AMPK-regulated miRNAs in pancreatic β-cells: underlying molecular mechanisms and high throughput target identification

Grazia Pizza

Imperial College London

Department of Metabolism, Digestion and Reproduction
Section of Cell Biology and Functional Genomics

Thesis presented for the degree of Doctor of
Philosophy and diploma of Imperial College London

Declaration of Authenticity and Copyright

II declare that the work presented in this thesis is original and generated by myself, except where indicated by special reference in the text, and that no part of the dissertation has been submitted for any other degree.

To date, some of the results contained here have been presented at conferences and seminars.

Abstract

MicroRNAs (miRNAs) are small noncoding RNAs that silence gene expression post-transcriptionally and play a key role in the development of Type 2 Diabetes (T2D). In pancreatic β-cells, miRNAs control glucose-induced insulin secretion and content and their expression is altered in T2D and hyperglycaemia through unknown molecular mechanisms. MiR-184, an important regulator of insulin secretion, and miR-125b, whose role in β-cells is still unknown, are both regulated by glucose through AMP-activated protein kinase (AMPK), a nutrient sensor essential for cellular energy homeostasis and suggested target for anti-diabetes drugs.

With this thesis we aimed to identify the mechanisms linked to AMPK-mediated transcription regulation of miR-184 and miR-125b in β-cells. We also aimed to elucidate miR-125b function in β-cells through high-throughput identification of its direct targets. Finally, we aimed to develop a high-throughput method (AGO2 eiCLIP-seq) to identify all specific miRNA-mRNA interactions in human β-cells.

Our results suggest that AMPK mediates glucose-dependent miR-184 expression regulation independently of CTCF and that miR-184 TSS is located ~121 Kb upstream of *MIR184* in a less accessible region in LKB1KO islets, suggesting a positive regulatory role in miR-184 transcription for AMPK. We also found that miR-125b expression is negatively regulated by SMAD3 and that the AMPK-mediated down-regulation of miR-125b might occur independently of SMAD3. Importantly, we identified miR-125b direct targets in β-cells and found that miR-125b overexpression results in abnormally enlarged lysosomes that could contribute to impaired nutrient signalling and/or insulin biosynthesis. Finally, I optimized the usage of enhanced single-nucleotide resolution UV-crosslinking and immunoprecipitation (eiCLIP) against Argonaute 2 (AGO2 – an essential miRISC component) in EndoC-βH1 human β-cells, which will support the precise mapping of the entire β-cell miRNA targetome.

My findings provide new insights into the regulation and role of β-cell miRNAs and may contribute to the identification of future therapeutic targets for the treatment of T2D.

Acknowledgments

I would like to thank my supervisor, Dr. Aida Martinez Sanchez, for her guidance and support through each stage of the process and Professor Guy Rutter for giving me the opportunity to carry out the PhD in his section.

Thanks to all the collaborators that supported this thesis with additional experiments and/or analysis including Dr. Alejandra Thomas for generating electron microscopy images, Dr. Chris Sibley for his help with the optimization of eiCLIP and Dr. Nejc Haberman for performing the computational analyses for both SLIC-CAGE and eiCLIP. I would also like to thank Dr. Ines Cebola for her support with the transcriptomics part and Dr. Anthony Beucher for providing me with the ChIP-qPCR protocol and much more.

I would also like to thank all the members of the Section of Cell Biology and Functional Genomics at Imperial College; I have learned so much from each one of you. Marie-Sophie, Livia Lopez and Steven Parks come to mind for their everyday support and company; you were always there with a kind word of encouragement or a listening ear.

I would also like to thank Rebecca Cheung for her part in this journey. It wouldn't be so much fun without you.

Special, gigantic, thanks go to Shazi for providing me with constant support and Neda for the everyday company (inside and outside the lab). This journey would have been very different without the precious gift of your contagious joy that you both have given me aplenty.

My dear Monica, a big thank goes to you as well! As I hope you know already, your support and encouragement were worth more than I can express on paper. I consider myself very lucky to have been gifted with your friendship.

Vorrei anche ringraziare la mia famiglia per il loro costante supporto. Un grazie speciale va a mia sorella per essere sempre stata lì anche quando, inaspettatamente, si è ritrovata ad affrontare qualcosa di molto spaventoso. Sono molto orgogliosa di te, sappilo! Ringrazio infinitamente i miei

genitori che mi hanno sempre sostenuto, appoggiando ogni mia decisione, di sicuro non avrei mai potuto completare questo traguardo senza di voi. In questi ringraziamenti non possono certamente mancare Bruna e Alfredo. Grazie per il vostro sostegno e per farmi sentire parte integrante della vostra famiglia. Non vedo l'ora di riabbracciarvi tutti e festeggiare questo traguardo insieme.

Per finire vorrei ringraziare Dario, la persona che è stata costretta a sopportare i miei lamenti più di tutti, senza nemmeno avere la possibilita' di scappare a causa dei ripetuti lockdowns. Sono state avviate le pratiche per la Santificazione, non ti preoccupare! Scherzi a parte, non sarebbe stato possibile completare questo percorso senza il tuo supporto quindi grazie di cuore.

Table of Contents

List of Figures

List of Tables

Abbreviations

ACC1	Acetyl-CoA carboxylase 1
ACC2	Acetyl-CoA carboxylase 2
AGO	Argonaute
AICAR	5-aminoimidazole-4-carboxamide 1-β-D-ribofuranoside
AMPK	AMP-activated protein kinase
ATAC-seq	Transposase-Accessible Chromatin followed by high-throughput sequencing
C-13	Compound 13
C-991	Compound 991
CAGE	Cap analysis of gene expression
CAMKK2	Ca2+/Calmodulin (CaM)-dependent serine–threonine protein kinase
cAMP	Cyclic adenosine monophosphate
CDKN1B	Cyclin Dependent Kinase Inhibitor 1B
CDS	coding sequence
ChIP	Chromatin Immunoprecipitation
CLASH	crosslinking ligation and sequencing of hybrids
CLIP	Crosslinking and immunoprecipitation
CTCF	CCCTC-binding factor
CVD	Cardiovascular disease
DGCR8	George syndrome critical region 8
DMEM	Dulbecco's modified eagle media
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
eiCLIP	enhanced Individual nucleotide resolution CLIP
EM	Electron Microscopy
EMT	Epithelial-to-mesenchymal transition
ER	Endoplasmic reticulum
ES	Enrichment Score
FAS	Fatty acid synthase
FBS	foetal bovine serum
FFAs	Free Fatty Acids
Foxa2	Forkhead box protein A2
GDM	Gestational diabetes mellitus
GLP-1	Glucagon-like peptide-1
GLP-1R	Glucagon-like peptide-1 receptor
GLUT1	Glucose transporter type 1
GLUT4	Glucose transporter type 4
GNPAT	Glyceronephosphate O-Acyltransferase
GSEA	Gene Set Enrichment Analysis

