	ChREBP$_{A2-E298}$D254A	luciferase
pGampac-GSM_T255A	mouse	CL4168	ChREBP$_{A2-E298}$T255A	luciferase
pGampac-GSM_L256A	mouse	CL4169	ChREBP$_{A2-E298}$L256A	luciferase
pGampac-GSM_F257A	mouse	CL4170	ChREBP$_{A2-E298}$F257A	luciferase
pGampac-GSM_T258A	mouse	CL4171	ChREBP$_{A2-E298}$F258A	luciferase
pGampac-GSM_SDTLFT253-8A	mouse	CL4184	ChREBP$_{A2-E298}$SDTLFT253-8A	luciferase
peGFP-C1-LID	mouse	CL4244	ChREBP$_{S37-S196}$	subcellular localisation
peGFP-C1-GRACE	mouse	CL4228	ChREBP$_{S196-E298}$	subcellular localisation
peGFP-C1-ChREBP_P104A	mouse	CL4224	ChREBP P104A	subcellular localisation

Table 6.1- *Continued from previous page*

Construct	Organism	Database ID	Insert	Experiments
peGFP-C1-ChREBP_W106A	mouse	CL4211	ChREBP W106A	subcellular localisation
peGFP-C1-ChREBP_W106F	mouse	CL4209	ChREBP W106F	subcellular localisation
peGFP-C1-ChREBP_W106L	mouse	CL4212	ChREBP W106L	subcellular localisation
peGFP-C1-ChREBP_K107A	mouse	CL4225	ChREBP K107A	subcellular localisation
peGFP-C1-ChREBP_F109A	mouse	CL4213	ChREBP F109A	subcellular localisation
peGFP-C1-ChREBP_F109L	mouse	CL4215	ChREBP F109L	subcellular localisation
peGFP-C1-ChREBP_K110A	mouse	CL4214	ChREBP K110A	subcellular localisation
peGFP-C1-ChREBP_L114A	mouse	CL4202	ChREBP L114A	subcellular localisation
peGFP-C1-ChREBP_I120A	mouse	CL4226	ChREBP I120A	subcellular localisation
peGFP-C1-ChREBP_N123A	mouse	CL4227	ChREBP N123A	subcellular localisation
peGFP-C1-ChREBP_N124A	mouse	CL4200	ChREBP N124A	subcellular localisation
peGFP-C1-ChREBP_W127A	mouse	CL4223	ChREBP W127A	subcellular localisation
peGFP-C1-ChREBP_W170F	mouse	CL4218	ChREBP W170F	subcellular localisation
peGFP-C1-ChREBP_W170L	mouse	CL4201	ChREBP W170L	subcellular localisation
peGFP-C1-ChREBP_K171A	mouse	CL4219	ChREBP K171A	subcellular localisation
peGFP-C1-ChREBP_K171R	mouse	CL4220	ChREBP K171R	subcellular localisation
peGFP-C1-ChREBP_KRR171-3A	mouse	CL4216	ChREBP KRR171-3A	subcellular localisation
peGFP-C1-ChREBP_V177A	mouse	CL4208	ChREBP V177A	subcellular localisation

Table 6.1- *Continued from previous page*

Construct	Organism	Database ID	Insert	Experiments
peGFP-C1-ChREBP_V177I	mouse	CL4178	ChREBP V177I	subcellular localisation
peGFP-C1-ChREBP_Q241H	mouse	CL4210	ChREBP Q241H	subcellular localisation
peGFP-C1-ChREBP_S253A	mouse	CL4203	ChREBP S253A	subcellular localisation
peGFP-C1-ChREBP_D254A	mouse	CL4204	ChREBP D254A	subcellular localisation
peGFP-C1-ChREBP_T255A	mouse	CL4205	ChREBP T255A	subcellular localisation
peGFP-C1-ChREBP_L256A	mouse	CL4206	ChREBP L256A	subcellular localisation
peGFP-C1-ChREBP_F257A	mouse	CL4207	ChREBP F257A	subcellular localisation
peGFP-C1-ChREBP_T258A	mouse	CL4217	ChREBP T258A	subcellular localisation
peGFP-C1-ChREBP_SDTLFT253-8A	mouse	CL4221	ChREBP SDTLFT253-8A	subcellular localisation

[a] Indicates a construct that was *not* cloned specifically for this work.

6.3.2 PCR

Unless noted otherwise, constructs were cloned by standard or inverse PCR amplification followed by restriction digest or phosphorylation, respectively. Primers were designed manually using SnapGene Viewer (from Insightful Science; `https://www.snapgene.com`) and purchased from metabion. PCRs were performed using Phusion® High-Fidelity DNA Polymerase (NEB) in 50 µl reaction volumes following manufacturer instructions, as detailed in Table 6.2.

Table 6.2: **PCR reaction setup**

Reagent	Quantity
5X Phusion buffer	10 µl
10 mM dNTPs	1 µl
10 µM forward primer	2.5 µl
10 µM reverse primer	2.5 µl
template DNA	1-100 ng
Phusion polymerase	0.5 µl
water	to 50 µl

Two-step thermal cycler protocols were selected for predicted annealing temperatures of 72 °C or greater. Otherwise a three-step protocol was used. Annealing temperatures were predicted using the NEB T_m calculator (`http://tmcalculator.neb.com/`). Two-step thermal cycler reactions followed the protocol recommended by NEB for Phusion® amplifications and are detailed in Table 6.3.

Table 6.3: **Two-step PCR protocol used**

Phase	Cycles	Step	Temperature	Time
initial	1	denaturation	98 °C	30 seconds
main	25-35	denaturation	98 °C	10 seconds
		annealing & extension	72 °C	30 seconds/kb
final	1	annealing & extension	72 °C	5 minutes
hold			12 °C	

Three-step PCRs were programmed using a modified slowdown PCR protocol [145] with a heating ramp rate of 2.5 °C per second and a cooling ramp rate of 1.5 °C per second. The initial touchdown phase had sufficient cycles so that the annealing temperature of the final cycle was equal to the final annealing temperature of the PCR. Thermal cyclers were preheated to 98 °C before use. The three-step protocol is detailed in Table 6.4.

6.3.3 Standard subcloning

For standard subcloning protocols, regions of interest were amplified by PCR using primers with restriction sites compatible with the multiple cloning site of the target vector. Primers used and

Table 6.4: **Three-step PCR protocol used**

Phase	Cycles	Step	Temperature	Time
initial	1	denaturation	98 °C	30 seconds
touchdown	3-51	denaturation	98 °C	10 seconds
		annealing	70 °C (-0.3 °C/cycle)	10 seconds
		extension	72 °C	30 seconds/kb
final	15	denaturation	98 °C	10 seconds
		annealing	50 °C - 71 °C	10 seconds
		extension	72 °C	30 seconds/kb
hold			12 °C	

the restriction sites they contain are detailed in Table 6.7. Amplification success was checked for using agarose gel electrophoresis of 1 µl of PCR product. Successful PCRs were column purified using the mi-PCR Purification Kit (metabion). Purified PCR products and target vectors were digested at 37 °C for at least 1 hour using NEB enzymes following manufacturer protocol. The reaction setup is detailed in Table 6.5 All enzymes used here were compatible with CutSmart buffer and 37 °C digest.

Table 6.5: **Restriction digest setup**

Reagent	Quantity
10X CutSmart Buffer	5 µl
DNA	1 µg
enzyme(s)	1 µl each
water	to 50 µl

Digests were column purified using the mi-PCR Purification Kit if excised DNA fragments were less than 20 bp. Otherwise, digests were run on an agarose gel and the DNA fragment of the correct size was excised and purified using the mi-Gel Extraction Kit (metabion). DNA concentration was measured using the NanoDrop 2000C (Thermo). Purified digests were ligated together for 10 minutes at room temperature or in the fridge overnight. Ligations were performed in a 1:3 vector:insert ratio alongside a water control using T4 DNA Ligase (NEB) in a 20 µl reaction volume as detailed in Table 6.6.

Table 6.6: **Ligation setup**

Reagent	Quantity
10X T4 DNA Ligase Buffer	2 µl
DNA	3 mol insert per 1 mol vector
T4 DNA Ligase	1 µl
water	to 20 µl

If the region of interest was in a vector with compatible restriction sites, the PCR step was skipped and cloning began with restriction digest and gel purification. The remainder of this section details the specific primers, templates, and restriction enzymes used to make the wild type constructs in Table 6.1.

To make the wild type F2H vectors listed in Table 6.1, pMS-RQ-MLXIPL or pCMV6-Entry-14-3-3 as relevant were amplified using the 206548/206549, 206550/206551 or 206557/206558 primer pairs (Table 6.7) to make LID, GRACE, and 14-3-3 amplicons, respectively. PCR products were column purified as described above and double digested with HindIII-HF and SalI-HF followed by a second round of column purification. Full length ChREBP was excised directly from pMS-RQ-MLXIPL by HindIII-HF SalI-HF double digest, followed by gel extraction and purification as described above. peGFP-C1 and pmCherry-C1-LacI target vectors were linearised by HindIII-HF SalI-HF double digestion and column purified. The purified ChREBP, LID, GRACE, and 14-3-3 digests and a water negative control were ligated into linearised peGFP-C1 and pmCherry-C1-LacI.

A similar protocol was followed to make the N1-GRACE-FLAG-Cherry vector used in the CoIP. GRACE was amplified by the 207484/207485 primer pair (Table 6.7) from a pMS-RQ-MLXIPL template. The PCR product was column purified, digested with SalI-HF and BamHI-HF, column purified, and ligated into the purified SalI-HF BamHI-HF digest of the pmCherry-N1 vector (Clontech).

The peGFP-C1-GSM and peGFP-C1-GSM-Cherry vectors used in the FRET-FLIM assay were made using the same process. The GSM was amplified from the pMS-RQ-MLXIPL template using the 206548/207048 and 207071/207072 primer pairs, creating amplicons with and without a 3' STOP codon, respectively. PCR products were column purified and double digested with HindIII-HF and SalI-HF (206548/207048) or XhoI and KpnI-HF (207071/207072). Following column purification, digests were ligated into peGFP-C1 or peGFP-C1-pmCherry (Database ID CL3781, made by Dr. Hari Singh), digested with the respective enzymes and purified.

The mouse ChREBP vectors used in the FRAP, and subcellular localisation assays were made as above. Full-length ChREBP, LID and GRACE were amplified from a peGFP-N1-ChREBP template using 206413/206414, 206362/206363 and 206364/206365 primer pairs, respectively. PCR products were column purified and double digested with SalI-HF and HindIII-HF and ligated into a peGFP-C1 plasmid which had been digested with the same vectors.

The pPAGFP-C1-ChREBP vector used in the photoactivation assay was made by excising mouse ChREBP from peGFP-C1-ChREBP made above using a SalI-HF and HindIII-HF double digest and ligating the gel purified insert into the purified digest of pPAGFP-C1 (Database ID CL2354).

Plasmids pGampac-GSM_I120A, pGampac-GSM_N123A, pGampac-GSM_N124A, pGampac-GSM_W127A, and pGampac-GSM_Q241H were made by digesting pETDuet_14-3-3_GSM vectors containing the same mutations (Database IDs CL4012, CL4013 and CL4014 for pETDuet_14-3-3_GSM N124A, W127A, and Q241H, respectively; pETDuet_14-3-3_GSM I120A and N123A not in database at time of writing; made by Thomas Pysik) with XmaI and XbaI. GSM fragments containing the indicated mutations were gel purified and ligated into a gel purified pGampac-GSM vector digested with the same restriction enzymes.

peGFP-C1-ChREBP mutations of residues between P104 and Q241 were generated by digesting pGampac-GSM containing the relevant mutation with PpuMI and XbaI and ligating the excised

fragment into a peGFP-C1-ChREBP wild type plasmid digested with the same enzymes. The remaining peGFP-C1-ChREBP mutations (i.e. those of residues S253 to T258) were generated using a similar subcloning protocol. pGampac-GSM constructs encoding the relevant mutation were amplified using the 207503/207504 primer pair. 207503 overlaps an endogenous XbaI site in wild type ChREBP and 207504 ends with the upstream half of one of the two EcoRV cut sites in ChREBP (the other EcoRV site is located within the 207503/207504 amplicon). The PCR product was gel purified and digested with XbaI generating a fragment with a 5' XbaI site and a 3' blunt end aligning with and EcoRV site in the full-length protein. The peGFP-C1-ChREBP wild type vector was double digested with EcoRV and XbaI and gel purified. A 1 hour simultaneous ligation and phosphorylation was performed with 1 µl each of insert and vector in a reaction mix containing 2 µl T4 ligase buffer (NEB), 14 µl water, 1 µl T4 Polynukleotide Kinase (PNK) (NEB) and 1 µl T4 DNA ligase. All other point mutations used in the text were generated by inverse PCR mutagenesis, described in detail below.

Table 6.7: **Details of primers used for standard PCR subcloning**

Template	Primer ID	Database ID	Sequence (5'-3')
human ChREBP	LID_HindIII	206548	CTTGAAGCTTCTagc gcgggcggcttgctc
	LID_STOP_SalI	206549	GTGAGTCGACTTAg agccgcttcttgtagtagatgcg ccacttgtg
	GRACE_HindIII	206550	GTTGAAGCTTCTagg gaagatgacctcctg
	GRACE_STOP_SalI	206551	GAGAGTCGACTTAg tccatgaagtcatccagg
	GRACE_Kozak_SalI	207484	ATATGTCGACgccacc ATGGCCagggaagatgac ctcctg
	GRACE_FLAG_BamHI	207485	ATATGGATCCgcCT TATCGTCGTCATCC TTGTAATCgtccatgaa gtcatccagg
	GSM_XhoI	207071	CTTACTCGAGCCagc gcgggcggcttgctc
	GSM_KpnI	207072	GTGTGGTACCgtccat gaagtcatccaggcttggctg
	GSM_STOP_SalI	207048	GAGAGTCGACTTAg tccatgaagtcatccaggcttg gctgcagtg

Table 6.7- *Continued from previous page*

Template	Primer ID	Database ID	Sequence (5'-3')
human 14-3-3	14-3-3_HindIII	206557	GATGAAGCTTCTatg acaatggataaaagtgagctg
	14-3-3_STOP_SalI	206558	CAGAGTCGACTTAa gttctctccctccccag
mouse ChREBP	ChREBP_HindIII	206413	GATGAAGCTTCTatg gcgcgcgcgctggcggatctat ccgtg
	ChREBP_STOP_SalI	206414	CAGAGTCGACttataa tggtctccccagggtgc
	LID_HindIII	206362	GATGAAGCTTCTagc gcgggtggcctgcatcg
	LID_STOP_SalI	206363	CAGAGTCGACTTAg ctggacttacggagccgcttttt gtagtag
	GRACE_HindIII	206364	GATGAAGCTTCTagg gaaggggatttcctg
	GRACE_STOP_Sal	206365	CAGAGTCGACTTAc tccatgaagtcatctaggctg
	MCR6-5_f	207503	agcttctagacctggac
	MCR6-5_r	207504	atctccatgaagtcatctagg

6.3.4 Inverse PCR mutagenesis

Site-directed mutagenesis was performed using an inverse PCR mutagenesis protocol modified from [146]. DNA was amplified as described in Section 6.3.2 using the primers detailed in Table 6.8. After amplification, 1 µl of DpnI (NEB) was added to the PCR tube and the reaction was incubated at 37 °C for 1 hour to digest all of the original plasmid template. 2 µl of DpnI digest was then added to a phosphorylation/ligation reaction mixture consisting of 2 µl T4 ligase buffer (NEB), 14 µl water, 1 µl T4 PNK (NEB) and 1 µl T4 DNA ligase. The phosphorylation/ligation reaction was incubated for 1 hour at room temperature, then transformed into XL10 chemically competent cells.

All point mutations used were performed on a mouse ChREBP template, specifically peGFP-C1-ChREBP and pGampac-GSM. One deletion was made using a pGampac-GSM template (i.e. the pGampac empty vector negative control) and one using a pmCherry-C1-LacI-GRACE template (i.e. the pmCherry-C1-LacI-GRACEΔMCR6 construct).

The pGampac empty vector negative control was made using a slightly modified version of the inverse PCR protocol detailed above in which, following an amplification of the pGampac vector with the 207470/207471 primer pair, 1 µl of both DpnI and SalI-HF were added directly to the PCR mix to simultaneously remove template DNA and create complementary sticky ends on the amplified region. Following a 1 hour digestion, the plasmid was gel purified and ligated for 10 minutes at room temperature before transformation into XL10 chemically competent cells.

Table 6.8: **Details of primers used for PCR mutagenesis**

Primer ID	Database ID	Sequence (5'-3')
ΔMCR6_f	207368	cagtccggcccttcgccc
ΔMCR6_r	207369	aaagcaattgaggtccaggagctgc
ΔGSM_f	207470	atggaggtcgacgagcag
ΔGSM_r	207471	gagagtcgacgaattccggcgatacag
P104A_f	207448	aagtggaagaacttcaaagg
P104A_r	207449	tgcagagaccagcttgccac
W106A_f	207450	aagaacttcaaaggcctc
W106A_r	207451	ggccttgggagagaccagc
W106F_f	207490	ttcaagaacttcaaaggcctcaag
W106L_f	207489	cagaagaacttcaaaggcctcaag
W106_r[a]	207488	cttgggagagaccagcttg
K107A_f	207452	aacttcaaaggcctcaag
K107A_r	207453	ggcccacttgggagagacc
F109A_f	207454	aaaggcctcaagttgctatg
F109A_r	207455	ggcgttcttccacttgggagag
F109L_f	207492	ctgaaaggcctcaagttgctatg
F109L_r	207491	gttcttccacttgggagag
K110A_f	207456	ggcctcaagttgctatgc
K110A_r	207457	ggcgaagttcttccacttggg
L114A_f	207495	gccctatgccgggacaagatc
L114A_r	207494	cttgaggcctttgaagttc
W170F_f	207473	ttcaagcggcgcatcgag
W170L_f	207472	ctgaagcggcgcatcgag
W170_r[a]	207474	gtaattaccctccaggatgacag
K171A_f	207475	gcccggcgcatcgaggtg

Table 6.8- *Continued from previous page*

Mutation	Database ID	Sequence (5'-3')
K171R_f	207476	aggcggcgcatcgaggtg
K171_r[a]	207477	ccagtaattaccctccaggatgacag
KRR171-3A_f	207482	gccgccgccatcgaggtggtgatgc
KRR171-3A_r	207483	ccagtaattaccctccagg
V177A_f	207478	gccatgcgcgaataccacaagtggag
V177I_f	207479	atcatgcgcgaataccacaagtggag
V177_r[a]	207480	cacctcgatgcgccgcttc
S253A_f	207458	gccgacacactcttcacc
S253A_r	207459	gatatcggacaagaagcag
D254A_f	207460	gccacactcttcaccatgacac
D254A_r	207461	ggagatatcggacaagaagc
T255A_f	207462	gccctcttcaccatgacacagc
T255A_r	207463	gtcggagatatcggacaag
L256A_f	207464	gccttcaccatgacacagcc
L256A_r	207465	tgtgtcggagatatcggac
F257A_f	207466	gccaccatgacacagccc
F257A_r	207467	gagtgtgtcggagatatcg
T258A_f	207468	gccatgacacagcccagtc
T258A_r	207469	gaagagtgtgtcggagatatcg
SDTLFT253-8A_f	207487	gccgccgccatgacacagcccagtc
SDTLFT253-8A_r	207486	ggcggcggcgatatcggacaagaagcagtc

[a] Indicates a reverse primer compatible with more than one forward primer.