GSIS	Glucose-stimulated insulin secretion
GWASs	Genome-wide association studies
HNF4α	Hepatocyte Nuclear Factor 4α
hnNRPC	Heterogeneous nuclear ribonucleoproteins C1/C2
hnRNPs	Heterogeneous nuclear ribonucleoproteins
iCLIP	Individual nucleotide resolution CLIP
IP3	Inositol 1,4,5 -trisphosphate
IP-RNA	RIP-seq
ISGs	Insulin secretory granules
KATP	ATP-dependent K+ channel
LKB1	Liver kinase B1
M6P	Mannose 6-Phosphate
M6PR	Mannose 6-Phosphate Receptor
MAFA	V-maf musculoaponeurotic fibrosarcoma oncogene homologue A
MeCP2	Methyl CpG-binding protein 2
miRISC	miRNA inducing silencing complex
MTFP1	Mitochondrial fission process protein 1
mTORC1	Mammalian target of rapamycin complex 1
NES	Normalized Enrichment Score
NEUROD1	Neurogenic differentiation 1
PAR-CLIP	Photoactivatable Ribonucleoside Enhanced CLIP
PC1	Endoproteases prohormone convertase
PCR	polymerase chain reaction
PDAC	Pancreatic duct adenocarcinoma
PDX1	Pancreatic and duodenal homeobox- 1
piRNAs	PIWI-interacting RNA
qRT-PCR	Quantitative Real-Time PCR
reverse ChIP	Reverse chromatin immunoprecipitation assay
RIP	RNA immunoprecipitation
ROS	Reactive oxidative species
RP	Reserve pool
RRP	Readily releasable pool
RT	Reverse transcription
siRNA	Short interfering RNAs
SLC30A8	Solute Carrier Family 30 Member 8
SLIC-CAGE	Super-Low Input Carrier-Cap analysis of gene expression
SMI	Size matched input
SNARE	N-ethylmaleimide-sensitive factor attachment protein receptor
SNPs	Single nucleotide polymorphism
SREBP1c	Sterol regulatory element-binding transcription factor 1
STAT3	Signal transducer and activator of transcription 3

T1D	Type 1 Diabetes
T2D	Type 2 Diabetes
TF	Transcription factor
TOR2A	Torsin Family 2, member A
TPM1	Tumour suppressor tropomyosin 1
TRBP	Trans-activation response RNA-binding protein
T-RNA	Total RNA-seq
TSS	Transcription start site
ULK1	Autophagy-initiating kinase
UTR	Untranslated region
VSMC	Vascular smooth muscle cells
XPO5	Exportin 5 protein

Introduction

1.1 Pancreas and glucose homeostasis

1.1.1 Glucose homeostasis

The human body is highly specialized in maintaining normoglycemia, the normal level of glucose in blood, within a very narrow range of 4-6 mmol/l [1]. This is accomplished by a complex regulatory neurohormonal system [2] that suppresses the release of pancreatic insulin in hypoglycaemia, which happens when glucose levels in blood drop (~3mmol/l), for example after exercise [3] or moderate fasting [4], and stimulates the release of counterregulatory hormones such as glucagon in order to increase blood glucose concentration through hepatic glycogenolysis and gluconeogenesis [5]. While the first promotes the breakdown of glycogen, a polysaccharide that serves as a form of energy storage, the second is a metabolic pathway that generates glucose from certain non-carbohydrate carbon substrates. On the other hand, when the body senses a 0.5 mmol/l increment in plasma glucose it stimulates insulin release from pancreatic β-cells and suppresses glucagon secretion from alpha cells to prevent further increases and to restore normoglycemia [2]. Insulin from the pancreas moves to muscle and adipose tissue, binds its specific receptors and enables the insulin-dependent uptake of glucose. In addition, insulin also induces formation of glycogen from sugar via glycogenesis in liver, muscle and adipose tissues [6, 7], fat production in liver and adipose tissue through lipogenesis [6, 8] and incorporation of amino acids into proteins in muscles [9]. The fact that the body finely regulates all these processes is of critical importance to human health due to the central importance of glucose as a source of energy, and because brain tissues are not able to synthesize it. Therefore, maintaining adequate glucose levels in the blood is necessary for survival [1].

1.1.2 Pancreas

The Pancreas is an organ characterised by two functionally and morphologically different tissues, exocrine and endocrine, that develop from one single epithelium [10]. The exocrine pancreas con-

sists of ductal and acinar cells and its main role is to secrete digestive enzymes into the intestine[11]. Most of these enzymes are secreted as precursors and then activated once they reach the duodenum where their main role is to catalyse the breakdown of proteins, carbohydrates, and lipids [11, 12]. The endocrine pancreas instead consists of five main different types of cells grouped in the islets of Langerhans (Figure 1.1), which are compact spherical clusters enclosed in the exocrine tissue [11]. These clusters of cells secrete hormones into the bloodstream that are responsible for the regulation of glucose homeostasis.

1.1.3 Islets of Langerhans

The islets of Langerhans occupy 1-2% of the total mass of the pancreas [13] and play a key role in the regulation of metabolism at the systemic level. More than 1 million islets are present in a human pancreas, each of them consisting in a group of thousands of cells [13]. The most abundant cells are β-cells, which constitute 50-70% of the islets in humans [14] and 60-80% in mice [15]. They are responsible for the secretion of insulin, the hormone that promotes the absorption of glucose and other carbohydrates from the blood into liver, skeletal muscle and fat tissues, and amylin, which has a role in decreasing gastric emptying and stimulate satiety after food intake [16]. Another cell type accounting for 35% and 10-20% of the islet cells in human and mouse respectively, are α-cells. These are responsible for the secretion of glucagon, the hormone that mobilizes hepatic glucose in the fasting state to maintain normal level of glucose in the bloodstream [17]. Other cell types present in the islets of Langerhans are the somatostatin-producing δ-cells and the pancreatic polypeptide-producing PP cells [11]. These occupy a small fraction of the islets and, while δ-cells produce somatostatin to inhibit the secretion of pancreatic hormones including glucagon and insulin, PP cells produce pancreatic polypeptide to suppress exocrine pancreas secretion and possibly to slow gastric emptying [18]. Finally, a very small fraction (1%) of the islets is formed by the ε-cells responsible for the secretion of ghrelin, the hormone that stimulates appetite[19]. Humans and mice have a different spatial distribution of the endocrine cells within the islets: in humans, they don't distribute with a specific organization while in mice the β-cells occupy the core of the islet and the other endocrine cells are on the periphery [20].