6.3.5 Transformation

After ligation, DNA was transformed into XL10 *E. coli* competent cells using the following protocol. 10 µl ligation reaction was mixed with 50 µl XL10 competent cells in pre-cooled eppendorfs on ice. The cell mixture was incubated on ice for 15 minutes then heat shocked at 42 °C for 35 seconds. Cells were placed on ice for 5 min. 250 µl room temperature lysogeny broth (LB) was then added to the cells which were then incubated in a shaking heat block at 37 °C and 500 rpm. After 1 hour, the transformation mixture was centrifuged at 6000 xg for 1 minute. The supernatant was then decanted and the pellet resuspended in the remaining LB. Resuspended pellet was then

spread on LB plates containing 50 µg/ml ampicillin or 30 µg/ml kanamycin, as necessary. Plates were incubated overnight at 37 °C then stored in the fridge.

6.3.6 Plasmid preparation

When many ligations and transformations were run in parallel, positive clones were screened for by a colony PCR of the transformation plate. Colony PCRs were preformed in 20 µl reaction volumes with colonies scraped and dipped in the PCR mix, one colony per tube. PCRs were performed using Taq polymerase (NEB). The reaction mix and thermal cycler settings used are detailed in Tables 6.9 and 6.10, respectively.

Table 6.9: **Colony PCR reaction setup**

Reagent	Quantity
10X Standard Taq Reaction Buffer	10 µl
10 mM dNTPs	20 µl
10 µM forward primer	50 µl
10 µM reverse primer	50 µl
Taq DNA Polymerase	20 µl
water	700 µl
XL10 cells (template DNA)	fraction of colony[a]

[a] Transformed XL10 cells were added by dipping a scraped colony into the reaction mix.

Table 6.10: **Colony PCR protocol used**

Phase	Cycles	Step	Temperature	Time
initial	1	denaturation	95 °C	1 minute
main	30	denaturation	95 °C	30 seconds
		annealing	58 °C	30 seconds/kb
		extension	68 °C	1 min/kb
final	1	extension	68 °C	5 minutes
hold			12 °C	

Colonies identified as positive by colony PCR were amplified in a 37 °C shaker overnight in 2 ml LB with 50 µg/ml ampicillin or 30 µg/ml kanamycin, as relevant and miniprepped using the mi-Plasmid Miniprep Kit (metabion).

When a small amount of ligations and transformations were being run simultaneously, colonies were amplified and miniprepped without colony PCR screening. After miniprepping, positive constructs were assessed for successful ligation (or successful inverse PCR deletion) by restriction digest.

Positive constructs identified by colony PCR or restriction digest were verified by sequencing. Constructs where positives could not be distinguished by colony PCR or restriction digest (i.e. point mutants generated by inverse PCR) were directly screened by sequencing. Sequencing was performed by GATC/Eurofins.

Clones that were verified by sequencing were grown in a 37 °C shaker overnight in 80 ml LB and midiprepped using the PureYield™ Plasmid Midiprep System (Promega), following manufacturer instructions.

6.4 Western Blotting

6.4.1 Sample preparation

Cells were lysed in ice cold RIPA buffer freshly supplemented with Complete EDTA free protease inhibitor (Roche) and 1 mM phenylmethylsulfonyl fluoride (PMSF) (SERVA). Lysate was incubated for 15 minutes on ice, then spun down at 20000 xg for 10 min at 4 °C. The pellet was discarded and the protein concentration in the supernatant was measured using the Pierce® BCA Protein Assay Kit (Thermo Scientific). Samples were boiled at 95 °C in laemmli for 5 minutes and either used immediately or stored at -20 °C.

6.4.2 SDS-PAGE

For blots in which a wide range of protein sizes needed to be resolved, 4–20% Mini-PROTEAN® TGX Stain-Free™ Protein Gels (Biorad) were used, otherwise sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels were made by hand and stored in the fridge or used immediately. Table 6.11 shows the setup for the 10% and 12% gels used in this thesis. Samples were run along with the PageRuler™ Plus Prestained Protein Ladder (Thermo), on SDS-PAGE gels at 150 V for 90 min.

Table 6.11: **SDS-PAGE gel setup**

Gel type	Reagent	10% gel, 12% gel
resolving gel	water	3.8 ml, 3.2 ml
	30% Acrylamide	3.4 ml, 4 ml
	1.5 M Tris pH 8.8	2.6 ml
	10% SDS	100 µl
	10% APS	100 µl
	TEMED	10 µl
stacking gel	water	3 ml
	30% Acrylamide	670 µl
	0.5 M Tris pH 6.8	1.25 ml
	10% SDS	50 µl
	10% APD	50 µl
	TEMED	5 µl

6.4.3 Transfer, antibody staining and imaging

Protein was transferred from SDS-PAGE gels to Amersham™ Protran® nitrocellulose membranes (GE Healthcare) in transfer buffer on ice at 100 V for 1 hour. Blocking was done for 1 hour

in tris-buffered saline 0.2% Tween-20 (TBST) with 5% milk freshly added. Primary antibody incubation was performed in the cold room overnight (4-7 °C) with the indicated concentrations of antibody diluted in TBST with 5% milk freshly added. The following primary antibodies were used in the indicated dilutions: anti-Cherry 1:2000 (ab167453, Abcam), anti-GAPDH 1:2500 (ab9485, Abcam), anti-GFP 1:1000 (#2955, Cell signalling), anti-GFP 1:1000 (G3, in-house), anti-Myc 1:200 (9E10, courtesy of the Helmholz Center), anti-Tubulin 1:1000 (T9026 clone DM1A, Sigma). After primary antibody incubation, membranes were washed three times in TBST at room temperature, 5 minutes per wash. They were then incubated for 30 minutes in a 1:10000 dilution of horseradish peroxidase (HRP) conjugated secondary antibody (BioRad) in TBST. This was followed by three TBST washes at room temperature, 5 minutes per wash and one quick phosphate-buffered saline (PBS) rinse. The membranes were then imaged using Immobilion Western HRP substrate (Milipore).

6.4.4 Membrane stripping

When a membrane was to be probed with more than one primary antibody, as in the case of a protein of interest versus a loading control, the membrane was stripped immediately after imaging by two 10 minute incubations in stripping buffer, followed by two 10 minute PBS washes, and a 5 minute TBST wash. After stripping, the membrane was blocked and incubated in primary antibody as described above.

6.5 Knock-in cell line generation

"Wild type" U2OS 2B2 cells were transfected with either the peGFP-C1 empty vector (Database ID CL0530) or peGFP-C1-LID (Database ID CL3929). Forty-eight hours after transfection, the cells were put on G418 selection (400 µg/ml). After 36 hours of G418 selection, GFP positive cells were monoclonalised by sorting onto 96 well plates, one cell per well using the BD FACSAria Fusion cell sorter operated by Dr. Lisa Richter. Cells were expanded and tested for mycoplasma contamination before freezing down. Successful integration of GFP and GFP-LID was verified by western blotting.

6.6 Co-immunoprecipitation (CoIP)

Pulldowns were performed using anti-GFP-sepharose beads made by conjugating recombinantly expressed anti GFP-nanobody [131], to NHS-activated Sepharose^{TM}beads (GE Healthcare). Nanobody purification and bead conjugation was performed by Christiane Kotthoff.

The indicated U2OS-derived cell lines (stably expressing either GFP alone or GFP-LID) were plated on 6-well plates and transfected the following day. The media was changed the day after transfection. Two days after transfection cells were scraped in 50 µl RIPA. The lysate was then

sonicated in two 30 second pulses with a 30 second rest at 90% amplitude in a Q800R sonicator (Qsonica). Sonicated lysate was spun down at 20000 xg for 10 min at 4 °C. The supernatant was then split into two fractions to be bound and washed in the presence or absence of 100 µM G6P (Cayman), respectively. Supernatant was loaded onto anti-GFP-sepharose beads equilibrated in wash buffer either containing or lacking G6P, as relevant. A small fraction of supernatant not added to the beads was boiled in laemmli for western blotting (input fraction). Wash buffer was added to the supernatant–bead solution to a total final volume of 300 µl. The lysate–bead suspension was incubated on a rotator in the fridge for 1 hour. The sample was then spun down at 2000 xg for 2 min at 4 °C. Supernatant was removed and beads were washed three times in 500 µl wash buffer and then boiled in laemmli (bound fraction).

6.7 Fluorescence two-hybrid (F2H) assays

For the negative F2H assay, U2OS 2B2 cells were plated on 8-well Nunc™ Lab-Tek™ chambered cover glasses (Thermo) in 2.5 mM (low) glucose phenol-free DMEM. Twenty-four hours after plating, the cells were transfected with the indicated fluorescent proteins. Sixteen hours after transfection, a 2.5 M glucose stock solution or equivalent volume of water was added to final concentrations of 25 and 2.5 mM glucose, respectively. One hour after glucose addition, the cells were washed in PBS then fixed and permeabilised for 10 min in a 1:1 solution of methanol:acetone at -20 °C. Cells were then washed three times in PBS, stained for 10 minutes in 1 µg/ml Hoechst 33342 (Thermo) at room temperature, washed three more times and then imaged immediately or stored in the fridge for later imaging. Negative F2H assays were performed with help from Giorgi Beroshvili.

For the positive F2H assay, U2OS 2B2 GFP-LID cell lines were plated on black Cycloolefin-coated 96-well plates (Greiner) in 25 mM (high) glucose phenol-free DMEM. Twenty-four hours after plating, the cells were transfected with the indicated fluorescent proteins. Twenty-four hours after transfection, the glucose concentration in the media was diluted out by the addition of glucose-free DMEM, or an equivalent volume of high glucose DMEM was added, to final concentrations of 2.5 and 25 mM glucose, respectively. Two hours after glucose dilution, the cells were washed in PBS then fixed and permeabilised for 10 min in a 1:1 solution of methanol:acetone at -20 °C. Cells were then washed three times in PBS, stained for 10 minutes in 1 µg/ml Hoechst 33342 (Thermo) at room temperature, washed three more times and then imaged immediately or stored in the fridge for later imaging.

Imaging was performed using a Zeiss AxioObserver Z1 microscope equipped with a Yokogawa confocal spinning disk unit and recorded using a Hamamatsu Orca Flash 4.0 camera. Images were taken using a Zeiss Plan-Apochromat 40x/0.95 Korr air objective. An ImageJ plugin was written for automated quantification (Appendix 7.4.1).

6.8 Fluorescence resonance energy transfer-fluorescence lifetime imaging (FRET-FLIM)

832/13 cells were plated on 4-well Nunc™ Lab-Tek™ chambered cover glasses (Thermo) in 11 mM glucose RPMI. Twenty-four hours after plating, cells were transfected with peGFP-C1-GSM (Database ID CL3931) or peGFP-C1-GSM-Cherry (Database ID CL3933). Twenty-four hours after transfection, cells were changed to either 2.5 mM (low) or 25 mM (high) glucose DMEM, supplemented with 25 mM HEPES to allow CO_2 independence. Microscopy was begun at least one hour after media change. Imaging was performed on an inverted Leica SP8X confocal microscope using an HC PL APO 40x/1.30 oil objective. FLIM was performed by time-correlated single-photon counting (TCSPC) using a white light laser with a repetition rate of 80 MHz. Samples were excited at 488 nm and emission detected between 500 and 550 nm using a Hybrid photodetector. The pixel dwell time was set to 3.1625 µs/pixel. Each acquisition recorded a 512x512 pixel field of view. 100 frame repetitions were performed per acquisition. Lifetime analysis was done on individual cells called from manually drawn regions of interest using LASX software (Leica). Instrument response function reconvolution fitting of the lifetime decay was used to determine the amplitude-weighted lifetime. The number of lifetimes for each condition was determined using the FLIM module of LASX by empirical determination of which lifetime produced the optimal χ^2 values and residuals. In this way, the peGFP-C1-GSM negative control cells were found to fit best with one lifetime overall, while the peGFP-C1-GSM-Cherry cells fit best with two lifetimes overall. Images were recorded in three separate experiments on three separate days. 234 total cells were recorded in low glucose GFP-GSM. 240 total cells were recorded in high glucose GFP-GSM. 223 total cells were recorded in low glucose GFP-GSM-Cherry. 245 total cells were recorded in high glucose GFP-GSM-Cherry.

6.9 Subcellular localisation of ChREBP proteins

832/13 cells were plated 200 µl/well on black Cycloolefin-coated 96-well plates (Greiner) in 11 mM RPMI. Twenty-four hours after plating, cells were transfected with the indicated GFP-tagged constructs using Lipofectamine® 2000 Transfection Reagent (Thermo) following manufacturer protocol. Twenty-four hours after transfection, media was changed to either 2.5 mM (low) or 27.5 mM (high) glucose RPMI. Twenty-four hours after media change, cells were washed in PBS and fixed in 4% paraformaldehyde for 10 minutes. Cells were then washed three times in PBS and permeabilised for 10 minutes in PBS containing 0.1% Triton X-100 (Sigma). Cells were then washed three times in PBS and stained for half hour in a PBS solution containing 1 µg/ml Hoechst 33342 (Thermo) and 12.5 ng/ml CellMask® Deep Red plasma membrane stain (Thermo).

Imaging was performed using a Zeiss AxioObserver Z1 microscope equipped with a Yokogawa confocal spinning disk unit and recorded using a Hamamatsu Orca Flash 4.0 camera. Images were recorded automatically using a Zeiss Plan-Apochromat 20x/0.8 air objective. Each automated imaging run recorded 60 non-overlapping tiles per well (30% of each well) with focus maintained

via one round of definite focus stabilisation per well. An ImageJ plugin was written for automated quantification (Appendix 7.4.2). Imaging was performed in six separate experiments.

6.10 Fluorescence recovery after photobleaching (FRAP) imaging

832/13 cells were prepared for imaging as in Section 6.8, with cells transfected with peGFP-C1-ChREBP (Database ID CL4181). Leptomycin B (LMB) was purchased from Sigma. In wells where LMB was used, it was added in a 1:1000 dilution (to a final concentration of 9.25 nM) immediately before the respective well was imaged. ChREBP became nuclear localised in all cells within the first hour after LMB addition. Imaging was performed by Dr. Imke Mandemaker on an inverted Leica SP8X confocal microscope using an HC PL APO 40x/1.30 oil objective. Images were taken using a white light laser at 1400 Hz with a line averaging of 1. Nuclear peGFP-ChREBP was bleached at 100% laser power using a 488 nm laser covering either the whole nucleus or a line through the nucleus, as indicated. For a whole nuclear bleach, intensity was detected using a photomultiplier tube (PMT) and recorded every 0.15 seconds for a total of 800 frames. For a linear nuclear bleach, intensity was detected using a PMT and recorded every 0.2 seconds for a total of 500 frames.

6.11 Photoactivation of ChREBP

832/13 cells were prepared for imaging as in Section 6.8, with cells doubly transfected with pPAGFP-C1-ChREBP (Database ID CL4257) and pmCherry-C1-LacI (Database ID CL4236), the latter to identify the nuclei of transfected cells. Imaging was performed by Dr. Imke Mandemaker on an inverted Leica SP8X confocal microscope using an HC PL APO 40x/1.30 oil objective. Images were taken using a white light laser at 1400 Hz with a line averaging of 1. Nuclear pPAGFP-ChREBP was photoactivated using a 405 nm laser. Intensity was detected using a PMT and recorded every 0.37 seconds for a total of 200 frames.

6.12 Luciferase-based transcription reporter assay

Luciferase assays were performed as described previously [69]. 832/13 cells were plated in 11 mM glucose RPMI on a white 384-well plate (Greiner) 67 µl/well. 24 hours after plating, cells were transfected with equimolar amounts of pRL-TK (Promega, containing a constitutive TK-promoter driven Renilla luciferase), pG5luc (Promega, containing a Gal4 UAS-promoter driven Firefly luciferase), and a pGampac vector (a gift from Professor Lawrence Chan and Professor Benny Hung-Junn Chang) expressing the indicated Gal4 fusion protein. 24 hours after transfection, the media was aspirated and replaced with RPMI containing the indicated concentration of glucose.

Twenty-four hours after media change, luciferase expression was determined using the Dual-Glo® Luciferase Assay System (Promega) and read in an Infinite M1000 PRO plate reader (Tecan) using a 1 second integration time and a 15 millisecond settle time.

6.13 Quantification and statistics

Quantification of F2H and subcellular localisation images were performed using Fiji ImageJ [147]. The code of the Jython plugins written for analysis is included in the Appendix 7.4. Quantification of FRET-FLIM, FRAP, and photoactivation experiments was done using LASX software (Leica). Graphing and statistics were performed in R [148] using the RStudio Integrated Development Environment [149] with the extrafont [150], ggplot2 [151], Rmisc [152], rstatix [153] and xlsx [154] packages. Error bars are standard error of the mean and p-values were calculated using the two-tailed t-test and adjusted when necessary using hommel adjustment for multiple comparisons.

6.14 Buffers

4x Laemmli
200 mM Tris-HCl pH 6.8, 277 mM SDS, 3 mM Bromophenol blue, 40% glycerol, 20% β-Mercaptoethanol

LB
10 g Tryptone, 5 g yeast extract, 10 g NaCl, dissolved in water to a total volume of 1 L with the pH adjusted to 7.0

PBS
137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4

RIPA buffer
50mM Tris-HCl pH8, 150mM NaCl, 1mM EDTA, 1% NP-40, 0.1% SDS, 0.5% Sodium Deoxycholate, Complete EDTA free protease inhibitor (Sigma) and PMSF (Sigma) to 1 mM were added immediately before use.

Stripping buffer
200 mM glycine, 3.5 mM SDS, 1% Tween20, pH adjusted to 2.2

10x tris-buffered saline (TBS)
200 mM Tris pH 7.6, 1.5 M NaCl

TBST
1x TBS, 0.2% Tween20

Transfer Buffer

25 mM Tris, 192 mM glycine, 20% (v/v) methanol

Wash Buffer
10 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA, Complete EDTA free protease inhibitor (Sigma) and PMSF to 1 mM were added immediately before use. G6P was also added to a final concentration of 100 µM immediately before use where indicated.

References

[1] H.-S. Han, G. Kang, J. S. Kim, B. H. Choi, and S.-H. Koo, "Regulation of glucose metabolism from a liver - centric perspective," *Experimental & Molecular Medicine*, vol. 48, no. 3, p. e218, 2016.

[2] Z. Song, A. M. Xiaoli, and F. Yang, "Regulation and metabolic significance of de novo lipogenesis in adipose tissues," *Nutrients*, vol. 10, Sep 2018.

[3] P. Ortega-Prieto and C. Postic, "Carbohydrate sensing through the transcription factor chrebp," *Front Genet*, vol. 10, p. 472, 2019.

[4] A. Moslehi and Z. Hamidi-Zad, "Role of srebps in liver diseases: A mini-review," *J Clin Transl Hepatol*, vol. 6, pp. 332–338, Sep 2018.

[5] M. B. Patel, N. A. Oza, I. S. Anand, S. S. Deshpande, and C. N. Patel, "Liver x receptor: a novel therapeutic target," *Indian J Pharm Sci*, vol. 70, no. 2, pp. 135–44, 2008.

[6] B. Wang and P. Tontonoz, "Liver x receptors in lipid signalling and membrane homeostasis," *Nat Rev Endocrinol*, vol. 14, pp. 452–463, 08 2018.

[7] Y. Wang, T. Nakajima, F. J. Gonzalez, and N. Tanaka, "Ppars as metabolic regulators in the liver: Lessons from liver-specific ppar-null mice," *Int J Mol Sci*, vol. 21, Mar 2020.

[8] J. J. Repa, G. Liang, J. Ou, Y. Bashmakov, J. M. Lobaccaro, I. Shimomura, B. Shan, M. S. Brown, J. L. Goldstein, and D. J. Mangelsdorf, "Regulation of mouse sterol regulatory element-binding protein-1c gene (srebp-1c) by oxysterol receptors, lxralpha and lxrbeta," *Genes Dev*, vol. 14, pp. 2819–30, Nov 2000.

[9] J.-Y. Cha and J. J. Repa, "The liver x receptor (lxr) and hepatic lipogenesis: The carbohydrate-response element-binding protein is a target gene of lxr," *Journal of Biological Chemistry*, vol. 282, no. 1, pp. 743–751, 2007.

[10] C. Bindesbøll, Q. Fan, R. C. Nørgaard, L. MacPherson, H.-B. Ruan, J. Wu, T. Å. Pedersen, K. R. Steffensen, X. Yang, J. Matthews, S. Mandrup, H. I. Nebb, and L. M. Grønning-Wang, "Liver x receptor regulates hepatic nuclear o-glcnac signaling and carbohydrate responsive element-binding protein activity," *Journal of Lipid Research*, vol. 56, no. 4, pp. 771–785, 2015.

[11] J. Meng, M. Feng, W. Dong, Y. Zhu, Y. Li, P. Zhang, L. Wu, M. Li, Y. Lu, H. Chen, X. Liu, Y. Lu, H. Sun, and X. Tong, "Identification of hnf-4α as a key transcription factor to promote chrebp expression in response to glucose," *Sci Rep*, vol. 6, p. 23944, Mar 2016.

[12] A. Jalil, T. Bourgeois, L. Ménégaut, L. Lagrost, C. Thomas, and D. Masson, "Revisiting the role of lxrs in pufa metabolism and phospholipid homeostasis," *Int J Mol Sci*, vol. 20, Aug 2019.

[13] M. A. Herman, O. D. Peroni, J. Villoria, M. R. Schon, N. A. Abumrad, M. Bluher, S. Klein, and B. B. Kahn, "A novel chrebp isoform in adipose tissue regulates systemic glucose metabolism," *Nature*, vol. 484, no. 7394, pp. 333–338, 2012.

[14] J. Xu, B. Christian, and D. B. Jump, "Regulation of rat hepatic l-pyruvate kinase promoter composition and activity by glucose, n-3 polyunsaturated fatty acids, and peroxisome proliferator-activated receptor-alpha agonist," *J Biol Chem*, vol. 281, pp. 18351–62, Jul 2006.

[15] K. Iizuka, W. Wu, Y. Horikawa, M. Saito, and J. Takeda, "Feedback looping between chrebp and pparα in the regulation of lipid metabolism in brown adipose tissues," *Endocr J*, vol. 60, no. 10, pp. 1145–53, 2013.

[16] T. Yoshikawa, T. Ide, H. Shimano, N. Yahagi, M. Amemiya-Kudo, T. Matsuzaka, S. Yatoh, T. Kitamine, H. Okazaki, Y. Tamura, M. Sekiya, A. Takahashi, A. H. Hasty, R. Sato, H. Sone, J.-I. Osuga, S. Ishibashi, and N. Yamada, "Cross-talk between peroxisome proliferator-activated receptor (ppar) alpha and liver x receptor (lxr) in nutritional regulation of fatty acid metabolism. i. ppars suppress sterol regulatory element binding protein-1c promoter through inhibition of lxr signaling," *Mol Endocrinol*, vol. 17, pp. 1240–54, Jul 2003.

[17] R. Brunmeir and F. Xu, "Functional regulation of ppars through post-translational modifications," *Int J Mol Sci*, vol. 19, 06 2018.

[18] N. Liang, T. Jakobsson, R. Fan, and E. Treuter, "The nuclear receptor-co-repressor complex in control of liver metabolism and disease," *Front Endocrinol (Lausanne)*, vol. 10, p. 411, 2019.

[19] M. Balanarasimha, A. M. Davis, F. L. Soman, S. D. Rider, Jr, and H. A. Hostetler, "Ligand-regulated

heterodimerization of peroxisome proliferator-activated receptor α with liver x receptor α," *Biochemistry*, vol. 53, pp. 2632–43, Apr 2014.

[20] M. Boergesen, T. Å. Pedersen, B. Gross, S. J. van Heeringen, D. Hagenbeek, C. Bindesbøll, S. Caron, F. Lalloyer, K. R. Steffensen, H. I. Nebb, J.-Å. Gustafsson, H. G. Stunnenberg, B. Staels, and S. Mandrup, "Genome-wide profiling of liver x receptor, retinoid x receptor, and peroxisome proliferator-activated receptor α in mouse liver reveals extensive sharing of binding sites," *Mol Cell Biol*, vol. 32, pp. 852–67, Feb 2012.

[21] K. S. Miyata, S. E. McCaw, H. V. Patel, R. A. Rachubinski, and J. P. Capone, "The orphan nuclear hormone receptor lxr alpha interacts with the peroxisome proliferator-activated receptor and inhibits peroxisome proliferator signaling," *J Biol Chem*, vol. 271, pp. 9189–92, Apr 1996.

[22] L. Yue, F. Ye, C. Gui, H. Luo, J. Cai, J. Shen, K. Chen, X. Shen, and H. Jiang, "Ligand-binding regulation of lxr/rxr and lxr/ppar heterodimerizations: Spr technology-based kinetic analysis correlated with molecular dynamics simulation," *Protein Sci*, vol. 14, pp. 812–22, Mar 2005.

[23] A. Fernández-Alvarez, M. S. Alvarez, R. Gonzalez, C. Cucarella, J. Muntané, and M. Casado, "Human srebp1c expression in liver is directly regulated by peroxisome proliferator-activated receptor alpha (pparalpha)," *J Biol Chem*, vol. 286, pp. 21466–77, Jun 2011.

[24] C. Améen, U. Edvardsson, A. Ljungberg, L. Asp, P. Akerblad, A. Tuneld, S.-O. Olofsson, D. Lindén, and J. Oscarsson, "Activation of peroxisome proliferator-activated receptor alpha increases the expression and activity of microsomal triglyceride transfer protein in the liver," *J Biol Chem*, vol. 280, pp. 1224–9, Jan 2005.

[25] H. Niwa, K. Iizuka, T. Kato, W. Wu, H. Tsuchida, K. Takao, Y. Horikawa, and J. Takeda, "Chrebp rather than shp regulates hepatic vldl secretion," *Nutrients*, vol. 10, no. 3, p. 321, 2018.

[26] A. Iroz, A. Montagner, F. Benhamed, F. Levavasseur, A. Polizzi, E. Anthony, M. Régnier, E. Fouché, C. Lukowicz, M. Cauzac, E. Tournier, M. Do-Cruzeiro, M. Daujat-Chavanieu, S. Gerbal-Chalouin, V. Fauveau, S. Marmier, A.-F. Burnol, S. Guilmeau, Y. Lippi, J. Girard, W. Wahli, R. Dentin, H. Guillou, and C. Postic, "A specific chrebp and pparα cross-talk is required for the glucose-mediated fgf21 response," *Cell Rep*, vol. 21, pp. 403–416, Oct 2017.

[27] C. Tezze, V. Romanello, and M. Sandri, "Fgf21 as modulator of metabolism in health and disease," *Front Physiol*, vol. 10, p. 419, 2019.

[28] X. Yuan, T. C. Ta, M. Lin, J. R. Evans, Y. Dong, E. Bolotin, M. A. Sherman, B. M. Forman, and F. M. Sladek, "Identification of an endogenous ligand bound to a native orphan nuclear receptor," *PLoS One*, vol. 4, no. 5, p. e5609, 2009.

[29] C. P. Martinez-Jimenez, I. Kyrmizi, P. Cardot, F. J. Gonzalez, and I. Talianidis, "Hepatocyte nuclear factor 4alpha coordinates a transcription factor network regulating hepatic fatty acid metabolism," *Mol Cell Biol*, vol. 30, pp. 565–77, Feb 2010.

[30] H. Lu, "Crosstalk of hnf4α with extracellular and intracellular signaling pathways in the regulation of hepatic metabolism of drugs and lipids," *Acta Pharm Sin B*, vol. 6, pp. 393–408, Sep 2016.

[31] S. J. Burke, J. J. Collier, and D. K. Scott, "camp opposes the glucose-mediated induction of the l-pk gene by preventing the recruitment of a complex containing chrebp, hnf4α, and cbp," *The FASEB Journal*, vol. 23, no. 9, pp. 2855–2865, 2009.

[32] Q. Fan, R. C. Nørgaard, I. Grytten, C. M. Ness, C. Lucas, K. Vekterud, H. Soedling, J. Matthews, R. B. Lemma, O. S. Gabrielsen, C. Bindesbøll, S. M. Ulven, H. I. Nebb, L. M. Grønning-Wang, and T. Sæther, "Lxrα regulates chrebpα transactivity in a target gene-specific manner through an agonist-modulated lbd-lid interaction," *Cells*, vol. 9, May 2020.

[33] E. W. SUTHERLAND and T. W. RALL, "Fractionation and characterization of a cyclic adenine ribonucleotide formed by tissue particles," *J Biol Chem*, vol. 232, pp. 1077–91, Jun 1958.

[34] T. Yamamoto, H. Shimano, N. Inoue, Y. Nakagawa, T. Matsuzaka, A. Takahashi, N. Yahagi, H. Sone, H. Suzuki, H. Toyoshima, and N. Yamada, "Protein kinase a suppresses sterol regulatory element-binding protein-1c expression via phosphorylation of liver x receptor in the liver," *J Biol Chem*, vol. 282, pp. 11687–95, Apr 2007.

[35] H. Yang and L. Yang, "Targeting camp/pka pathway for glycemic control and type 2 diabetes therapy," *J Mol Endocrinol*, vol. 57, pp. R93–R108, 08 2016.

[36] S. Ros and A. Schulze, "Balancing glycolytic flux: the role of 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatases in cancer metabolism," *Cancer Metab*, vol. 1, p. 8, Feb 2013.

[37] R. Bartrons, L. Hue, E. Van Schaftingen, and H. G. Hers, "Hormonal control of fructose 2,6-bisphosphate concentration in isolated rat hepatocytes," *Biochem J*, vol. 214, pp. 829–37, Sep 1983.

[38] C. Arden, S. J. Tudhope, J. L. Petrie, Z. H. Al-Oanzi, K. S. Cullen, A. J. Lange, H. C. Towle, and L. Agius, "Fructose 2,6-bisphosphate is essential for glucose-regulated gene transcription of glucose-6-phosphatase and other chrebp target genes in hepatocytes," *Biochemical Journal*, vol. 443, no. 1, pp. 111–123, 2012.

[39] D. G. Hardie, "Minireview: the amp-activated protein kinase cascade: the key sensor of cellular energy status," *Endocrinology*, vol. 144, pp. 5179–83, Dec 2003.

[40] D. G. Hardie, S. A. Hawley, and J. W. Scott, "Amp-activated protein kinase–development of the energy sensor concept," *J Physiol*, vol. 574, pp. 7–15, Jul 2006.

[41] C. Longuet, E. M. Sinclair, A. Maida, L. L. Baggio, M. Maziarz, M. J. Charron, and D. J. Drucker, "The glucagon receptor is required for the adaptive metabolic response to fasting," *Cell Metab*, vol. 8, pp. 359–71, Nov 2008.

[42] M. Grabacka, M. Pierzchalska, M. Dean, and K. Reiss, "Regulation of ketone body metabolism and the role of pparα," *Int J Mol Sci*, vol. 17, Dec 2016.

[43] H. Yadav, S. Devalaraja, S. T. Chung, and S. G. Rane, "Tgf-β1/smad3 pathway targets pp2a-ampk-foxo1 signaling to regulate hepatic gluconeogenesis," *J Biol Chem*, vol. 292, pp. 3420–3432, 02 2017.

[44] X. Gu, Y. Yan, S. J. Novick, A. Kovach, D. Goswami, J. Ke, M. H. E. Tan, L. Wang, X. Li, P. W. de Waal, M. R. Webb, P. R. Griffin, H. E. Xu, and K. Melcher, "Deconvoluting amp-activated protein kinase (ampk) adenine nucleotide binding and sensing," *J Biol Chem*, vol. 292, pp. 12653–12666, 07 2017.

[45] J. J. Holst, "The physiology of glucagon-like peptide 1," *Physiological Reviews*, vol. 87, pp. 1409–1439, Oct 2007.

[46] A. R. Saltiel and C. R. Kahn, "Insulin signalling and the regulation of glucose and lipid metabolism," *Nature*, vol. 414, pp. 799–806, Dec 2001.

[47] R. L. Hurley, L. K. Barré, S. D. Wood, K. A. Anderson, B. E. Kemp, A. R. Means, and L. A. Witters, "Regulation of amp-activated protein kinase by multisite phosphorylation in response to agents that elevate cellular camp," *J Biol Chem*, vol. 281, pp. 36662–72, Dec 2006.

[48] G. A. Rutter, G. Da Silva Xavier, and I. Leclerc, "Roles of 5'-amp-activated protein kinase (ampk) in mammalian glucose homoeostasis," *Biochem J*, vol. 375, pp. 1–16, Oct 2003.

[49] I. Leclerc and G. A. Rutter, "Amp-activated protein kinase: a new beta-cell glucose sensor?: Regulation by amino acids and calcium ions," *Diabetes*, vol. 53 Suppl 3, pp. S67–74, Dec 2004.

[50] R. H. Eckel, S. M. Grundy, and P. Z. Zimmet, "The metabolic syndrome," *Lancet*, vol. 365, no. 9468, pp. 1415–28, 2005.

[51] S. Azhar, "Peroxisome proliferator-activated receptors, metabolic syndrome and cardiovascular disease," *Future Cardiol*, vol. 6, pp. 657–91, Sep 2010.

[52] P. Joseph, D. Leong, M. McKee, S. S. Anand, J.-D. Schwalm, K. Teo, A. Mente, and S. Yusuf, "Reducing the global burden of cardiovascular disease, part 1: The epidemiology and risk factors," *Circ Res*, vol. 121, pp. 677–694, Sep 2017.

[53] P. Saeedi, I. Petersohn, P. Salpea, B. Malanda, S. Karuranga, N. Unwin, S. Colagiuri, L. Guariguata, A. A. Motala, K. Ogurtsova, J. E. Shaw, D. Bright, R. Williams, and IDF Diabetes Atlas Committee, "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the international diabetes federation diabetes atlas, 9th edition," *Diabetes Res Clin Pract*, vol. 157, p. 107843, Nov 2019.

[54] K. Blaslov, F. S. Naranđa, I. Kruljac, and I. P. Renar, "Treatment approach to type 2 diabetes: Past, present and future," *World J Diabetes*, vol. 9, pp. 209–219, Dec 2018.

[55] A. V. Haas and M. E. McDonnell, "Pathogenesis of cardiovascular disease in diabetes," *Endocrinol Metab Clin North Am*, vol. 47, pp. 51–63, 03 2018.

[56] D. G. Simons-Morton, "The glycemia-cardiovascular disease hypothesis," *Clin Transl Sci*, vol. 1, pp. 185–6, Dec 2008.

[57] R. A. Defronzo, "Banting lecture. from the triumvirate to the ominous octet: a new paradigm for the treatment of type 2 diabetes mellitus," *Diabetes*, vol. 58, pp. 773–95, Apr 2009.

[58] F. Giorgino, P. D. Home, and J. Tuomilehto, "Glucose control and vascular outcomes in type 2 diabetes: Is the picture clear?," *Diabetes Care*, vol. 39 Suppl 2, pp. S187–95, Aug 2016.