Due to their important role in secreting the hormones directly into the bloodstream, a significant volume of the islets (7–8%) is made by capillaries [21]. Because of this high vascularization, even if the islets constitute only the 1-2% of the total mass of the pancreas, they receive 15% of total pancreas blood flow [22].
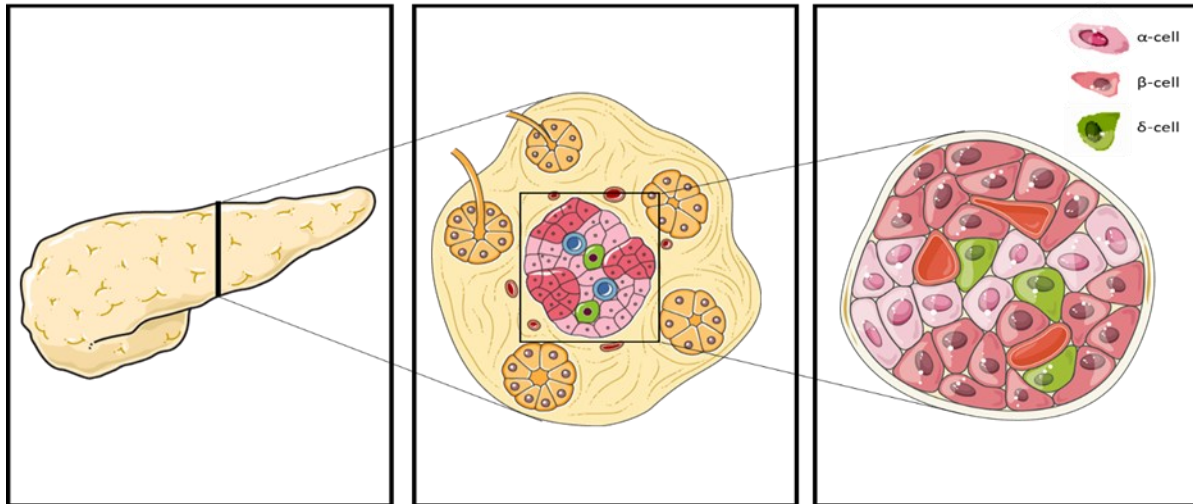
Figure 1.1: Schematic representation of the endocrine pancreas
The human pancreas is an organ characterized by both endocrine and digestive exocrine function. The exocrine pancreas occupies the majority of the pancreatic mass and is responsible for the secretion of digestive peptides. The endocrine pancreas is composed of five different cell types that form island-like structures known as islets of Langerhans and regulate the levels of glucose in the blood through the regulated secretion of different types of hormones into the circulation. This figure was created using Servier Medical Art templates (https://smart.servier.com)

1.1.4 Pancreatic β-cells and insulin secretory granules

Pancreatic β-cells are highly specialized cells able to secrete insulin in response to increased level of nutrients in the bloodstream. In fact, using specific transporters (see section 1.2.1) they uptake glucose from the bloodstream and subsequently activate a cascade of molecular events that culminates with the exocytosis of insulin. Insulin is stored in β-cells within secretory granules (ISGs), small vesicles with a diameter of approximately 300nM that, under the electron microscope, appear to have a central dense core surrounded by a halo[23] (Figure 1.2 A), with each β-cell containing a total of ~9000 granules. While most of these granules are localized into the cytoplasm, 7% are pre-docked beneath the cellular surface [24] and are involved in the rapid insulin secretion

known as the "first phase", constituting the readily releasable pool (RRP)[25]. The remaining granules are called reserve pool (RP) and are responsible for the prolonged insulin secretion called 'the second phase' where insulin is released in a slower and more sustained manner over time (Figure 1.2 B) [26].
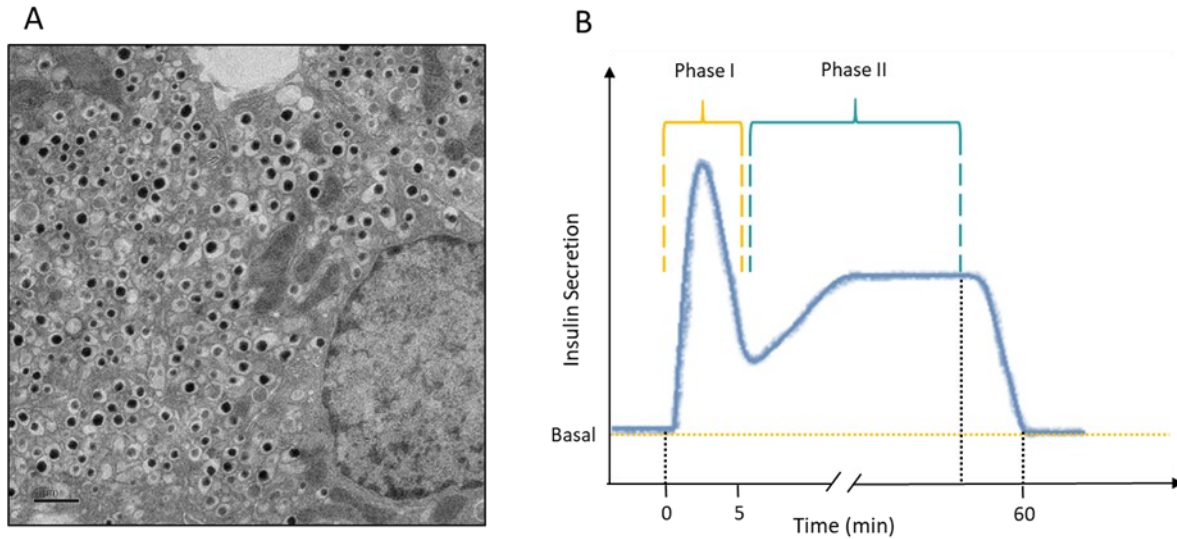
Figure 1.2: Representation of insulin secretory granules and insulin secretion phases in response to hyperglycaemia.
A. Representative β-cell from lean C57BL/6 mouse containing abundant mature secretory granules (dark and dense granules) and few immature granules containing pro-insulin (light granules). Image from Boland et al. [27].
B. Schematic representation of insulin secretion phases in response to hyperglycaemia. High level of glucose in the blood triggers a large and rapid insulin secretion (Phase I) before falling to levels above basal for an extended period of time (Phase II).