[59] P. Singh and D. M. Irwin, "Contrasting patterns in the evolution of vertebrate mlx interacting protein (mlxip) and mlx interacting protein-like (mlxipl) genes," *PLoS ONE*, vol. 11, no. 2, 2016.

[60] R. C. Edgar, "Muscle: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Res*, vol. 32, no. 5, pp. 1792–7, 2004.

[61] A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, and G. J. Barton, "Jalview version 2–a multiple sequence alignment editor and analysis workbench," *Bioinformatics*, vol. 25, pp. 1189–91, May 2009.

[62] E. Havula, M. Teesalu, T. Hyötyläinen, H. Seppälä, K. Hasygar, P. Auvinen, M. Orešič, T. Sandmann, and V. Hietakangas, "Mondo/chrebp-mlx-regulated transcriptional network is essential for dietary sugar tolerance in drosophila," *PLOS Genetics*, vol. 9, no. 4, p. e1003438, 2013.

[63] J. Mattila and V. Hietakangas, "Regulation of carbohydrate energy metabolism in drosophila melanogaster," *Genetics*, vol. 207, pp. 1231–1253, 12 2017.

[64] A. V. Samuelson, C. E. Carr, and G. Ruvkun, "Gene activities that mediate increased life span of c. elegans insulin-like signaling mutants," *Genes Dev*, vol. 21, pp. 2976–94, Nov 2007.

[65] D. W. Johnson, J. R. Llop, S. F. Farrell, J. Yuan, L. R. Stolzenburg, and A. V. Samuelson, "The caenorhabditis elegans myc-mondo/mad complexes integrate diverse longevity signals," *PLoS Genet*, vol. 10, p. e1004278, Apr 2014.

[66] I. Leclerc, G. A. Rutter, G. Meur, and N. Noordeen, "Roles of ca2+ ions in the control of chrebp nuclear translocation," *Journal of Endocrinology*, vol. 213, no. 2, pp. 115–122, 2012.

[67] M. J. Plevin, M. M. Mills, and M. Ikura, "The lxxll motif: a multifunctional binding sequence in transcriptional regulation," *Trends Biochem Sci*, vol. 30, pp. 66–9, Feb 2005.

[68] L. G. McFerrin and W. R. Atchley, "A novel n-terminal domain may dictate the glucose response of mondo proteins," *PLoS ONE*, vol. 7, no. 4, p. e34803, 2012.

[69] M. V. Li, B. Chang, M. Imamura, N. Poungvarin, and L. Chan, "Glucose-dependent transcriptional regulation by an evolutionarily conserved glucose-sensing module," *Diabetes*, vol. 55, no. 5, pp. 1179–1189, 2006.

[70] M. N. Davies, B. L. O'Callaghan, and H. C. Towle, "Glucose activates chrebp by increasing its rate of nuclear entry and relieving repression of its transcriptional activity," *Journal of Biological Chemistry*, vol. 283, no. 35, pp. 24029–24038, 2008.

[71] G. Jing, J. Chen, G. Xu, and A. Shalev, "Islet chrebp-β is increased in diabetes and controls chrebp-α and glucose-induced gene expression via a negative feedback loop," *Molecular Metabolism*, 2016.

[72] K. Iizuka, R. K. Bruick, G. Liang, J. D. Horton, and K. Uyeda, "Deficiency of carbohydrate response element-binding protein (chrebp) reduces lipogenesis as well as glycolysis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 19, pp. 7281–7286, 2004.

[73] J. S. Kooner, J. C. Chambers, C. A. Aguilar-Salinas, D. A. Hinds, C. L. Hyde, G. R. Warnes, F. J. Gómez Pérez, K. A. Frazer, P. Elliott, J. Scott, P. M. Milos, D. R. Cox, and J. F. Thompson, "Genome-wide scan identifies variation in mlxipl associated with plasma triglycerides," *Nat Genet*, vol. 40, pp. 149–51, Feb 2008.

[74] K. Nakayama, Y. Yanagisawa, A. Ogawa, Y. Ishizuka, L. Munkhtulga, P. Charupoonphol, S. Supannnatas, S. Kuartei, U. Chimedregzen, Y. Koda, T. Ishida, Y. Kagawa, and S. Iwamoto, "High prevalence of an anti-hypertriglyceridemic variant of the mlxipl gene in central asia," *J Hum Genet*, vol. 56, pp. 828–33, Dec 2011.

[75] R. Kursawe, S. Caprio, C. Giannini, D. Narayan, A. Lin, E. D'Adamo, M. Shaw, B. Pierpont, S. W. Cushman, and G. I. Shulman, "Decreased transcription of chrebp-α/β isoforms in abdominal subcutaneous adipose tissue of obese adolescents with prediabetes or early type 2 diabetes," *Associations With Insulin Resistance and Hyperglycemia*, vol. 62, no. 3, pp. 837–844, 2013.

[76] O. de Luis, M. C. Valero, and L. A. Jurado, "Wbscr14, a putative transcription factor gene deleted in williams-beuren syndrome: complete characterisation of the human gene and the mouse ortholog," *Eur J Hum Genet*, vol. 8, no. 3, pp. 215–22, 2000.

[77] P.-D. Denechaud, P. Bossard, J.-M. A. Lobaccaro, L. Millatt, B. Staels, J. Girard, and C. Postic, "Chrebp, but not lxrs, is required for the induction of glucose-regulated genes in mouse liver," *The Journal of Clinical Investigation*, vol. 118, no. 3, pp. 956–964, 2008.

[78] Q. Fan, R. C. Nørgaard, C. Bindesbøll, C. Lucas, K. T. Dalen, E. Babaie, H. M. Itkonen,

J. Matthews, H. I. Nebb, and L. M. Grønning-Wang, "Lxrα regulates hepatic chrebpα activity and lipogenesis upon glucose, but not fructose feeding in mice," *Nutrients*, vol. 9, Jun 2017.

[79] A. G. Linden, S. Li, H. Y. Choi, F. Fang, M. Fukasawa, K. Uyeda, R. E. Hammer, J. D. Horton, L. J. Engelking, and G. Liang, "Interplay between chrebp and srebp-1c coordinates postprandial glycolysis and lipogenesis in livers of mice," *J Lipid Res*, vol. 59, pp. 475–487, 03 2018.

[80] F. Benhamed, P.-D. Denechaud, M. Lemoine, C. Robichon, xE, line, M. Moldes, J. Bertrand-Michel, V. Ratziu, L. Serfaty, C. Housset, J. Capeau, J. Girard, H. Guillou, xE, and C. Postic, "The lipogenic transcription factor chrebp dissociates hepatic steatosis from insulin resistance in mice and humans," *The Journal of Clinical Investigation*, vol. 122, no. 6, pp. 2176–2194, 2012.

[81] T. Jois, W. Chen, V. Howard, R. Harvey, K. Youngs, C. Thalmann, P. Saha, L. Chan, M. A. Cowley, and M. W. Sleeman, "Deletion of hepatic carbohydrate response element binding protein (chrebp) impairs glucose homeostasis and hepatic insulin sensitivity in mice," *Mol Metab*, vol. 6, no. 11, pp. 1381–1394, 2017.

[82] A. Vijayakumar, P. Aryal, J. Wen, I. Syed, R. P. Vazirani, P. M. Moraes-Vieira, J. P. Camporez, M. R. Gallop, R. J. Perry, O. D. Peroni, G. I. Shulman, A. Saghatelian, T. E. McGraw, and B. B. Kahn, "Absence of carbohydrate response element binding protein in adipocytes causes systemic insulin resistance and impairs glucose transport," *Cell Rep*, vol. 21, pp. 1021–1035, Oct 2017.

[83] A. M. Nuotio-Antar, N. Poungvarin, M. Li, M. Schupp, M. Mohammad, S. Gerard, F. Zou, and L. Chan, "Fabp4-cre mediated expression of constitutively active chrebp protects against obesity, fatty liver, and insulin resistance," *Endocrinology*, vol. 156, no. 11, pp. 4020–4032, 2015.

[84] L. S. Katz, S. Xu, K. Ge, D. K. Scott, and M. C. Gershengorn, "T3 and glucose coordinately stimulate chrebp-mediated ucp1 expression in brown adipocytes from male mice," *Endocrinology*, vol. 159, pp. 557–569, 01 2018.

[85] J. Sanchez-Gurmaches, Y. Tang, N. Z. Jespersen, M. Wallace, C. Martinez Calejman, S. Gujja, H. Li, Y. J. K. Edwards, C. Wolfrum, C. M. Metallo, S. Nielsen, C. Scheele, and D. A. Guertin, "Brown fat akt2 is a cold-induced kinase that stimulates chrebp-mediated de novo lipogenesis to optimize fuel storage and thermogenesis," *Cell Metab*, 2017.

[86] M. Kim, I. Astapova, S. N. Flier, S. A. Hannou, L. Doridot, A. Sargsyan, H. H. Kou, A. J. Fowler, G. Liang, and M. A. Herman, "Intestinal, but not hepatic, chrebp is required for fructose tolerance," *JCI Insight*, vol. 2, no. 24, 2017.

[87] D. Zhang, X. Tong, K. VanDommelen, N. Gupta, K. Stamper, G. F. Brady, Z. Meng, J. Lin, L. Rui, M. B. Omary, and L. Yin, "Lipogenic transcription factor chrebp mediates fructose-induced metabolic adaptations to prevent hepatotoxicity," *The Journal of Clinical Investigation*, vol. 127, no. 7, pp. 2855–2867, 2017.

[88] L. He, E. Chang, J. Peng, H. An, S. M. McMillin, S. Radovick, C. A. Stratakis, and F. E. Wondisford, "Activation of the camp-pka pathway antagonizes metformin suppression of hepatic glucose production," *J Biol Chem*, vol. 291, pp. 10562–70, May 2016.

[89] K. Iizuka, "The transcription factor carbohydrate-response element-binding protein (chrebp): A possible link between metabolic disease and cancer," *Biochim Biophys Acta*, vol. 1863, no. 2, pp. 474–485, 2017.

[90] M. Shaked, M. Ketzinel-Gilad, E. Cerasi, N. Kaiser, and G. Leibowitz, "Amp-activated protein kinase (ampk) mediates nutrient regulation of thioredoxin-interacting protein (txnip) in pancreatic beta-cells," *PLoS One*, vol. 6, no. 12, p. e28804, 2011.

[91] K. Iizuka, B. Miller, and K. Uyeda, "Deficiency of carbohydrate-activated transcription factor chrebp prevents obesity and improves plasma glucose control in leptin-deficient (ob/ob) mice," *American Journal of Physiology-Endocrinology and Metabolism*, vol. 291, no. 2, pp. E358–E364, 2006.

[92] W. Wu, H. Tsuchida, T. Kato, H. Niwa, Y. Horikawa, J. Takeda, and K. Iizuka, "Fat and carbohydrate in western diet contribute differently to hepatic lipid accumulation," *Biochemical and Biophysical Research Communications*, vol. 461, no. 4, pp. 681–686, 2015.

[93] J. E. Docherty, J. E. Manno, J. E. McDermott, and J. R. DiAngelo, "Mio acts in the drosophila brain to control nutrient storage and feeding," *Gene*, vol. 568, no. 2, pp. 190–5, 2015.

[94] S. A. Kliewer and D. J. Mangelsdorf, "A dozen years of discovery: Insights into the physiology and pharmacology of fgf21," *Cell Metab*, vol. 29, pp. 246–253, 02 2019.

[95] T. Kawaguchi, M. Takenoshita, T. Kabashima, and K. Uyeda, "Glucose and camp regulate the l-type pyruvate kinase gene by phosphorylation/dephosphorylation of the carbohydrate response element binding protein," *Proceedings of the National Academy of Sciences*, vol. 98, no. 24, pp. 13710–13715, 2001.

[96] T. Kabashima, T. Kawaguchi, B. E. Wadzinski, and K. Uyeda, "Xylulose 5-phosphate mediates glucose-induced lipogenesis by xylulose 5-phosphate-activated protein phosphatase in rat liver," *Proceedings of the National Academy of Sciences*, vol. 100, no. 9, pp. 5107–5112, 2003.

[97] N. G. Tsatsos and H. C. Towle, "Glucose activation of chrebp in hepatocytes occurs via a two-step mechanism," *Biochem Biophys Res Commun*, vol. 340, pp. 449–56, Feb 2006.

[98] H. Sakiyama, R. M. Wynn, W.-R. Lee, M. Fukasawa, H. Mizuguchi, K. H. Gardner, J. J. Repa, and K. Uyeda, "Regulation of nuclear import/export of carbohydrate response element-binding protein (chrebp): Interaction of an α-helix of chrebp with the 14-3-3 proteins and regulation by phosphorylation," *Journal of Biological Chemistry*, vol. 283, no. 36, pp. 24899–24908, 2008.

[99] M. R. Christie and S. J. Ashcroft, "Cyclic amp-dependent protein phosphorylation and insulin secretion in intact islets of langerhans," *Biochem J*, vol. 218, pp. 87–99, Feb 1984.

[100] N. A. Noordeen, G. Meur, G. A. Rutter, and I. Leclerc, "Glucose-induced nuclear shuttling of chrebp is mediated by sorcin and ca2+ ions in pancreatic β-cells," *Diabetes*, vol. 61, no. 3, pp. 574–585, 2012.

[101] T. Brun, E. Roche, K. H. Kim, and M. Prentki, "Glucose regulates acetyl-coa carboxylase gene expression in a pancreatic beta-cell line (ins-1)," *J Biol Chem*, vol. 268, pp. 18905–11, Sep 1993.

[102] M. V. Li, W. Chen, R. N. Harmancey, A. M. Nuotio-Antar, M. Imamura, P. Saha, H. Taegtmeyer, and L. Chan, "Glucose-6-phosphate mediates activation of the carbohydrate responsive binding protein (chrebp)," *Biochem Biophys Res Commun*, vol. 395, pp. 395–400, May 2010.

[103] S. Marie, M. J. Diaz-Guerra, L. Miquerol, A. Kahn, and P. B. Iynedjian, "The pyruvate kinase gene as a model for studies of glucose-dependent regulation of gene expression in the endocrine pancreatic beta-cell type," *J Biol Chem*, vol. 268, pp. 23881–90, Nov 1993.

[104] F. Foufelle, B. Gouhot, J. P. Pégorier, D. Perdereau, J. Girard, and P. Ferré, "Glucose stimulation of lipogenic enzyme gene expression in cultured white adipose tissue. a role for glucose 6-phosphate," *J Biol Chem*, vol. 267, pp. 20543–6, Oct 1992.

[105] R. Dentin, F. Benhamed, J.-P. Pégorier, F. Foufelle, B. Viollet, S. Vaulont, J. Girard, and C. Postic, "Polyunsaturated fatty acids suppress glycolytic and lipogenic genes through the inhibition of chrebp nuclear protein translocation," *J Clin Invest*, vol. 115, pp. 2843–54, Oct 2005.

[106] M. Panse, O. Kluth, E. Lorza-Gil, G. Kaiser, E. Mühlbauer, A. Schürmann, H.-U. Häring, S. Ullrich, and F. Gerst, "Palmitate and insulin counteract glucose-induced thioredoxin interacting protein (txnip) expression in insulin secreting cells via distinct mechanisms," *PLoS One*, vol. 13, no. 5, p. e0198016, 2018.

[107] C. M. Jenkins, J. Yang, H. F. Sims, and R. W. Gross, "Reversible high affinity inhibition of phosphofructokinase-1 by acyl-coa: a mechanism integrating glycolytic flux with lipid metabolism," *J Biol Chem*, vol. 286, pp. 11937–50, Apr 2011.

[108] N. G. Tsatsos, M. N. Davies, B. L. O'Callaghan, and H. C. Towle, "Identification and function of phosphorylation in the glucose-regulated transcription factor chrebp," *Biochemical Journal*, vol. 411, no. 2, pp. 261–270, 2008.

[109] T. Nakagawa, Q. Ge, R. Pawlosky, R. M. Wynn, R. L. Veech, and K. Uyeda, "Metabolite regulation of nucleo-cytosolic trafficking of carbohydrate response element-binding protein (chrebp): Role of ketone bodies," *Journal of Biological Chemistry*, vol. 288, no. 39, pp. 28358–28367, 2013.

[110] J. J. Collier, P. Zhang, K. B. Pedersen, S. J. Burke, J. W. Haycock, and D. K. Scott, "c-myc and chrebp regulate glucose-mediated expression of the l-type pyruvate kinase gene in ins-1-derived 832/13 cells," *Am J Physiol Endocrinol Metab*, vol. 293, no. 1, pp. E48–56, 2007.

[111] S. Sato, H. Jung, T. Nakagawa, R. Pawlosky, T. Takeshima, W.-R. Lee, H. Sakiyama, S. Laxman, R. M. Wynn, B. P. Tu, J. B. MacMillan, J. K. De Brabander, R. L. Veech, and K. Uyeda, "Metabolite regulation of nuclear localization of carbohydrate-response element-binding protein (chrebp): Role of amp as an allosteric inhibitor," *Journal of Biological Chemistry*, vol. 291, no. 20, pp. 10515–10527, 2016.

[112] M. V. Li, W. Chen, N. Poungvarin, M. Imamura, and L. Chan, "Glucose-mediated transactivation of carbohydrate response element-binding protein requires cooperative actions from mondo conserved regions and essential trans-acting factor 14-3-3," *Molecular Endocrinology*, vol. 22, no. 7, pp. 1658–1672, 2008.

[113] R. Palanivel, R. Veluthakal, and A. Kowluru, "Regulation by glucose and calcium of the carboxylmethylation of the catalytic subunit of protein phosphatase 2a in insulin-secreting ins-1 cells," *Am J Physiol Endocrinol Metab*, vol. 286, pp. E1032–41, Jun 2004.

[114] M. M. Adeva-Andany, N. Pérez-Felpete, C. Fernández-Fernández, C. Donapetry-García, and C. Pazos-García, "Liver glucose metabolism in humans," *Biosci Rep*, vol. 36, 12 2016.

[115] C. J. Hedeskov and K. Capito, "The pentose cycle and insulin release in isolated mouse pancreatic islets during starvation," *Biochem J*, vol. 152, pp. 571–6, Dec 1975.

[116] M. Huang and J. W. Joseph, "Metabolomic analysis of pancreatic β-cell insulin release in response to glucose," *Islets*, vol. 4, no. 3, pp. 210–22, 2012.

[117] O. K. Vintermyr, R. Bøe, T. Bruland, G. Houge, and S. O. Døskeland, "Elevated camp gives short-term inhibition and long-term stimulation of hepatocyte dna replication: roles of the camp-dependent protein kinase subunits," *J Cell Physiol*, vol. 156, pp. 160–70, Jul 1993.

[118] G. H. Thoresen, T. E. Sand, M. Refsnes, O. F. Dajani, T. K. Guren, I. P. Gladhaug, A. Killi, and T. Christoffersen, "Dual effects of glucagon and cyclic amp on dna synthesis in cultured rat hepatocytes: stimulatory regulation in early g1 and inhibition shortly before the s phase entry," *J Cell Physiol*, vol. 144, pp. 523–30, Sep 1990.

[119] G. Merla, C. Howald, S. E. Antonarakis, and A. Reymond, "The subcellular localization of the chore-binding protein, encoded by the williams–beuren syndrome critical region gene 14, is regulated by 14-3-3," *Human Molecular Genetics*, vol. 13, no. 14, pp. 1505–1514, 2004.

[120] Q. Ge, N. Huang, R. M. Wynn, Y. Li, X. Du, B. Miller, H. Zhang, and K. Uyeda, "Structural characterization of a unique interface between carbohydrate response element-binding protein (chrebp) and 14-3-3β protein," *Journal of Biological Chemistry*, vol. 287, no. 50, pp. 41914–41921, 2012.

[121] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, "Ucsf chimera–a visualization system for exploratory research and analysis," *J Comput Chem*, vol. 25, pp. 1605–12, Oct 2004.

[122] H. Yamashita, M. Takenoshita, M. Sakurai, R. K. Bruick, W. J. Henzel, W. Shillinglaw, D. Arnot, and K. Uyeda, "A glucose-responsive transcription factor that regulates carbohydrate metabolism in the liver," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 16, pp. 9116–9121, 2001.

[123] A. K. Stoeckman, L. Ma, and H. C. Towle, "Mlx is the functional heteromeric partner of the carbohydrate response element-binding protein in glucose regulation of lipogenic enzyme genes," *Journal of Biological Chemistry*, vol. 279, no. 15, pp. 15662–15669, 2004.

[124] N. Poungvarin, J. K. Lee, V. K. Yechoor, M. V. Li, T. Assavapokee, P. Suksaranjit, J. J. Thepsongwajja, P. K. Saha, K. Oka, and L. Chan, "Carbohydrate response element-binding protein (chrebp) plays a pivotal role in beta cell glucotoxicity," *Diabetologia*, vol. 55, no. 6, pp. 1783–1796, 2012.

[125] M. N. Davies, B. L. O'Callaghan, and H. C. Towle, "Activation and repression of glucose-stimulated chrebp requires the concerted action of multiple domains within the mondoa conserved region," *American Journal of Physiology - Endocrinology and Metabolism*, vol. 299, no. 4, pp. E665–E674, 2010.

[126] C. W. Peterson, C. A. Stoltzman, M. P. Sighinolfi, K.-S. Han, and D. E. Ayer, "Glucose controls nuclear accumulation, promoter binding, and transcriptional activity of the mondoa-mlx heterodimer," *Mol Cell Biol*, vol. 30, pp. 2887–95, Jun 2010.

[127] U. K. Jung, H., "Crystal structure of chrebp nls peptide bound to importin alpha.," Sep 2019.

[128] C. Sae-Lee, K. Moolsuwan, L. Chan, and N. Poungvarin, "Chrebp regulates itself and metabolic genes implicated in lipid accumulation in β–cell line," *PLoS ONE*, vol. 11, no. 1, p. e0147411, 2016.

[129] N. Heppner, *Interaction of Transcription Factor ChREBP with Glucose Metabolites*. PhD thesis, Ludwig-Maximilians-University Munich, Dec 2016.

[130] A. N. Billin, A. L. Eilers, K. L. Coulter, J. S. Logan, and D. E. Ayer, "Mondoa, a novel basic helix-loop-helix-leucine zipper transcriptional activator

that constitutes a positive branch of a max-like network," *Mol Cell Biol*, vol. 20, pp. 8845–54, Dec 2000.

[131] M. H. Kubala, O. Kovtun, K. Alexandrov, and B. M. Collins, "Structural and thermodynamic analysis of the gfp:gfp-nanobody complex," *Protein Sci*, vol. 19, pp. 2389–401, Dec 2010.

[132] K. Zolghadr, O. Mortusewicz, U. Rothbauer, R. Kleinhans, H. Goehler, E. E. Wanker, M. C. Cardoso, and H. Leonhardt, "A fluorescent two-hybrid assay for direct visualization of protein interactions in living cells," *Mol Cell Proteomics*, vol. 7, no. 11, pp. 2279–2287, 2008.

[133] X. Yang, W. H. Lee, F. Sobott, E. Papagrigoriou, C. V. Robinson, J. G. Grossmann, M. Sundström, D. A. Doyle, and J. M. Elkins, "Structural basis for protein-protein interactions in the 14-3-3 protein family," *Proc Natl Acad Sci U S A*, vol. 103, pp. 17237–42, Nov 2006.

[134] H. E. Hohmeier, H. Mulder, G. Chen, R. Henkel-Rieger, M. Prentki, and C. B. Newgard, "Isolation of ins-1-derived cell lines with robust atp-sensitive k+ channel-dependent and -independent glucose-stimulated insulin secretion," *Diabetes*, vol. 49, no. 3, pp. 424–430, 2000.

[135] Q. Ge, T. Nakagawa, R. M. Wynn, Y. M. Chook, B. C. Miller, and K. Uyeda, "Importin-alpha protein binding to a nuclear localization signal of carbohydrate response element-binding protein (chrebp)," *J Biol Chem*, vol. 286, pp. 28119–27, Aug 2011.

[136] N. Kudo, N. Matsumori, H. Taoka, D. Fujiwara, E. P. Schreiner, B. Wolff, M. Yoshida, and S. Horinouchi, "Leptomycin b inactivates crm1/exportin 1 by covalent modification at a cysteine residue in the central conserved region," *Proc Natl Acad Sci U S A*, vol. 96, pp. 9112–7, Aug 1999.

[137] E. M. RENKIN, "Filtration, diffusion, and molecular sieving through porous cellulose membranes," *J Gen Physiol*, vol. 38, pp. 225–43, Nov 1954.

[138] F. Cardarelli, R. Bizzarri, M. Serresi, L. Albertazzi, and F. Beltram, "Probing nuclear localization signal-importin alpha binding equilibria in living cells," *J Biol Chem*, vol. 284, pp. 36638–46, Dec 2009.

[139] A. García-González, E. Jacchetti, R. Marotta, M. Tunesi, J. F. Rodríguez Matas, and M. T. Raimondi, "The effect of cell morphology on the permeability of the nuclear envelope to diffusive factors," *Front Physiol*, vol. 9, p. 925, 2018.

[140] M. Misheva, G. Kaur, K. R. W. Ngoei, Y. Y. Yeap, I. H. W. Ng, K. M. Wagstaff, D. C. H. Ng, D. A. Jans, and M. A. Bogoyevitch, "Intracellular mobility and nuclear trafficking of the stress-activated kinase jnk1 are impeded by hyperosmotic stress," *Biochim Biophys Acta*, vol. 1843, pp. 253–64, Feb 2014.

[141] R. Gronning, *Visualization and quantitative analysis of nuclear shuttling of carbohydrate response element binding protein (ChREBP) in pancreatic beta cells*. PhD thesis, University of Southern Denmark, January 2018.

[142] M. Köster, T. Frahm, and H. Hauser, "Nucleocytoplasmic shuttling revealed by frap and flip technologies," *Curr Opin Biotechnol*, vol. 16, pp. 28–34, Feb 2005.

[143] D. Daelemans, S. V. Costes, S. Lockett, and G. N. Pavlakis, "Kinetic and molecular analysis of nuclear export factor crm1 association with its cargo in vivo," *Mol Cell Biol*, vol. 25, pp. 728–39, Jan 2005.

[144] P. Richards, L. Rachdi, M. Oshima, P. Marchetti, M. Bugliani, M. Armanet, C. Postic, S. Guilmeau, and R. Scharfmann, "Mondoa is an essential glucose-responsive transcription factor in human pancreatic β-cells," *Diabetes*, vol. 67, no. 3, pp. 461–472, 2018.

[145] U. H. Frey, H. S. Bachmann, J. Peters, and W. Siffert, "Pcr-amplification of gc-rich regions: 'slowdown pcr'," *Nat Protoc*, vol. 3, no. 8, pp. 1312–7, 2008.

[146] A. Hemsley, N. Arnheim, M. D. Toney, G. Cortopassi, and D. J. Galas, "A simple method for site-directed mutagenesis using the polymerase chain reaction," *Nucleic Acids Res*, vol. 17, pp. 6545–51, Aug 1989.

[147] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: an open-source platform for biological-image analysis," *Nat Methods*, vol. 9, pp. 676–82, Jun 2012.

[148] R. C. Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.

[149] R. Team, *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA, 2016 ed., `http://www.rstudio.com/`.

[150] W. Chang, *extrafont: Tools for using fonts*, 2014. R package version 0.17.

[151] H. Wickham, *Ggplot2: elegant graphics for data analysis*. Use R!, New York: Springer, 2009.

[152] R. M. Hope, *Rmisc: Rmisc: Ryan Miscellaneous*, 2013.

[153] A. Kassambara, *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*, 2019. R package version 0.3.1.

[154] A. A. Dragulescu and C. Arendt, *xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files*, 2018. R package version 0.6.1.

[155] S. Besson, *czi.py*, `https://gist.github.com/sbesson/42bcb21dc30f3644c17aabb6f5a1d917.js`.

[156] A. Cardona, *A Fiji Scripting Tutorial*, `https://www.ini.uzh.ch/~acardona/fiji-tutorial/`.

[157] Christian, *Flatten an irregular list of lists*, `https://stackoverflow.com/questions/2158395/flatten-an-irregular-list-of-lists`.

Chapter 7

Appendix

7.1 Data

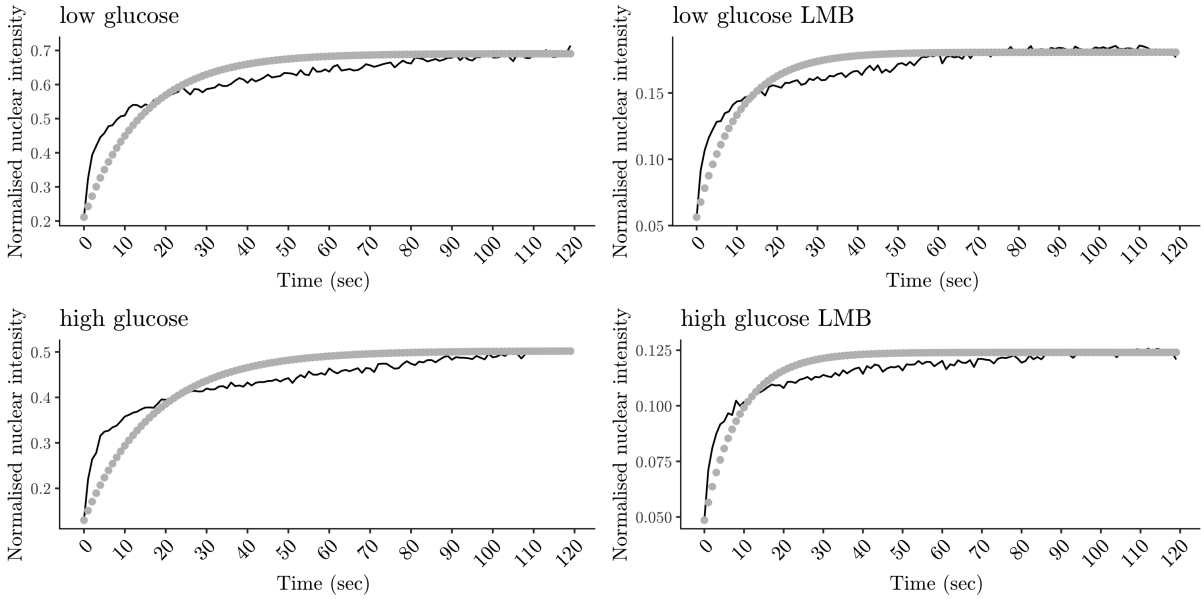


Figure 7.1: **Nuclear FRAP data do not fit a monoexponential model of nuclear import** Equation 4.2 was fit to the data in Figure 4.1b. Time of bleach was set to 0. Data is shown as a black line. The fitted monoexponential model is shown in grey.

7.2 Reagents

Table 7.1: **Reagents used**

Item	Catalogue number	Vendor
4-20% MP TGX Stain-Free Gel	4568095	BioRad
β-Mercaptoethanol	M3148	Sigma Aldrich

Table 7.1- *Continued from previous page*

Item	Catalogue number	Vendor
Acetone	34850-1L-M	Sigma Aldrich
Acrylamide/Bis Sol. 37.5:1 (30%)	10688.02	SERVA
Ammonium persulfate (APS)	09913-100G	Sigma Aldrich
Ampicillin sodium salt	A9518	Sigma Aldrich
Anti-GAPDH antibody	ab9485	Abcam
Anti-mCherry antibody	ab167453	Abcam
BamHI-HF	R3136S	New England Biolabs
Bromophenol Blue sodium salt	114405-5G	Sigma Aldrich
CellMask^{TM} Deep Red Plasma membrane Stain	C10046	Thermo Fisher Scientific
Complete EDTA free protease inhibitor	5056489001	Roche
D-Glucose-6-phosphate (sodium salt)	20376	Cayman
D-(+)-Glucose monohydrate	1.04074.9050	VWR
DMEM high glucose	D5671	Sigma Aldrich
DMEM high glucose, no phenol red	31053028	Thermo Fisher Scientific
DMEM no glucose, no phenol red	A1443001	Thermo Fisher Scientific
DpnI	R0176L	New England Biolabs
Dual-Glo® Luciferase Assay System	E2940	Promega
EcoRV-HF	R3195S	New England Biolabs
EDTA	E1644-1KG	Sigma Aldrich
fetal bovine serum	10270106	Thermo Fisher Scientific
G418 disulfate salt	A1720-5G	Sigma Aldrich
GFP (4B10) Mouse mAb	2955S	Cell Signaling
Glycerol	G6279-1L	Sigma Aldrich
Glycine	G7126-5kg	Sigma Aldrich
Goat Anti-Mouse HRP	170-6516	BioRad
Goat anti rabbit HRP	170-6515	BioRad
HEPES (1 M)	15630080	Thermo Fisher Scientific
HindIII-HF	R3104S	New England Biolabs
Hoechst 33342	H3570	Thermo Fisher Scientific
Hydrochloric acid fuming 37%	4625.2	Roth
Immobilion Western (HRP substrate)	WBKLS0500	Millipore
Kanamycin sulfate	K4000-25G	Sigma Aldrich
KCl	1049361000	Sigma Aldrich
KH_2PO_4	P9791	Sigma Aldrich
KpnI-HF	R3142S	New England Biolabs
L-Glutamine	25030024	Thermo Fisher Scientific
Leptomycin B	L2913	Sigma Aldrich
Lipofectamine® 2000 Reagent	11668019	Thermo Fisher Scientific
Methanol	20903.368P	VWR
Monoclonal Anti-alpha-Tubulin	T9026-0,2ML	Sigma Aldrich

Table 7.1- *Continued from previous page*

Item	Catalogue number	Vendor
NaCl	S7653-5KG	Sigma Aldrich
Na_2HPO_4	1.06586.1000	Roth
NHS-Activated Sepharose	17090601	GE Healthcare
Nitrocellulose, Protran®, 0.2 uM	10600001	VWR
Nonidet™ P 40 Substitute	74385-1L	Sigma Aldrich
OptiMEM	31985062	Thermo Fisher Scientific
Paraformaldehyde	158127	Sigma-Aldrich
PMSF	32395.04	SERVA
RPMI 1640	21875091	Thermo Fisher Scientific
Rabbit Anti-Goat HRP Conjugate	1721034	BioRad
RPMI 1640 no glucose	11879020	Thermo Fisher Scientific
PageRuler™ Plus Prestained Protein Ladder	26620	Thermo Fisher Scientific
Penicillin-Streptomycin Solution	15140122	Thermo Fisher Scientific
Phusion® High-Fidelity DNA Polymerase	M0530L	New England Biolabs
PpuMI	R0506S	New England Biolabs
SalI-HF	R3138L	New England Biolabs
SDS (Pellets)	20765.03	SERVA
Skim milk powder	42590.03	SERVA
Sodium deoxycholate	D6750-100g	Sigma Aldrich
Sodium hydroxide pellets	1064679010	Sigma Aldrich
Sodium Pyruvate (100 mM)	11360070	Thermo Fisher Scientific
T4 DNA Ligase	M0202L	New England Biolabs
T4 PNK	M0201L	New England Biolabs
Taq DNA Polymerase	M0267L	New England Biolabs
TEMED	T9281	Sigma Aldrich
Triton X-100	T9284-500ML	Sigma Aldrich
Trizma® base	T1503-1KG	Sigma Aldrich
Tween20	P1379-1L	Sigma Aldrich
Tryptone	95039-5kg	Sigma Aldrich
XbaI	R0145S	New England Biolabs
XhoI	RO146S	New England Biolabs
XmaI	R0180S	New England Biolabs
X-tremeGENE HP DNA Transfection Reagent	XTGHP-RO	Sigma Aldrich
Yeast extract	212730	Otto Nordwald

7.3 Sequences

The translated open reading frames of the indicated plasmids are listed below in the order in which they appear in Table 6.1. Insert species and database ID are listed in parentheses.