1.2 Glucose-stimulated insulin secretion

1.2.1 Glucose import and metabolism in pancreatic β-cells

Glucose uptake from β-cells is mediated by a family of glucose transporters (GLUTs) that are expressed on the plasma membrane (Figure 1.3). In both humans and rodents, this family is characterized by 14 members, each of them with distinct amino acid sequences, kinetic properties, substrate specificities, and tissue and cellular localisations [28]. However, while in rodents glucose

transporter number 2 (GLUT2) is responsible for glucose uptake [29], human islets predominantly express GLUT1 and, to a much lesser extent, GLUT3 on their surface [30]. These transporters with different capacity and affinity for glucose act as glucose sensors [31] and are important for maintaining normal glucose metabolism. In fact, mice deficient of GLUT2 are hyperglycaemic and have elevated plasma levels of glucagon and free fatty acids. These mice also present abnormal glucose tolerance *in vivo* and loss of glucose stimulated insulin gene expression *in vitro* with a substantial loss of the first phase of insulin secretion [32]. Additionally, a rescue experiment that used transgenic expression of GLUT1 or GLUT2 in the β-cells of the same mice was able to restore normal glucose-stimulated insulin secretion (GSIS) and glucose-stimulated insulin biosynthesis [33]. Moreover, genome-wide association studies have reported that GLUT2 variants increase the risks of fasting hyperglycaemia, transition to type 2 diabetes, hypercholesterolaemia and cardiovascular diseases[34].

Once inside the β-cell, glucose is phosphorylated to glucose-6-phospate by the high Michaelis constant (Km) enzyme glucokinase (GK, hexokinase IV) [29],[31, 35] and enters the glycolysis process. Glucokinase is able to regulate the rate of glucose phosphorylation acting as a rate-limiting factor for glucose metabolism and for this reason is widely considered a glucose sensor [35-37]. In humans, a mutation in GK that causes low stability or low activity of the enzyme can be in part responsible for the maturity-onset diabetes of the young (MODY)[38, 39], while in mice an ablation or reduction of glucokinase activity leads to impaired glucose-stimulated insulin secretion [40]. During glycolysis (Figure 1.3), glucose is transformed into two products, pyruvate and NADH, both of which enter the mitochondria to produce ATP, although at different stages of the respiratory pathway [41].

Pancreatic β-cell membranes are characterized by the presence of ATP-sensitive K^+ channels (K_{ATP}) that have a fundamental role in the first phase of insulin secretion. The channel is an octameric complex formed by four inwardly rectifying potassium (K^+) Kir6.x-type subunits surrounded by four sulfonylurea receptors (SUR1) [42]. Increased glucose in the cells induces ATP accumulation that binds the cytosolic ATP-binding site located on each Kir6.x-type subunit and leads to the closure of ATP-sensitive K^+ channels [43]. Such closure prevents the efflux of K^+ ions which causes a plasma

membrane depolarization from -70 mV, its resting potential, to the -30 to -40 mV range [44]. At this potential, the high voltage gated Ca^{2+} channels are activated, Ca2+ enters the cell and its influx leads to exocytosis of the insulin-containing granules (Figure 1.3) [45, 46]. β-cells surface is characterized by several types of high voltage-gated Ca^{2+} channels including L-, N-, P/Q- and R-type [47], however, L-type Ca^{2+} channels are the most abundant and therefore believed to be mainly responsible for KATP channel-regulated Ca^{2+} influx [48].
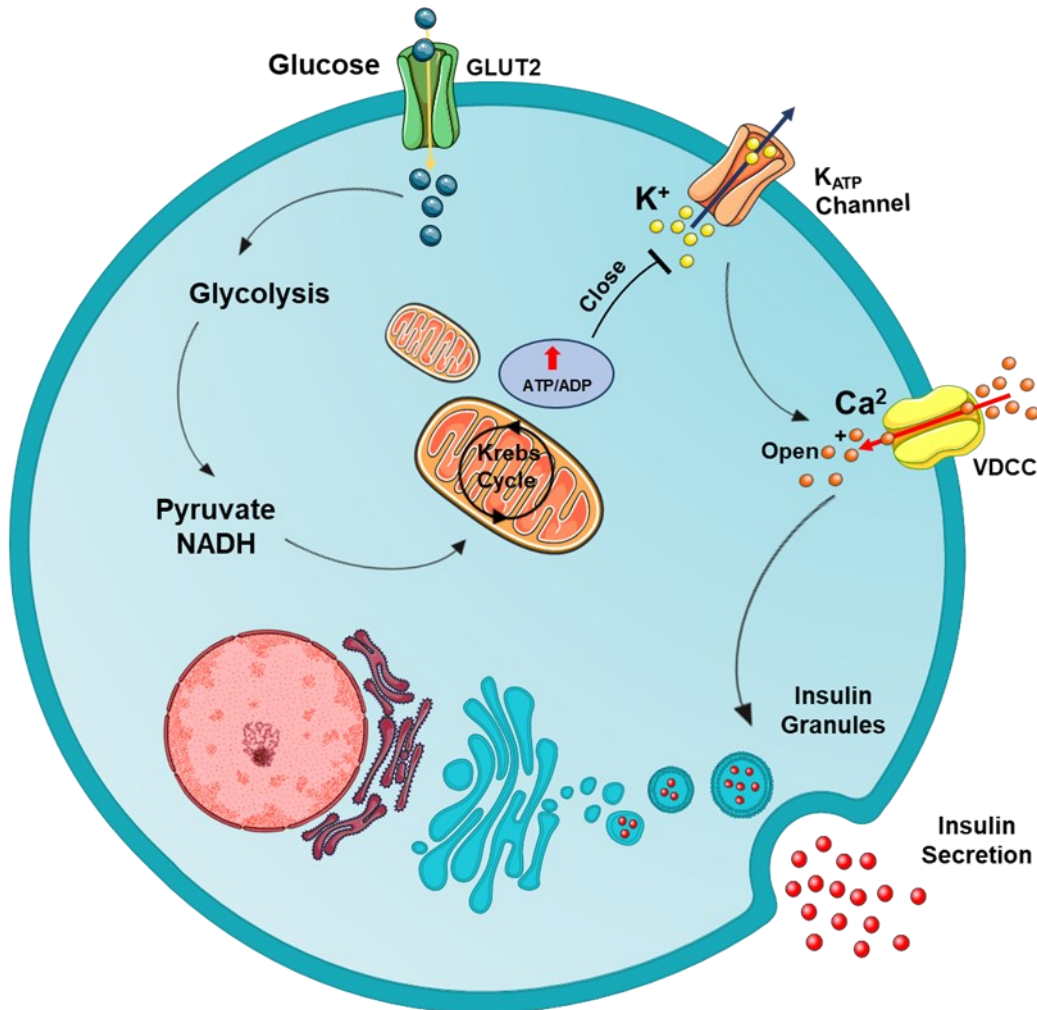
Figure 1.3: Schematic representation of glucose-stimulated insulin secretion in pancreatic β-cells. Glucose enters the β-cell by glucose transporters and is metabolized by oxidative glycolysis that leads to an increase in the ATP/ADP ratio. Increased levels of ATP results in the closure of ATP-sensitive K+ channels (KATP) that then induces plasma membrane depolarisation and the opening of voltage-dependent Ca2+ channels (VDCC). Calcium influx triggers the exocytosis of insulin granules.

1.2.2 Insulin structure, transcription and expression

The human insulin molecule has a molecular mass of 5.8 kDa and consists of 51 amino acids divided in two short peptide chains linked by two invariant disulphide bridges. Transcribed as a precursor of 12kDa (pre-proinsulin) from a gene located on the short arm of chromosome 11, this gene is characterised by three exonic regions interspersed with two introns. Exon 2 encodes for the signal peptide, a part of the C-peptide and the B-chain, while the A-chain and the remaining portion of C-peptide are encoded by exon 3 [49]. In rats and mice insulin is transcribed from two genes, *Ins1* and *Ins2*, with only the second being homologous to human insulin [50].

The transcription of the insulin gene is mainly controlled by a 340bp conserved region located upstream of the transcription starting site identified as the insulin promoter. The insulin promoter is characterized by a number of cis-acting sequence motifs contained within the 5' flanking region of the insulin gene which serve as binding sites for β-cell specific transcription factors [51, 52]. In fact, many transcription factors have been identified to play a role in the regulation of insulin transcription, with *Pdx-1* (pancreatic and duodenal homeobox- 1), *MafA* (V-maf musculoaponeurotic fibrosarcoma oncogene homologue A) and *NeuroD1* (neurogenic differentiation 1) having a crucial role in the glucose-mediated insulin gene transcription [53-55]. Glucose, as a major nutrient regulator of pancreatic β-cell function, plays a fundamental role in the regulation of insulin gene expression acting to promote the binding of these transcription factors to Cis-regulatory elements [56]. Furthermore, it has also been demonstrated that transcription factor regulation can also occur through recruitment of specific co-activators. *Pdx-1* and *NeuroD1*, for example, recruit p300 histone acetylase [57, 58], which through the acetylation of histone H4 contributes to the glucose-mediated insulin gene expression regulation [59].

Mature insulin is the result of enzymatic cleavages (Figure 1.4) that start in the endoplasmic reticulum (ER) where the pre-pro-insulin, consisting of a signal peptide, C-peptide, chain A and chain B peptides, loses its N-terminal signal and is converted in proinsulin (MW 9-10 kDa) by the enzyme signal peptidase [60]. This cleavage is important for the folding of proinsulin and the oxidation of the three cysteine disulphide bonds between the A and B chain (Figure .4). The proinsulin is then transferred to the Golgi apparatus where it is packaged into dense storage granules, cleaved by

two endoproteases, prohormone convertase PC1 and PC3, and processed by the exoprotease carboxypeptidase H (CPH) into C-peptide and mature insulin. Zinc is then imported into the secretory granule causing hexamerisation of insulin and allowing insulin to be stored and packaged within Insulin secretory granules (ISGs) (Figure 1.4).[23, 61].
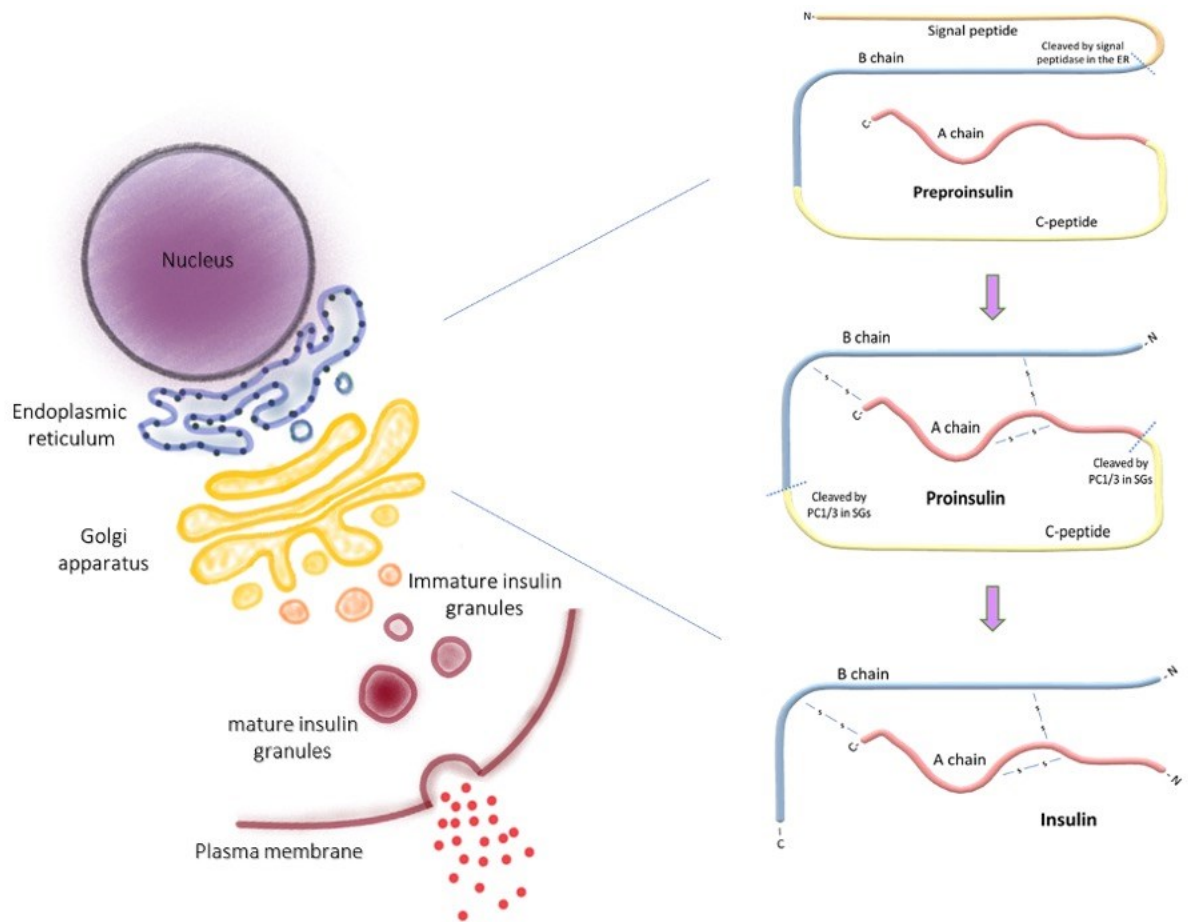
Figure 1.4: Insulin biosynthesis