peGFP-C1-ChREBP, (human, CL3932)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
PTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM
DELYKSGLRSRAQASMAGALAGLAAGLQVPRVAPSPDSDSDTDSEDPSLRRSAGGLLRS
QVIHSGHFMVSSPHSDSLPRRRDQEGSVGPSDFGPRSIDPTLTRLFECLSLAYSGKLVSPK
WKNFKGLKLLCRDKIRLNNAIWRAWYIQYVKRRKSPVCGFVTPLQGPEADAHRKPEA
VVLEGNYWKRRIEVVMREYHKWRIYYKKRLRKPSREDDLLAPKQAEGRWPPPEQWC
KQLFSSVVPVLLGDPEEEPGGRQLLDLNCFLSDISDTLFTMTQSGPSPLQLPPEDAYVGN
ADMIQPDLTPLQPSLDDFMDISDFFTNSRLPQPPMPSNFPEPPSFSPVVDSLFSSGTLGPE
VPPASSAMTHLSGHSRLQARNSCPGPLDSSAFLSSDFLLPEDPKPRLPPPPVPPPLLHYPP
PAKVPGLEPCPPPPFPPMAPPTALLQEEPLFSPRFPFPTVPPAPGVSPLPAPAAFPPTPQ
SVPSPAPTPFPIELLPLGYSEPAFGPCFSMPRGKPPAPSPRGQKASPPTLAPATASPPTT
AGSNNPCLTQLLTAAKPEQALEPPLVSSTLLRSPGSPQETVPEFPCTFLPPTPAPTPPRP
PPGPATLAPSRPLLVPKAERLSPPAPSGSERRLSGDLSSMPGPGTLSVRVSPPQPILSRGR
PDSNKTENRRITHISAEQKRRFNIKLGFDTLHGLVSTLSAQPSLKVSKATTLQKTAEYIL
MLQQERAGLQEEAQQLRDEIEELNAAINLCQQQLPATGVPITHQRFDQMRDMFDDYVR
TRTLHNWKFWVFSILIRPLFESFNGMVSTASVHTLRQTSLAWLDQYCSLPALRPTVLNSL
RQLGTSTSILTDPGRIPEQATRAVTEGTLGKPL

peGFP-C1-LID (human, CL3929)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
PTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM
DELYKSGLRSRAQASSAGGLLRSQVIHSGHFMVSSPHSDSLPRRRDQEGSVGPSDFGPRSI
DPTLTRLFECLSLAYSGKLVSPKWKNFKGLKLLCRDKIRLNNAIWRAWYIQYVKRRKSP
VCGFVTPLQGPEADAHRKPEAVVLEGNYWKRRIEVVMREYHKWRIYYKKRL

peGFP-C1-GRACE (human, CL3930)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
PTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM
DELYKSGLRSRAQASREDDLLAPKQAEGRWPPPEQWCKQLFSSVVPVLLGDPEEEPGG
RQLLDLNCFLSDISDTLFTMTQSGPSPLQLPPEDAYVGNADMIQPDLTPLQPSLDDFMD

peGFP-C1-14-3-3 (human, CL3935)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
PTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM
DELYKSGLRSRAQASMTMDKSELVQKAKLAEQAERYDDMAAAMKAVTEQGHELSNEE
RNLLSVAYKNVVGARRSSWRVISSIEQKTERNEKKQQMGKEYREKIEAELQDICNDVLE
LLDKYLIPNATQPESKVFYLKMKGDYFRYLSEVASGDNKQTTVSNSQQAYQEAFEISKK
EMQPTHPIRLGLALNFSVFYYEILNSPEKACSLAKTAFDEAIAELDTLNEESYKDSTLIMQ
LLRDNLTLWTSENQGDEGDAGEGEN

pmCherry-C1-LacI-ChREBP, (human, CL3949)

MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGP
LPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDS
SLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKD
GGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMD
ELYKSGLRSMVKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAELNYI
PNRVAQQLAGKQSLLIGVATSSLALHAPSQIVAAIKSRADQLGASVVVSMVERSGVEACK
AAVHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSIIFSHEDGTRL
GVEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSAMSGFQQ
TMQMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSCYIPPLT
TIKQDFRLLGQTSVDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNTQTASPRALADSL
MQLARQVSRSSLRPPKKKRKVRSRAQASMAGALAGLAAGLQVPRVAPSPDSDSDTDSE
DPSLRRSAGGLLRSQVIHSGHFMVSSPHSDSLPRRRDQEGSVGPSDFGPRSIDPTLTRLFE
CLSLAYSGKLVSPKWKNFKGLKLLCRDKIRLNNAIWRAWYIQYVKRRKSPVCGFVTPL
QGPEADAHRKPEAVVLEGNYWKRRIEVVMREYHKWRIYYKKRLRKPSREDDLLAPKQ
AEGRWPPPEQWCKQLFSSVVPVLLGDPEEEPGGRQLLDLNCFLSDISDTLFTMTQSGPS
PLQLPPEDAYVGNADMIQPDLTPLQPSLDDFMDISDFFTNSRLPQPPMPSNFPEPPSFSP
VVDSLFSSGTLGPEVPPASSAMTHLSGHSRLQARNSCPGPLDSSAFLSSDFLLPEDPKPRL
PPPPVPPPLLHYPPPAKVPGLEPCPPPPFPPMAPPTALLQEEPLFSPRFPFPTVPPAPGV
SPLPAPAAFPPTPQSVPSPAPTPFPIELLPLGYSEPAFGPCFSMPRGKPPAPSPRGQKAS
PPTLAPATASPPTTAGSNNPCLTQLLTAAKPEQALEPPLVSSTLLRSPGSPQETVPEFPC
TFLPPTPAPTPPRPPPGPATLAPSRPLLVPKAERLSPPAPSGSERRLSGDLSSMPGPGTL
SVRVSPPQPILSRGRPDSNKTENRRITHISAEQKRRFNIKLGFDTLHGLVSTLSAQPSLKV
SKATTLQKTAEYILMLQQERAGLQEEAQQLRDEIEELNAAINLCQQQLPATGVPITHQR
FDQMRDMFDDYVRTRTLHNWKFWVFSILIRPLFESFNGMVSTASVHTLRQTSLAWLDQ
YCSLPALRPTVLNSLRQLGTSTSILTDPGRIPEQATRAVTEGTLGKPL

pmCherry-C1-LacI-LID (human, CL3950)

MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGP
LPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDS
SLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKD
GGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMD

ELYKSGLRSMVKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAELNYI
PNRVAQQLAGKQSLLIGVATSSLALHAPSQIVAAIKSRADQLGASVVVSMVERSGVEACK
AAVHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSIIFSHEDGTRL
GVEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSAMSGFQQ
TMQMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSCYIPPLT
TIKQDFRLLGQTSVDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNTQTASPRALADSL
MQLARQVSRSSLRPPKKKRKVRSRAQASSAGGLLRSQVIHSGHFMVSSPHSDSLPRRRD
QEGSVGPSDFGPRSIDPTLTRLFECLSLAYSGKLVSPKWKNFKGLKLLCRDKIRLNNAIW
RAWYIQYVKRRKSPVCGFVTPLQGPEADAHRKPEAVVLEGNYWKRRIEVVMREYHK
WRIYYKKRL

pmCherry-C1-LacI-GRACE (human, CL3951)
MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGP
LPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDS
SLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKD
GGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMD
ELYKSGLRSMVKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAELNYI
PNRVAQQLAGKQSLLIGVATSSLALHAPSQIVAAIKSRADQLGASVVVSMVERSGVEACK
AAVHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSIIFSHEDGTRL
GVEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSAMSGFQQ
TMQMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSCYIPPLT
TIKQDFRLLGQTSVDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNTQTASPRALADSL
MQLARQVSRSSLRPPKKKRKVRSRAQASREDDLLAPKQAEGRWPPPEQWCKQLFSSVV
PVLLGDPEEEPGGRQLLDLNCFLSDISDTLFTMTQSGPSPLQLPPEDAYVGNADMIQPDL
TPLQPSLDDFMD

pmCherry-C1-LacI-14-3-3 (human, CL3952)
MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGP
LPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDS
SLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKD
GGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMD
ELYKSGLRSMVKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAELNYI
PNRVAQQLAGKQSLLIGVATSSLALHAPSQIVAAIKSRADQLGASVVVSMVERSGVEACK
AAVHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSIIFSHEDGTRL
GVEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSAMSGFQQ
TMQMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSCYIPPLT
TIKQDFRLLGQTSVDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNTQTASPRALADSL
MQLARQVSRSSLRPPKKKRKVRSRAQASMTMDKSELVQKAKLAEQAERYDDMAAAMK
AVTEQGHELSNEERNLLSVAYKNVVGARRSSWRVISSIEQKTERNEKKQQMGKEYREKI
EAELQDICNDVLELLDKYLIPNATQPESKVFYLKMKGDYFRYLSEVASGDNKQTTVSNS
QQAYQEAFEISKKEMQPTHPIRLGLALNFSVFYYEILNSPEKACSLAKTAFDEAIAELDTL

NEESYKDSTLIMQLLRDNLTLWTSENQGDEGDAGEGEN

pmCherry-C1-GRACE (human, CL4279)
MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGP
LPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDS
SLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKD
GGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMD
ELYKSGLRSRAQASREDDLLAPKQAEGRWPPPEQWCKQLFSSVVPVLLGDPEEEPGGR
QLLDLNCFLSDISDTLFTMTQSGPSPLQLPPEDAYVGNADMIQPDLTPLQPSLDDFMD

pmCherry-C1-LacI-GRACE∆MCR6 (human, CL4048)
MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGP
LPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDS
SLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKD
GGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMD
ELYKSGLRSMVKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAELNYI
PNRVAQQLAGKQSLLIGVATSSLALHAPSQIVAAIKSRADQLGASVVVSMVERSGVEACK
AAVHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSIIFSHEDGTRL
GVEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSAMSGFQQ
TMQMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSCYIPPLT
TIKQDFRLLGQTSVDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNTQTASPRALADSL
MQLARQVSRSSLRPPKKKRKVRSRAQASREDDLLAPKQAEGRWPPPEQWCKQLFSSVV
PVLLGDPEEEPGGRQLLDLNCFQSGPSPLQLPPEDAYVGNADMIQPDLTPLQPSLDDFM
D

N1-GRACE-FLAG-Cherry(human, CL4193)
MAREDDLLAPKQAEGRWPPPEQWCKQLFSSVVPVLLGDPEEEPGGRQLLDLNCFLSDI
SDTLFTMTQSGPSPLQLPPEDAYVGNADMIQPDLTPLQPSLDDFMDDYKDDDDKADPP
VATMVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTK
GGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVT
QDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLK
LKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTG
GMDELYK

peGFP-C1-GSM (human, CL3931)
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
PTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM
DELYKSGLRSRAQASSAGGLLRSQVIHSGHFMVSSPHSDSLPRRRDQEGSVGPSDFGPRSI

DPTLTRLFECLSLAYSGKLVSPKWKNFKGLKLLCRDKIRLNNAIWRAWYIQYVKRRKSP
VCGFVTPLQGPEADAHRKPEAVVLEGNYWKRRIEVVMREYHKWRIYYKKRLRKPSRE
DDLLAPKQAEGRWPPPEQWCKQLFSSVVPVLLGDPEEEPGGRQLLDLNCFLSDISDTLF
TMTQSGPSPLQLPPEDAYVGNADMIQPDLTPLQPSLDDFMD

peGFP-C1-GSM-Cherry (human, CL3933)
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
PTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM
DELYKSGLRSRASAGGLLRSQVIHSGHFMVSSPHSDSLPRRRDQEGSVGPSDFGPRSIDPT
LTRLFECLSLAYSGKLVSPKWKNFKGLKLLCRDKIRLNNAIWRAWYIQYVKRRKSPVCG
FVTPLQGPEADAHRKPEAVVLEGNYWKRRIEVVMREYHKWRIYYKKRLRKPSREDDL
LAPKQAEGRWPPPEQWCKQLFSSVVPVLLGDPEEEPGGRQLLDLNCFLSDISDTLFTMT
QSGPSPLQLPPEDAYVGNADMIQPDLTPLQPSLDDFMDGTMVSKGEEDNMAIIKEFMRF
KVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAY
VKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFP
SDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKP
VQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKGSDPPDLDN

peGFP-C1-ChREBP, (mouse, CL4181)
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
PTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM
DELYKSGLRSRAQASMARALADLSVNLQVPRVVPSPDSDSDTDLEDPSPRRSAGGLHRS
QVIHSGHFMVSSPHSDSLTRRRDQEGPVGLADFGPRSIDPTLTHLFECLSLAYSGKLVSP
KWKNFKGLKLLCRDKIRLNNAIWRAWYIQYVQRRKSPVCGFVTPLQGSEADEHRKPEA
VILEGNYWKRRIEVVMREYHKWRIYYKKRLRKSSREGDFLAPKQVEGGWPPPERWCE
QLFSSVVPVLLGGSEEEPGGRQLLDLDCFLSDISDTLFTMTQPSPSSLQLPPEDAYVGNA
DMIQPDLTPLQPSLDDFMEISDFFTNYRPPQTPTSSNYIESPSFGPMADSLFSSGILAPEM
PSPASSSSSSGMTPHSGNTRLQARNSCSGPLDPNPFLSSEFLLPEDPKTKIPPAPGPTPLL
PFPTPVKVHGLEPCTPSPFPTMAPPPSLLPEESLLSARFPFTSAPPAPGVSTLPAPTTFV
PTPQPGPGPVPFSVDHLPHGYLEPVFGPHFTVPQGMQPRCKPSSPSPGGQKASPPTLAS
ATASPTATATARDNNPCLTQLLRAAKPEQALEPPTMPGTLLRPPESPQDTVSEIPRARA
FFPPIPAPTPPRPPPGPATLAPPRSLVVPKAERLSPPASSGSERRLSGDLNSIQPSGALSV
HLSPPQTVLSRGRVDNNKMENRRITHISAEQKRRFNIKLGFDTLHGLVSTLSAQPSLKVS
KATTLQKTAEYILMLQQERAAMQEEAQQLRDEIEELNAAINLCQQQLPATGVPITHQRF
DQMRDMFDDYVRTRTLHNWKFWVFSILIRPLFESFNGMVSTASLHSLRQTSLAWLEQY
CSLPALRPTVLNSLRQLSTSTSILTDPSLVPEQATRAVTEGTLGRPL

pPAGFP-C1-ChREBP, (mouse, CL4257)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
PTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFE
GDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKIRHNIEDGS
VQLADHYQQNTPIGDGPVLLPDNHYLSHQSKLSKDPNEKRDHMVLLEFVTAAGITLGM
DELYKSGLRSRAQASMARALADLSVNLQVPRVVPSPDSDSDTDLEDPSPRRSAGGLHRS
QVIHSGHFMVSSPHSDSLTRRRDQEGPVGLADFGPRSIDPTLTHLFECLSLAYSGKLVSP
KWKNFKGLKLLCRDKIRLNNAIWRAWYIQYVQRRKSPVCGFVTPLQGSEADEHRKPEA
VILEGNYWKRRIEVVMREYHKWRIYYKKRLRKSSREGDFLAPKQVEGGWPPPERWCE
QLFSSVVPVLLGGSEEEPGGRQLLDLDCFLSDISDTLFTMTQPSPSSLQLPPEDAYVGNA
DMIQPDLTPLQPSLDDFMEISDFFTNYRPPQTPTSSNYIESPSFGPMADSLFSSGILAPEM
PSPASSSSSSGMTPHSGNTRLQARNSCSGPLDPNPFLSSEFLLPEDPKTKIPPAPGPTPLL
PFPTPVKVHGLEPCTPSPFPTMAPPPSLLPEESLLSARFPFTSAPPAPGVSTLPAPTTFV
PTPQPGPGPVPFSVDHLPHGYLEPVFGPHFTVPQGMQPRCKPSSPSPGGQKASPPTLAS
ATASPTATATARDNNPCLTQLLRAAKPEQALEPPTMPGTLLRPPESPQDTVSEIPRARA
FFPPIPAPTPPRPPPGPATLAPPRSLVVPKAERLSPPASSGSERRLSGDLNSIQPSGALSV
HLSPPQTVLSRGRVDNNKMENRRITHISAEQKRRFNIKLGFDTLHGLVSTLSAQPSLKVS
KATTLQKTAEYILMLQQERAAMQEEAQQLRDEIEELNAAINLCQQQLPATGVPITHQRF
DQMRDMFDDYVRTRTLHNWKFWVFSILIRPLFESFNGMVSTASLHSLRQTSLAWLEQY
CSLPALRPTVLNSLRQLSTSTSILTDPSLVPEQATRAVTEGTLGRPL

pGampac, (-, CL4172)

MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTEVES
RLERLEQLFLLIFPREDLDMILKMDSLQDIKALLTGLFVQDNVNKDAVTDRLASVETDM
PLTLRQHRISATSSSEESSNKGQRQLTVSPEFVDEQKLISEEDLLDENLYFQGMSACLAQH
DEAVDNKFNKEQQNAFYEILHLPNLNEEQRNAFIQSLKDDPSQSANLLAEAKKLNDAQA
PKVDNKFNKEQQNAFYEILHLPNLNEEQRNAVIQSLKDDPSQSANLLAEAKKLNDAQAP
KVDANS

pGampac-LID, (mouse, CL4284)

MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTEVES
RLERLEQLFLLIFPREDLDMILKMDSLQDIKALLTGLFVQDNVNKDAVTDRLASVETDM
PLTLRQHRISATSSSEESSNKGQRQLTVSPEFARALADLSVNLQVPRVVPSPDSDSDTDLE
DPSPRRSAGGLHRSQVIHSGHFMVSSPHSDSLTRRRDQEGPVGLADFGPRSIDPTLTHLF
ECLSLAYSGKLVSPKWKNFKGLKLLCRDKIRLNNAIWRAWYIQYVQRRKSPVCGFVTPL
QGSEADEHRKPEAVILEGNYWKRRIEVVMREYHKWRIYYKKRLVDEQKLISEEDLLDE
NLYFQGMSACLAQHDEAVDNKFNKEQQNAFYEILHLPNLNEEQRNAFIQSLKDDPSQSA
NLLAEAKKLNDAQAPKVDNKFNKEQQNAFYEILHLPNLNEEQRNAVIQSLKDDPSQSAN
LLAEAKKLNDAQAPKVDANS

pGampac-GSM, (mouse, CL4185)
MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTEVES
RLERLEQLFLLIFPREDLDMILKMDSLQDIKALLTGLFVQDNVNKDAVTDRLASVETDM
PLTLRQHRISATSSSEESSNKGQRQLTVSPEFARALADLSVNLQVPRVVPSPDSDSDTDLE
DPSPRRSAGGLHRSQVIHSGHFMVSSPHSDSLTRRRDQEGPVGLADFGPRSIDPTLTHLF
ECLSLAYSGKLVSPKWKNFKGLKLLCRDKIRLNNAIWRAWYIQYVQRRKSPVCGFVTPL
QGSEADEHRKPEAVILEGNYWKRRIEVVMREYHKWRIYYKKRLRKSSREGDFLAPKQV
EGGWPPPERWCEQLFSSVVPVLLGGSEEEPGGRQLLDLDCFLSDISDTLFTMTQPSPSSL
QLPPEDAYVGNADMIQPDLTPLQPSLDDFMEVDEQKLISEEDLLDENLYFQGMSACLAQ
HDEAVDNKFNKEQQNAFYEILHLPNLNEEQRNAFIQSLKDDPSQSANLLAEAKKLNDAQ
APKVDNKFNKEQQNAFYEILHLPNLNEEQRNAVIQSLKDDPSQSANLLAEAKKLNDAQA
PKVDANS

7.4 Source code for ImageJ Jython plugins

The .czi opening portion of the below scripts was modified from the work of Sébastien Besson on GitHub [155]. The multithreading section was modified from Albert Cardona's Fiji scripting tutorial [156]. List flattening was streamlined using a function taken from a Stack Overflow thread [157].