Insulin is synthesised as a 110 amino-acid-long preproinsulin that consists of a signal peptide, a B chain, a connecting peptide and an A chain.The signal peptide targets the preproinsulin to the endoplasmic reticulum (ER), where it is cleaved by the signal peptidase and converted into proinsulin. In the ER, three disulfide bonds are formed between cysteine residues. Then, proinsulin is trafficked from the ER to the Golgi apparatus and packed into immature secretory granules (SGs). Before insulin release, the C-peptide is cleaved by PC1/3 to form mature insulin. During insulin secretion both C peptide and insulin are released from the cell.

1.2.3 Insulin granule dynamics and exocytosis

The fusion of insulin secretory granules to the β-cell membrane and the consequential release of insulin are the final steps of glucose-stimulated insulin secretion.

As previously described, GSIS is characterised by a biphasic time course with an initial peak which initiates rapidly and only lasts a few seconds, followed by a slowly developing but sustained second phase. Several studies (as reviewed in [62]) have suggested a relationship between biphasic insulin secretion and the fact that β-cells are characterized by two distinct populations of secretory granules. In fact, a smaller percentage (~7%) is formed by granules docked beneath the cell membrane, constituting the readily releasable pool (RRP) [24] that might be responsible for the rapid first phase of insulin secretion that occurs within 10 minutes after glucose-stimulated cytosolic Ca^{2+} increase [25]. The bigger percentage is known to be the reserve pool (RP), consisting of granules which are distant from the cell membrane and could play a role in the sustained second phase of insulin secretion [26].

It has been suggested that the recruitment of granules in RP to the RRP might occur through the microtubule network. The microtubule-based granule transport may act together with actin-based transport in the subplasmalemmal space to mobilize the granules and, when this process needs to be limited for example at low glucose concentration (basal condition), actin limits the directed microtubule-based motion [63]. Interestingly, it has also been suggested that microtubule organization and dynamics itself may spatially dictate where on the plasma membrane the vesicle can fuse [64].

It has also been suggested that cyclic adenosine monophosphate (cAMP) might have a role in the regulation of insulin granule dynamics. cAMP is an important 'second messenger' signalling molecule [65] whose intracellular levels are increased by activation of G-protein coupled receptors (GPCRs), which engage the actions of adenylate cyclase proteins through the activation of Gs proteins and produce cAMP from ATP [66]. cAMP regulates exocytosis in various secretory cells [67, 68] and it has been reported that after membrane depolarization levels of cAMP increase simultaneously with Ca^{2+} [69, 70]. Moreover, cAMP signalling has been associated with potentiation of

both first and second phase of insulin secretion without interfering with the number of docking granules present in the cell or with fusion events [71]. However, it appears that the activation of cAMP increases the size of both RRP and RP pools of insulin granules, suggesting that it could have a role in controlling the density of the granules near the plasma membrane [71].

The molecular mechanism involved in insulin exocytosis is similar to the one that controls neuro-transmitters release [72]. Several proteins participate in the insulin secretory vesicle fusion process, such as the soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) [73]. SNARE is a transmembrane protein family located in cell organelles and vesicles. In β-cells, it is essential for the fusion of the insulin granules to the membrane by a mechanism that involves the assembly of a complex consisting of VAMP-2 (R-SNARE) on the granule membrane, syntaxin-1a (Qc-SNARE) on the plasma membrane, and the membrane-associated protein SNAP-25 (Qa-Qb SNARE) [73, 74]. These proteins allow the vesicle membrane to come in closer contact with the plasma membrane and therefore fuse with it facilitating the release of insulin. The energy required for this process is provided by the conformational changes of the proteins themselves [73]. It has also been suggested that, in mouse β-cells, SNARE proteins have a role in Ca^{2+} influx in the areas where the plasma membrane is in contact with the secretory granules [75].

1.2.4 K_{ATP} channel-independent amplifying signals

Many experiments have confirmed the key role of K_{ATP} channels in controlling insulin secretion [76-78], however it is still possible for glucose to induce insulin secretion while K_{ATP} channels are partly open. In fact, it has been found that the potentiation of Ca^{2+}-induced exocytosis by glucose can be independent from the state of the K_{ATP} channels [79-81] and that glucose (and other metabolites) metabolism can generate additional signals that amplify insulin secretion [82]. Changes in the cytosolic ATP/ADP ratio has certainly been the first, but many others are included in this still growing list. Free Fatty Acids (FFAs) are one of them: when FFAs are transported into the cells they can be activated into long-chain acyl-CoA and therefore metabolised to generate lipid signalling molecules that initiates signals for insulin exocytosis [83]. FFAs can also bind and activate the G-

protein–coupled receptor FFAR1 (GPR40) on the cell membrane and initiate a cascade that increased intracellular levels of Ca^{2+}, mostly induced by raised levels of inositol 1,4,5 -trisphosphate (IP3) leading to the amplification of insulin secretion [84].

Increased levels of glutamate following high glucose challenge have also been associated with increased insulin exocytosis [85] . A recent study showed that cytosolic glutamate produced by glucose metabolism is key for the amplification of incretin-induced insulin secretion signal [85]. Incretin hormones are gut peptides that are secreted from different enteroendocrine cells into the bloodstream in response to food intake [86]. Incretins such as glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide bind to specific G protein-coupled receptors (GPCRs) on the β-cell surface [87], increase cyclic adenosine monophosphate (cAMP), promote Ca^{2+} influx and potentiate GSIS [88] through a so-called incretin effect, a mechanism defined by the greater difference in the insulin secretory responses obtained by oral glucose administration compared to the same dose administered intravenously [87].

1.3 Diabetes Mellitus

Diabetes is a complex chronic disease characterized by increased level of glucose in the blood [89]. This occurs either because the body is not able to produce enough insulin, or the produced insulin is not functional [89]. Diabetes is recognized as an important cause of premature death and disability with over 8.5% of the world's adult population affected and more than 1.5 million deaths annually [90]. Furthermore, the impact on the economy is considerably high, with an annual worldwide cost of more than US$ 827 billion [91].

Type 1 (T1D), Type 2 (T2D) and gestational are the three main types of diabetes. However, other types of diabetes have been described in recent years [89]. Amongst those we can find maturity-onset diabetes of the young (MODY), which while sharing some of the T2D symptoms is not linked to obesity and is characterized by a stronger genetic risk factor [92], and Latent Autoimmune Diabetes of Adulthood (LADA), a form of T1D that shares some characteristics with T2D. Similarly to T1D, LADA is an autoimmune disease in which the body's immune system attacks and destroys insulin producing cells, but develops over a longer period of time, like type 2 diabetes.

Finally, the secondary diabetes group clusters together all the diabetes that are developed as a consequence of another medical condition, such as cystic fibrosis, chronic pancreatitis and glucagonoma, or drug- or chemical-induced diabetes, for example following the treatment of HIV/AIDS or organ transplantation [89].

1.3.1 Type 1 Diabetes (T1D)

T1D is a chronic autoimmune disease characterized by the body deficiency to produce insulin due to a self-defence mechanism where T-cells mediate the destruction of insulin-producing pancreatic β-cells [93]. The majority of type 1 diabetes cases are first diagnosed in children and adolescents[90] and 10% of all diabetic patients live with this condition [94]. The risk factors for T1D have not been fully understood yet, making the onset of the condition not preventable. Nevertheless, it is known that both genetic and environmental factors play a role in the development of the disease [95]. Only recently the dogma by which the pancreatic β-cells are destroyed in T1D has been challenged: recent studies have in fact suggested that some β-cells can probably evade immune destruction and survive in T1D thanks to their de-differentiated status. Rui et al. [96] showed that, in T1D mouse, a small population of β-cells characterized by lower insulin content and glucose stimulated insulin secretion was also defined by a down-regulation of genes responsible for β-cell identity and up-regulation of genes associated with immune modulation. These discoveries suggest that the acquisition of a 'de-differentiated' phenotype might be responsible for protection against the immune system attack [96].

1.3.2 Type 2 Diabetes (T2D)

T2D, also called non-insulin-dependent or adult-onset diabetes in the past, is a result of the body not being able to use insulin. It is the most common type of diabetes with almost 90% of all the cases and, although it shares similar symptoms to T1D, they are generally less marked or even absent. This often results in a delay of the diagnosis and therefore causes unpredicted severe complication that damage different organs [90]. For many years, T2D has been associated to adults only, but its incidence has recently increased in children too [97]. Worldwide, almost 463 million of adult people are living with diabetes and, according the International Diabetes Federation, this

number will increase to 700 million by 2045 if preventative actions are not taken[94]. Different factors influence the development of T2D with the most influential being associated to lifestyle. In fact, up to 80% of the cases could be prevented with a healthy diet and physical activity [98, 99]. Also, ethnicity and family history of diabetes increase the risk suggesting that genetics play a role in the development of this condition [100, 101]. However, the identification of the genes responsible for this polygenic disease has been challenging.

In 2007, results from genome-wide association studies (GWASs) made possible for the first time the identification of common genetic variants associated with T2D [102]. The first GWAS for T2D was performed on normal-weight T2D individuals of European-ancestry compared to non-diabetic individuals, and identified variants at five novel *loci* associated with increased risk of being affected by the disease [102]. Only a subsequent meta-GWAS, which involved multiple cohorts in order to increase statistical power, led to the identification of new variants such as single nucleotide polymorphism (SNPs) [103]. Since then, the study of SNPs has become one of the mainstream ways to study the association of genetic variation to T2D. During the last few years, several new SNPs have been identified in genes that have been linked to T2D [104]. One example is transcription factor 7-like 2 (*TCF7L2*), a transcription factor downstream of wingless-related integration site (Wnt)/β-catenin pathway that plays a role in embryonic development and cell differentiation and division [105]. SNPs rs7903146 and rs12255372 in intron 3 of this gene are associated with a ~ 45% increase in T2D risk per allele and carriers show impaired insulin production [106] and β-cell dysfunction [107]. SNPs in *TCF7L2* represent one of the strongest known genetic determinant of T2D [108] alongside with those found in *SLC30A8*, a β-cell specific Zn^{2+} transporter [109]. SNP rs13266634 in the *SLC30A8* gene is responsible for a change in protein structure that impairs the accumulation of zinc by granules and impedes normal insulin storage, with variant carriers showing an increased T2D susceptibility of ~ 20% per allele [102].

Many other well established factors have been identified to be responsible for the pathophysiological abnormalities that contribute to impaired glucose homeostasis [104] with β-cell dysfunction and insulin resistance being the most studied [104, 110, 111]. However, it is not a novelty that in most developed countries the prevalence of T2D is increasing in parallel with the mounting incidence of obesity and that excess adiposity leading to a high BMI is the single strongest risk

factor for T2D [112]. Also diet and lifestyle can have an important impact on the incidence of T2D. It has been demonstrated that it is possible to prevent T2D through lifestyle interventions focused on increasing physical activity and adopting a healthy diet such as the Mediterranean diet [113, 114].

1.3.3 Gestational

Gestational diabetes mellitus (GDM) generally affects pregnant women during the second and third trimesters and is mainly due to the β-cells inefficiency to compensate for the increased insulin resistance [90].