7.4.1 F2H analysis

```
#### to call arrays and nuclei in F2H images and quantify GFP and Cherry intensity in these
### CZI opening section was modified from the work of Sebastien Besson on github
### the date parameters (i.e. biological replicates) in line 188 and the channel order in lines 67, 68, 152, and 162 are different between the positive F2H (parameter set used here) and the negative F2H.
### all other analysis parameters used are identical for both positive and negative F2H
### Multithreading section was modified from Albert Cardona's Fiji scripting tutorial available at https://www.ini.uzh.ch/~acardona/fiji-tutorial/


from ij import IJ, ImagePlus, Prefs
from ij.process import ImageStatistics as IS
from ij.plugin import ImageCalculator
from ij.measure import ResultsTable
from ij.gui import Overlay, Roi, ShapeRoi
import glob

# read in ImagePlus objects
```

```
16 from loci.plugins import BF
17 from loci.common import Region
18 from loci.plugins.in import ImporterOptions
19 from loci.plugins.util import LociPrefs
20 from loci.formats import ImageReader, MetadataTools
21 from ome.units import UNITS
22 # ZEISS STUFF
23 from loci.formats.in import ZeissCZIReader, DynamicMetadataOptions
24
25 #for multithreading
26 from java.util.concurrent import Executors, Callable, Future
27 from java.lang import Runtime
28
29 #for flattening lists
30 from collections import Iterable
31
32 def findnucleus(imp):
33   IJ.run(imp, "Gaussian Blur...", "sigma=3")
34   IJ.run(imp, "Enhance Contrast", "saturated=0.35")
35   IJ.run(imp, "Apply LUT", "")
36   IJ.run(imp, "Auto Threshold", "method=Default white")
37   IJ.run(imp, "Make Binary", "BlackBackground")
38   return imp
39
40 #function to flatten lists
41 def flatten(lis):
42      for item in lis:
43          if isinstance(item, Iterable) and not isinstance(item, basestring):
44              for x in flatten(item):
45                  yield x
46          else:
47              yield item
48
49 def resultsSaver(table, output):
50   if ResultsTable.size(table) > 0:
51     table.save(output + "F2H_Results.csv")
52     print "Image processed. Results saved."
53   else:
54     print "No positive cells found. Bye."
55
56 def measure(i, imagefile):
57   options = ImporterOptions()
58   options.setAutoscale(True)
59   options.setColorMode(ImporterOptions.COLOR_MODE_GRAYSCALE)
60   options.setId(imagefile)
61   imp = BF.openImagePlus(options)[0]
62   snapName = imp.getTitle()
63   width = imp.getWidth()
```

```
64  height = imp.getHeight()
65  print "processing", snapName
66  images = {}
67  images['DAPI_nuclei_calling'] = ImagePlus('DAPI_nuclei_calling',
        imp.getImageStack().getProcessor(1)).duplicate()
68  images['Cherry_array'] = ImagePlus('Cherry_array',
        imp.getImageStack().getProcessor(2)).duplicate()
69  """use the DAPI channel to call nuclei"""
70  findnucleus(images['DAPI_nuclei_calling'])
71  IJ.run(images['DAPI_nuclei_calling'], "Analyze Particles...", "size=4000-20000
        circularity=0.50-1.00 show=Overlay include")
72  DAPIoverlay = images['DAPI_nuclei_calling'].getOverlay()
73  if not DAPIoverlay:
74    print "no nuclei called in", snapName
75    return None
76  totalnuclei = Overlay.size(DAPIoverlay)
77  print totalnuclei, " nuclei found in", snapName
78  """use the called nuclei to select only nuclear regions in cherry channel to search
        for LacO arrays by setting the non-nuclear portions to black"""
79  nucShape = ShapeRoi(DAPIoverlay.get(0))
80  i = 1
81  while (i < totalnuclei):
82    shape = ShapeRoi(DAPIoverlay.get(i))
83    nucShape = nucShape.or(shape)
84    i = i + 1
85  bgShape = ShapeRoi(Roi(0, 0, width, height))
86  bgShape = bgShape.not(nucShape)
87  ip = images['Cherry_array'].getProcessor()
88  ip.setValue(0)
89  ip.setRoi(bgShape)
90  ip.fill(ip.getMask())
91  IJ.run(images['Cherry_array'], "8-bit", "")
92  IJ.setThreshold(images['Cherry_array'], 97, 195, "Black & White")
93  IJ.run(images['Cherry_array'], "Analyze Particles...", "size=5-200 pixel
        circularity=0.50-1.00 show=Overlay include")
94  Cherryoverlay = images['Cherry_array'].getOverlay()
95  if not Cherryoverlay:
96    print "no arrays coincident with called nuclei in", snapName
97    return None
98  totalarray = Overlay.size(Cherryoverlay)
99  print totalarray, " array(s) found in", snapName
100 completed = []
101 j = list(range(totalarray))
102 for i in range(0, totalnuclei):
103   roi = DAPIoverlay.get(i)
104   nucPoints = roi.getContainedPoints()
105   overlapCounter = 0
106   for k, element in enumerate(j):
```

```
107       roi2 = Cherryoverlay.get(element)
108       arrayPoints = roi2.getContainedPoints()
109       overlapTest = bool(set(arrayPoints) & set(nucPoints))
110       if overlapTest == True:
111         completed += [element]
112         overlapCounter = overlapCounter + 1
113     j = set(j) ^ set(completed)
114     if overlapCounter != 1:
115       Roi.setName(roi, "delete")
116   finalOverlay = Overlay()
117   completed = []
118   j = list(range(totalarray))
119   for i in range(0, totalnuclei):
120     tempNo = str(i + 1)
121     roi = DAPIoverlay.get(i)
122     Rname = Roi.getName(roi)
123     if Rname != "delete":
124       nucPoints = roi.getContainedPoints()
125       breakVal = 0
126       for k, element in enumerate(j):
127         if breakVal == 'completed': break
128         roi2 = Cherryoverlay.get(element)
129         arrayPoints = roi2.getContainedPoints()
130         overlapTest = bool(set(arrayPoints) & set(nucPoints))
131         if overlapTest == True:
132           nucleoplasm = ShapeRoi(roi).not(ShapeRoi(roi2))
133           Roi.setName(roi, "nucleus_" + tempNo)
134           Roi.setName(roi2, "array_" + tempNo)
135           Roi.setName(nucleoplasm, "nucleoplasm_" + tempNo)
136           finalOverlay.add(roi)
137           finalOverlay.add(roi2)
138           finalOverlay.add(nucleoplasm)
139           completed += [element]
140           breakVal = 'completed'
141       j = set(j) ^ set(completed)
142   if Overlay.size(finalOverlay) == 0:
143     print "no single coincident arrays and nuclei identified in snap",snapName
144     return None
145   table = []
146   finalCount = Overlay.size(finalOverlay)
147   imp.setOverlay(finalOverlay)
148   for i in range(0, finalCount):
149     roi = finalOverlay.get(i)
150     roi.setImage(imp)
151     Rname = Roi.getName(roi)
152     imp.setSlice(2)
153     roi.setPosition(imp)
154     roiStat = roi.getStatistics()
```

```
155     table += [imagefile]
156     table += [snapName]
157     table += ["Cherry"]
158     table += [Rname]
159     table += [roiStat.area]
160     table += [roiStat.mean]
161     table += [roiStat.median]
162     imp.setSlice(3)
163     roi.setPosition(imp)
164     roiStat = roi.getStatistics()
165     table += [imagefile]
166     table += [snapName]
167     table += ["GFP"]
168     table += [Rname]
169     table += [roiStat.area]
170     table += [roiStat.mean]
171     table += [roiStat.median]
172   if len(table) > 0:
173     return table
174   else:
175     print "table not generated for",snapName, "even though transfected cells
          overlapped specific nuclei."
176     return None
177
178 # A wrapper for concurrent execution
179 class Task(Callable):
180   def __init__(self, fn, *args):
181     self.fn = fn
182     self.args = args
183   def call(self):
184     return self.fn(*self.args)
185
186 inputDir = "/work/project/ladlad_001/F2H/"
187 outputDir = "/work/project/ladlad_001/F2H/output/"
188 dates = ["181121", "181130", "181219"]
189 pathList = []
190 for i, date in enumerate(dates):
191   pathList += [glob.glob(inputDir + date + "/Snap*.czi")]
192
193 pathList = list(flatten(pathList))
194 pathLen = len(pathList)
195
196 n_threads = Runtime.getRuntime().availableProcessors()
197 exe = Executors.newFixedThreadPool(n_threads)
198
199 try:
200   futures = [exe.submit(Task(measure,i, imagefile))
201     for i, imagefile in enumerate(pathList)]
```

```
202   steps = [f.get() for f in futures]
203 finally:
204   exe.shutdown()
205
206 print "image processing finished"
207
208 i = 0
209 while i < len(steps):
210   if steps[i] == None:
211     del steps[i]
212   else:
213     i = i + 1
214
215 posLen = len(steps)
216
217 print posLen, "positive images found"
218
219 print "about to flatten steps"
220
221 steps = list(flatten(steps))
222
223 pathList = []
224 channelList = []
225 nameList = []
226 roiList = []
227 areaList = []
228 meanList = []
229 medianList = []
230
231 for i in range(0, len(steps), 7):
232   pathList += [steps[i]]
233   nameList += [steps[i+1]]
234   channelList += [steps[i+2]]
235   roiList += [steps[i+3]]
236   areaList += [steps[i+4]]
237   meanList += [steps[i+5]]
238   medianList += [steps[i+6]]
239
240 table = ResultsTable()
241 resultLen = len(nameList)
242 print resultLen, "rows of results generated"
243 print "filling results table"
244 for i in range(0, resultLen):
245   table.incrementCounter()
246   table.addValue("Path", pathList[i])
247   table.addValue("Name", nameList[i])
248   table.addValue("Channel", channelList[i])
249   table.addValue("ROI", roiList[i])
```

```
250   table.addValue("Area", areaList[i])
251   table.addValue("Mean", meanList[i])
252   table.addValue("Median", medianList[i])
253 resultsSaver(table, outputDir)
```

7.4.2 Subcellular localisation analysis

```
1  ### to analyse GFP subcellular localisation CZI files
2  ### CZI opening section was modified from the work of Sebastien Besson on github
3  ### Multithreading section was modified from Albert Cardona's Fiji scripting tutorial
       available at https://www.ini.uzh.ch/~acardona/fiji-tutorial/
4
5  #@ String (label="Select the input file") imagefile
6  #@ String (label="Select the output directory") output
7
8  from ij import IJ, ImagePlus, Prefs
9  from ij.process import ImageStatistics as IS
10 from ij.plugin import ImageCalculator, filter
11 from ij.measure import ResultsTable
12 from ij.gui import Overlay, Roi, ShapeRoi
13 from java.io import File
14 import os
15 from os import path
16 # read in ImagePlus objects
17 from loci.plugins import BF
18 from loci.common import Region
19 from loci.plugins.in import ImporterOptions
20 from loci.plugins.util import LociPrefs
21 from loci.formats import ImageReader, MetadataTools
22 from ome.units import UNITS
23 #to read CZI files
24 from loci.formats.in import ZeissCZIReader, DynamicMetadataOptions
25 #for multithreading
26 from java.util.concurrent import Executors, Callable, Future
27 from java.lang import Runtime
28 #for flattening lists
29 from collections import Iterable
30 #regular expressions
31 import re
32
33 def binarize(imp, channel):
34   if channel == "DAPI":
35     IJ.run(imp, "Enhance Contrast", "saturated=0.35")
36     IJ.run(imp, "Apply LUT", "")
37     IJ.run(imp, "Gaussian Blur...", "sigma=1")
38   if channel == "DAPI-thresh":
```

```
39      ip = imp.getProcessor()
40      stats = IS.getStatistics(ip, IS.MIN_MAX, imp.getCalibration())
41      maxVal = stats.max
42      imp.setDisplayRange(450, maxVal)
43      IJ.run(imp, "Apply LUT", "")
44      IJ.run(imp, "Gaussian Blur...", "sigma=1")
45    if channel == "GFP":
46      ip = imp.getProcessor()
47      stats = IS.getStatistics(ip, IS.MIN_MAX, imp.getCalibration())
48      maxVal = stats.max
49      imp.setDisplayRange(450, maxVal)
50      IJ.run(imp, "Apply LUT", "")
51    if channel == "cell-mask-thresh":
52      ip = imp.getProcessor()
53      stats = IS.getStatistics(ip, IS.MIN_MAX, imp.getCalibration())
54      maxVal = stats.max
55      imp.setDisplayRange(500, maxVal)
56      IJ.run(imp, "Apply LUT", "")
57    if channel == "cell-mask":
58      IJ.run(imp, "Bandpass Filter...", "filter_large=40 filter_small=1 suppress=None
            tolerance=0 autoscale saturate")
59      IJ.run(imp, "8-bit", "")
60      IJ.run(imp, "Auto Threshold", "method=Default white")
61    IJ.run(imp, "Make Binary", "BlackBackground")
62    return imp
63
64  def resultsSaver(table, name, output):
65    if ResultsTable.size(table) > 0:
66      table.save(output + name + "_Results-threaded.csv")
67      print "Image processed. Results saved."
68    else:
69      print "No transfected cells found. Bye."
70
71
72  def getCZIinfo(imagefile, setflatres=False, setconcat=False):
73    CZIinfo = {}
74    options = DynamicMetadataOptions()
75    options.setBoolean("zeissczi.autostitch", False)
76    options.setBoolean("zeissczi.attachments", False)
77    czireader = ZeissCZIReader()
78    czireader.setFlattenedResolutions(setflatres)
79    czireader.setMetadataOptions(options)
80    czireader.setId(imagefile)
81    CZIinfo['rescount'] = czireader.getResolutionCount()
82    CZIinfo['seriesCount'] = czireader.getSeriesCount()
83    CZIinfo['flatres'] = czireader.hasFlattenedResolutions()
84    CZIinfo['getreslevel'] = czireader.getResolution()
85    CZIinfo['SizeT'] = czireader.getSizeT()
```

```
86    CZIinfo['SizeZ'] = czireader.getSizeZ()
87    CZIinfo['SizeC'] = czireader.getSizeC()
88    CZIinfo['SizeX'] = czireader.getSizeX()
89    CZIinfo['SizeY'] = czireader.getSizeY()
90    czireader.close()
91    print "CZIinfo generated"
92    return CZIinfo
93
94  def imageOpener(imagefile):
95    # Set the preferences in the ImageJ plugin
96    Prefs.set("bioformats.zeissczi.allow.autostitch", str(False).lower())
97    Prefs.set("bioformats.zeissczi.include.attachments", str(True).lower())
98    # read in and display ImagePlus(es) with arguments
99    options = ImporterOptions()
100   options.setOpenAllSeries(True)
101   options.setConcatenate(False)
102   options.setId(imagefile)
103   imps = BF.openImagePlus(options)
104   return imps
105
106 def imageProcessor(imps, width, height, frame):
107   imp = imps[frame]
108   snapName = imp.getTitle()
109   snapNo = snapName.split("#")[1]
110   print "beginning processing of frame", snapNo
111   images = {}
112   images['DAPI'] = ImagePlus("DAPI", imp.getImageStack().getProcessor(1)).duplicate()
113   images['GFP'] = ImagePlus("GFP", imp.getImageStack().getProcessor(2)).duplicate()
114   images['cell-mask'] = ImagePlus("cell-mask",
          imp.getImageStack().getProcessor(3)).duplicate()
115   stats = IS.getStatistics(images['cell-mask'].getProcessor(), IS.MIN_MAX,
          images['cell-mask'].getCalibration())
116   if stats.max < 500:
117     print "cell mask value below threshold. No cells in", snapNo
118     return None
119   binarize(images['DAPI'], "DAPI")
120   stats = IS.getStatistics(images['DAPI'].getProcessor(), IS.MEAN,
          images['DAPI'].getCalibration())
121   if stats.mean > 150:
122     print "Auto contrast didn't work for DAPI. Repeating with thresholding in", snapNo
123     images['DAPI'] = ImagePlus("DAPI", imp.getImageStack().getProcessor(1)).duplicate()
124     binarize(images['DAPI'], "DAPI-thresh")
125   binarize(images['GFP'], "GFP")
126   stats = IS.getStatistics(images['GFP'].getProcessor(), IS.MIN_MAX,
          images['GFP'].getCalibration())
127   if stats.min == stats.max:
128     print "binarized GFPmin = GFPmax. No cells in", snapNo
129     return None
```

```
130   binarize(images['cell-mask'], "cell-mask")
131   stats = IS.getStatistics(images['cell-mask'].getProcessor(), IS.MEAN,
          images['cell-mask'].getCalibration())
132   if stats.mean > 250:
133     print "FFT filtering didn't work for cell mask. Repeating with thresholding in",
            snapNo
134     images['cell-mask'] = ImagePlus("cell-mask",
            imp.getImageStack().getProcessor(3)).duplicate()
135     binarize(images['cell-mask'], "cell-mask-thresh")
136   for k in images:
137     images[k].setTitle(k)
138     """create a new image which contains only signal containing regions of the cell
            mask and GFP channels"""
139   images['GFP_cleaned'] = ImageCalculator().run("AND create", images['cell-mask'],
          images['GFP'])
140   images['GFP_cleaned'].setTitle('GFP_cleaned')
141   IJ.run(images['GFP'] , "Analyze Particles...", "size=300-2500 show=Overlay include")
142   GFPoverlay = images['GFP'].getOverlay()
143   if not GFPoverlay:
144     print "no transfected cells identified in snap", snapNo
145     return None
146   images['GFP_cleaned'].setOverlay(GFPoverlay)
147   cellCount = Overlay.size(GFPoverlay)
148   IJ.run(images['GFP'] , "Analyze Particles...", "size=300-2500 show=Overlay include")
149   GFPoverlay = images['GFP'].getOverlay()
150   if not GFPoverlay:
151     print "no transfected cells identified in snap", snapNo
152     return None
153   cellCount = Overlay.size(GFPoverlay)
154   i = 0
155   while i < cellCount:
156     roi = GFPoverlay.get(i)
157     images['GFP_cleaned'].setRoi(roi)
158     statMean = roi.getStatistics().mean
159     if statMean == 0:
160       GFPoverlay.remove(roi)
161 #     print "roi deleted"
162       cellCount = cellCount - 1
163       if cellCount == 0:
164         print "no cells found in", snapNo, "after cell mask filtering"
165         return None
166     else:
167       i = i + 1
168   """create an overlay of _all_ nuclei"""
169   IJ.run(images['DAPI'], "Analyze Particles...", "size=150-500 circularity=0.5-1.00
          show=Overlay include")
170   DAPIoverlay = images['DAPI'].getOverlay()
171   if not DAPIoverlay:
```

```
172     print "no nuclei found in snap", snapNo
173     return None
174   nucleiCount = Overlay.size(DAPIoverlay)
175   if nucleiCount == 1:
176     print "only 1 nucleus initially called in", snapNo, "likely background, skipping
          to next image."
177     return None
178 # print "initial nuclei:", nucleiCount
179   """iterate across all nuclei and delete those with no overlap to GFP cells"""
180   i = 0
181   while DAPIoverlay.contains(DAPIoverlay.get(i)):
182     roi = DAPIoverlay.get(i)
183     images['GFP_cleaned'].setRoi(roi)
184     if roi.getStatistics().mean < 20:
185       DAPIoverlay.remove(roi)
186       nucleiCount = nucleiCount - 1
187       if nucleiCount == 0:
188         break
189     else:
190       i = i + 1
191   if nucleiCount == 0:
192     print "no nuclei found to be coincident with called transfected cells in snap",
          snapNo
193     return None
194 # print "totalnuclei:", nucleiCount
195   i = 0
196   while i < cellCount:
197     overlapIndex = []
198     overlapTestList = []
199     roi = GFPoverlay.get(i)
200     GFPPoints = roi.getContainedPoints()
201     j = 0
202     while j < nucleiCount:
203       roi2 = DAPIoverlay.get(j)
204       nucPoints = roi2.getContainedPoints()
205       overlapTest = bool(set(GFPPoints) & set(nucPoints))
206       if overlapTest == True:
207         overlapIndex += [j]
208         overlapTestList += [len(set(GFPPoints) & set(nucPoints))]
209       if j == nucleiCount - 1 and len(overlapIndex) == 1:
210         roi4 = DAPIoverlay.get(overlapIndex[0])
211         roi4.setName("uniquefoo" + str(i)+ str(j))
212         roi.setName("uniquefoo" + str(i)+ str(j))
213       if j == nucleiCount - 1 and len(overlapIndex) > 1:
214         maxInd = overlapTestList.index(max(overlapTestList))
215         tempMax = overlapIndex[maxInd]
216         roi4 = DAPIoverlay.get(tempMax)
217         if roi4.getName() is None:
```

```
218         roi4.setName("foo" + str(i)+ str(j))
219         roi.setName("foo" + str(i)+ str(j))
220       else:
221         if "unique" in roi4.getName():
222           overlapIndex.remove(tempMax)
223           overlapTestList.remove(max(overlapTestList))
224           maxInd = overlapTestList.index(max(overlapTestList))
225           tempMax = overlapIndex[maxInd]
226           roi4 = DAPIoverlay.get(tempMax)
227           roi4.setName("foo" + str(i)+ str(j))
228           roi.setName("foo" + str(i)+ str(j))
229         else:
230           roi4.setName("foo" + str(i)+ str(j))
231           roi.setName("foo" + str(i)+ str(j))
232     j = j + 1
233   if len(overlapIndex) == 0:
234     GFPoverlay.remove(roi)
235     cellCount = cellCount - 1
236 #    print "cell deleted", i
237   else:
238     if i == cellCount - 1:
239       for k in range(nucleiCount):
240         roi3 = DAPIoverlay.get(i)
241         if DAPIoverlay.contains(roi3) == True:
242           if roi3.getName() is None:
243             DAPIoverlay.remove(roi3)
244             nucleiCount = nucleiCount - 1
245             if nucleiCount == 0:
246               print "no single transfected cells overlapped a single transfected
                      nucleus in snap", snapNo
247               return None
248         else:
249           nucleiCount = k
250           break
251     i = i + 1
252 j = 0
253 while j < nucleiCount:
254   overlapIndex = []
255   overlapTestList = []
256   roi2 = DAPIoverlay.get(j)
257   nucPoints = roi2.getContainedPoints()
258   i = 0
259   while i < cellCount:
260     roi = GFPoverlay.get(i)
261     GFPPoints = roi.getContainedPoints()
262     overlapTest = bool(set(GFPPoints) & set(nucPoints))
263     if overlapTest == True:
264       overlapIndex += [i]
```

```
265         overlapTestList += [len(set(GFPPoints) & set(nucPoints))]
266       if i == cellCount - 1 and len(overlapIndex) == 1:
267         roi4 = GFPoverlay.get(overlapIndex[0])
268         roi4.setName("uniquebar" + str(i)+ str(j))
269         roi2.setName("uniquebar" + str(i)+ str(j))
270       if i == cellCount - 1 and len(overlapIndex) > 1:
271         maxInd = overlapTestList.index(max(overlapTestList))
272         tempMax = overlapIndex[maxInd]
273         roi4 = GFPoverlay.get(tempMax)
274         if roi4.getName() is None:
275           roi4.setName("bar" + str(i)+ str(j))
276           roi2.setName("bar" + str(i)+ str(j))
277         else:
278           if "unique" in roi4.getName():
279             overlapIndex.remove(tempMax)
280             overlapTestList.remove(max(overlapTestList))
281             maxInd = overlapTestList.index(max(overlapTestList))
282             tempMax = overlapIndex[maxInd]
283             roi4 = GFPoverlay.get(tempMax)
284             roi4.setName("bar" + str(i)+ str(j))
285             roi2.setName("bar" + str(i)+ str(j))
286           else:
287             roi4.setName("bar" + str(i)+ str(j))
288             roi2.setName("bar" + str(i)+ str(j))
289       i = i + 1
290     if len(overlapIndex) == 0:
291       DAPIoverlay.remove(roi2)
292       nucleiCount = nucleiCount - 1
293 #     print "nuc deleted", i
294     else:
295       if j == nucleiCount - 1:
296         for k in range(cellCount):
297           roi3 = GFPoverlay.get(k)
298           if GFPoverlay.contains(roi3) == True:
299             if roi3.getName() is None:
300               GFPoverlay.remove(roi3)
301               cellCount = cellCount - 1
302               if cellCount == 0:
303                 print "no single transfected cells overlapped a single transfected
                        nucleus in snap", snapNo
304                 return None
305           else:
306             cellCount = k
307             break
308       j = j + 1
309   for i in range(cellCount):
310     roi = GFPoverlay.get(i)
311     Rname = roi.getName()
```

```
312     if DAPIoverlay.contains(DAPIoverlay.get(Rname)) == False:
313       roi.setName("unmatched" + str(i))
314   for i in range(nucleiCount):
315     roi = DAPIoverlay.get(i)
316     Rname = roi.getName()
317     if GFPoverlay.contains(GFPoverlay.get(Rname)) == False:
318       roi.setName("unmatched" + str(i))
319   for i in range(cellCount):
320     roi = GFPoverlay.get(i)
321     Rname = roi.getName()
322     if "unmatched" in Rname:
323       for j in range(nucleiCount):
324         roi2 = DAPIoverlay.get(i)
325         if DAPIoverlay.contains(roi2) == True:
326           Rname2 = roi2.getName()
327           if "unmatched" in Rname2:
328             GFPPoints = roi.getContainedPoints()
329             nucPoints = roi2.getContainedPoints()
330             overlapTest = bool(set(GFPPoints) & set(nucPoints))
331             if overlapTest == True:
332               roi.setName("name" + str(i))
333               roi2.setName("name" + str(i))
334   """even though the microscope is confocal, get more than one functional (with
      respect to the cell) z slice i.e. fluorescence is detected which corresponds to
      the maximal apparent (in the x, y plane) size of the subcellular compartment in
      which the stain is localised which can be captured in the given z plane of
      imaging. This means that the nuclear border in the DAPI channel shows the
      largest possible nucleus within the z plane of a given snap while, for
      cytoplasmically localised GFP in the same cell, the GFP channel shows the
      largest possible cytosol (i.e. smallest possible nucleus) in the z-plane of a
      given snap. The resulting differences in functional nuclear size in the GFP
      channel generate artefacts in the detected intensity of GFP in a given
      subcellular compartment. So need to adjust for this to accurately quantify
      subcellular localisation in the case of cytosolic localised GFP signal"""
335   i = 0
336   while i < cellCount:
337     roi = GFPoverlay.get(i)
338     Rname = roi.getName()
339     GFPPoints = roi.getContainedPoints()
340     roi2 = DAPIoverlay.get(Rname)
341     if DAPIoverlay.contains(roi2) == True:
342       nucPoints = roi2.getContainedPoints()
343       overlapTest = bool(set(GFPPoints) & set(nucPoints))
344       if overlapTest == True:
345         if len(set(GFPPoints) & set(nucPoints)) > 0.9 * len(set(nucPoints)):
346           print "GFP is not entirely cytoplasmic in cell", Rname, "snap", snapNo
347           roi.setName("nucGFP_" + Rname)
348           roi2.setName("nucleus_" + Rname)
```

```
      else:
        roi.setName("cell_" + Rname)
        nucleus = ShapeRoi(roi2).not(ShapeRoi(roi))
        DAPIoverlay.add(nucleus, "nucleus_" + Rname)
    i = i + 1
  else:
    GFPoverlay.remove(roi)
    cellCount = cellCount - 1
    print "DAPIoverlay.contains(roi2) == False in line 350 snap", snapNo
images['cell-mask_inv'] = images['cell-mask'].duplicate()
images['cell-mask_inv'].setTitle('cell-mask_inv')
IJ.run(images['cell-mask_inv'], "Invert", "")
IJ.run(images['cell-mask_inv'], "Watershed", "")
IJ.run(images['cell-mask_inv'] , "Analyze Particles...", "size=50-140000
    show=Overlay")
IJ.run(images['cell-mask_inv'], "Watershed", "")
bgOverlay = images['cell-mask_inv'].getOverlay()
if not bgOverlay:
  print "no bg identified in snap", snapNo
  return None
bgSize = Overlay.size(bgOverlay)
if bgSize == 1:
  bgRoi = bgOverlay.get(0)
  Roi.setName(bgRoi, "bg")
else:
  bgShape = ShapeRoi(bgOverlay.get(0))
  for i in range(1, bgSize):
    shape = ShapeRoi(bgOverlay.get(i))
    bgShape = bgShape.or(shape)
  bgRoi = bgShape.shapeToRoi()
  Roi.setName(bgRoi, "bg")
  bgOverlay.clear()
  bgOverlay.add(bgRoi)
  bgShape = ShapeRoi(bgOverlay.get(0))
print "bg shape generated in snap", snapNo
IJ.run(images['cell-mask'], "Watershed", "")
IJ.run(images['cell-mask'] , "Analyze Particles...", "size=50-5000 show=Overlay
    include")
maskoverlay = images['cell-mask'].getOverlay()
print "cell mask overlay generated in snap", snapNo
Celloverlay = Overlay()
for i in range(0, cellCount):
  counter = str(i)
  roi = GFPoverlay.get(i)
  Rname = Roi.getName(roi)
  tempNo = re.split("_", Rname)[-1]
  if "cell_" in Rname:
    roi2 = DAPIoverlay.get("nucleus_" + tempNo)
```

```
395       cell = ShapeRoi(roi).or(ShapeRoi(roi2))
396       cytoplasm = ShapeRoi(roi).not(ShapeRoi(roi2))
397       cellName = "cell_" + counter
398       Roi.setName(cell, cellName)
399       Roi.setName(roi2, "nucleus_" + counter)
400       Roi.setName(cytoplasm, "cytoplasm_" + counter)
401       Celloverlay.add(cell)
402       Celloverlay.add(roi2)
403       Celloverlay.add(cytoplasm)
404     if "nucGFP_" in Rname:
405       roi2 = DAPIoverlay.get("nucleus_" + tempNo)
406       cytoplasm = ShapeRoi(roi).not(ShapeRoi(roi2))
407       cytArea = cytoplasm.getStatistics().area
408       if cytArea > 0:
409         cell = ShapeRoi(roi).or(ShapeRoi(roi2))
410         Roi.setName(cell, "cell_" + counter)
411         Roi.setName(roi2, "nucleus_" + counter)
412         Roi.setName(cytoplasm, "cytoplasm_" + counter)
413         Celloverlay.add(cell)
414         Celloverlay.add(roi2)
415         Celloverlay.add(cytoplasm)
416       else:
417         for k in range(0,Overlay.size(maskoverlay)):
418           roi3 = maskoverlay.get(k)
419           cellPoints = roi3.getContainedPoints()
420           nucPoints = roi2.getContainedPoints()
421           overlapTest = bool(set(cellPoints) & set(nucPoints))
422           if overlapTest == False and k == Overlay.size(maskoverlay) - 1:
423             print "no cellmask cell identified for", Rname, "in snap", snapNo
424           if overlapTest == True:
425             cellShape = ShapeRoi(roi3)
426             nucleusShape = ShapeRoi(roi2)
427             cell = ShapeRoi(roi).or(ShapeRoi(roi2))
428             cytoplasm = ShapeRoi(roi).not(ShapeRoi(roi2))
429             Roi.setName(cell, "cell_" + counter)
430             Roi.setName(roi2, "nucleus_" + counter)
431             Roi.setName(cytoplasm, "cytoplasm_" + counter)
432             Celloverlay.add(cell)
433             Celloverlay.add(roi2)
434             Celloverlay.add(cytoplasm)
435             k = Overlay.size(maskoverlay) - 1
436   """add the background ROI to the final overlay"""
437   if Overlay.size(Celloverlay) == 0:
438     print "no transfected cells identified in snap", snapNo, "after mask filtering"
439     return None
440   if not isinstance(Overlay.size(Celloverlay)/3, (int, long)):
441     print "incorrect cell:nucleus:cytoplasm ratio in", snapNo, "after mask filtering"
442     return None
```

```
443   table = []
444   Overlay.add(Celloverlay, bgOverlay.get(0))
445   finalCell = Overlay.size(Celloverlay)
446   imp.setSlice(2)
447   imp.setOverlay(Celloverlay)
448   for i in range(0, finalCell):
449     roi = Celloverlay.get(i)
450     roi.setImage(imp)
451     roi.setPosition(1, 2, 1)
452     Rname = Roi.getName(roi)
453     roiStat = roi.getStatistics()
454     table += [name]
455     table += [snapNo]
456     table += [Rname]
457     table += [roiStat.area]
458     table += [roiStat.mean]
459     table += [roiStat.median]
460   if len(table) > 0:
461     return table
462   else:
463     print "table not generated for", snapNo, "even though transfected cells overlapped
          specific nuclei."
464     return None
465
466 #function to flatten lists
467 def flatten(lis):
468      for item in lis:
469          if isinstance(item, Iterable) and not isinstance(item, basestring):
470              for x in flatten(item):
471                  yield x
472          else:
473              yield item
474
475 # A wrapper for concurrent execution
476 class Task(Callable):
477   def __init__(self, fn, *args):
478     self.fn = fn
479     self.args = args
480   def call(self):
481     return self.fn(*self.args)
482
483 splitFile = imagefile.split("/")
484 fileLen = len(splitFile)
485 name = splitFile[fileLen-1].split(".")[0]
486
487 print "name is: ", name
488
489 CZIinfo = getCZIinfo(imagefile, setflatres=False, setconcat=False)
```

```
490
491 imps = imageOpener(imagefile)
492
493 n_threads = Runtime.getRuntime().availableProcessors()
494 exe = Executors.newFixedThreadPool(n_threads)
495
496 try:
497   futures = [exe.submit(Task(imageProcessor,
498     imps, CZIinfo['SizeX'], CZIinfo['SizeY'], frame))
499     for frame in xrange(0, CZIinfo['seriesCount'])]
500   steps = [f.get() for f in futures]
501 finally:
502   exe.shutdown()
503
504 print "image processing finished"
505
506 i = 0
507 while i < len(steps):
508   if steps[i] == None:
509     del steps[i]
510   else:
511     i = i + 1
512
513 posLen = len(steps)
514
515 print posLen, "positive images found"
516
517 print "about to flatten steps"
518
519 steps = list(flatten(steps))
520
521 #print "steps", steps
522
523 nameList = []
524 snapList = []
525 roiList = []
526 areaList = []
527 meanList = []
528 medianList = []
529
530 for i in range(0, len(steps), 6):
531     nameList += [steps[i]]
532     snapList += [steps[i+1]]
533     roiList += [steps[i+2]]
534     areaList += [steps[i+3]]
535     meanList += [steps[i+4]]
536     medianList += [steps[i+5]]
537
```

```
538 table = ResultsTable()
539
540 resultLen = len(nameList)
541 print resultLen, "rows of results generated"
542 print "filling results table"
543 for i in range(0, resultLen):
544   table.incrementCounter()
545   table.addValue("Name", nameList[i])
546   table.addValue("SnapNo", snapList[i])
547   table.addValue("ROI", roiList[i])
548   table.addValue("Area", areaList[i])
549   table.addValue("Mean", meanList[i])
550   table.addValue("Median", medianList[i])
551
552 resultsSaver(table, name, output)
```