During a healthy pregnancy, insulin sensitivity shifts over the stages of the gestation to fulfil the requirements of the pregnancy. In early gestation, insulin sensitivity increases to allow the storage of energy into adipose tissue and prepare the body for the future demand [115]. Later on, placental hormones promote insulin resistance[116] resulting in increased blood glucose that is then transported to the placenta and provides sustenance for the foetus. Pregnant women compensate for this phenomenon through hypertrophy and hyperplasia of pancreatic β-cells and increased glucose-stimulated insulin secretion [117]. In fact, it has been demonstrated that during a normal pregnancy the uptake of glucose mediated by peripheral organs decreases by 50% and, as a consequence, β-cells compensate for this by increasing insulin secretion by 200% [118]. Gestational diabetes occurs when pregnant mothers are unable to increase insulin secretion to compensate for the increased insulin resistance [119].

There are several risk factors involved in this condition, including obesity, excessive weight gain during gestation, advanced maternal age, genetic polymorphisms and ethnicity [120]. Consistently, 60% of the women with a past history of GD develop type 2 diabetes [121] and have a predisposition to cardiovascular disease (CVD)[122]. GD can also have long-term consequences for the infant: it has been demonstrated that babies that are born in GD pregnancies have higher possibilities of developing T2D, CVD and double the risk of developing obesity during childhood [123].

1.4 Pancreatic islet β-cells and type 2 diabetes

1.4.1 β-cells mass and apoptosis in T2D

Different factors that occur during obesity, pregnancy or increased insulin resistance, such as hypertrophy, enhanced replication and neogenesis, lead to β-cell mass increases [124-126]. On the contrary, the progression from insulin resistance to the onset of diabetes has been associated with β-cell dysfunction and reduced β-cell mass [127, 128].

In T2D, human β-cells are characterized in average by a ~50% β-cell mass reduction, with an even more marked reduction in obese diabetics (>60%) [128]. However, these findings have been questioned [124, 129]. A study from Rahier et al. [124] carried out with a large number of samples (57 T2D and 52 non-diabetic subjects) suggested that β-cell was only reduced by ~25% within 5 years of diagnosis in T2D patients, suggesting that β-cell mass alone cannot be responsible for the onset of diabetes.

Increased apoptosis in T2D subject is generally thought to contribute to the reduction in β-cell mass. Butler et al. [128] examined pancreatic tissue from diabetic and non-diabetic subjects and found that the frequency of β-cell apoptosis was increased 10-fold in the diabetic group compared to controls. The difference was reduced to 3-fold in the case of obese T2D subjects. Since new islet formation and β-cell replication was unchanged amongst the subjects, the authors claimed that apoptosis may be the major defect leading to a decrease in β-cells mass in type 2 diabetes [128].

Several factors seem to trigger apoptosis in β-cells, including the toxic effects of hyperglycaemia (glucotoxicity) [130] and elevated levels of free fatty acids (lipotoxicity [131]. Studies have demonstrated that also endoplasmic reticulum (ER) stress may participate in β-cell death [132, 133]. It has been suggested that, as compensatory mechanism for hyperinsulinemia, β-cells start to produce additional islet amyloid polypeptide (IAPP), a polypeptide pancreatic hormone that aggregates and accumulates at both intracellular and extracellular levels and mediate permeabilization and loss of membrane β-cell membrane integrity [134]. This mechanism leads to ER stress that

causes accumulation of unfolded proteins inside the ER, which in turn activates the unfolded protein response and leads to pro-apoptotic pathway initiation [135].

1.4.2 Insulin secretory deficiency in T2D

Defective glucose mediated insulin secretion is the hallmark of β-cell failure in T2D [136]. Indeed, islets isolated from patients with T2D present a 60% reduction in glucose insulin secretion compared to non-diabetic controls and this difference is supported by reduced β-cell insulin content [137]. It has also been reported that islets from diabetic and non-diabetic patients have a different threshold for glucose stimulated insulin secretion (7 mmol/l and 12 mmol/l respectively), indicating that diabetic islets are less sensitive to glucose [138]. Interestingly, mRNA levels of glucokinase (the first enzyme of the glycolytic breakdown of glucose) were also found to be reduced in T2D islets, supporting the defective glucose sensing and metabolism observed in these studies [139]. Indeed, reduced oxidative glucose metabolism has been observed in islets from T2D donors [140] and expression of glucokinase is also reduced in islets from T2D donors [139], indicating that an impairment of glucose metabolism could compromise the ability of glucose to initiate the insulin secretion cascade. Interestingly, defective secretion of insulin can be corrected by a glucokinase activator [140]. Moreover, the fact that T2D has been associated with a reduction of expression of key proteins involved in the exocytotic process [141] and that an age-dependent onset of type 2 diabetes has been linked to a SNP in the gene of syntaxin 1A, an important protein of the SNARE complex, suggests that in type 2 diabetic β-cells might also be characterized by a defective exocytosis machinery [142, 143].

1.4.3 Hyperglycaemia and β-cell identity

Hyperglycaemia is the result of a combination of insulin resistance in skeletal muscle, liver and adipose tissue and insufficient secretion of insulin from pancreatic β-cells [144]. It occurs when the blood glucose levels are greater than 7.0 mmol/L while fasting and 11.0 mmol/L 2 hours postprandial. Whereas in the pre-diabetes state β-cells are able to secrete more insulin in order to compensate for peripheral insulin resistance, in frank diabetes β-cells lose this ability leading to chronic hyperglycaemia [145]. Prolonged plasma glucose levels can have a toxic effect in a variety

of cell types, a process called glucotoxicity [146]. While most cells are able to maintain a relatively constant intracellular glucose concentration in an hyperglycaemic environment by reducing the rate of intracellular glucose transport [147], other cell types, including pancreatic β cells, neuronal cells, and vascular endothelial cells, cannot counteract this process because they equilibrate their intracellular glucose level to that of their extracellular environment [147-149], making them highly susceptible to hyperglycaemic damage as excess glucose enters the cell. Specifically, it has been demonstrated that, in mouse β-cells, glucose induces metabolic dysfunction through decreased oxidative metabolism and the accumulation of glycogen, providing one of the possible causes of reduced insulin release, β-cell number and identity [150].

Maintenance of cellular identity requires active regulation of gene expression. In β-cells, deletion of several important transcription factors such as FOXO1 [151], PDX1 [152], NKX6.1 [153] and PAX6 [154] have been shown to perturb cell identity by silencing the expression of set of genes that allow β-cells to be functional as such. Concomitantly, the expression of genes normally suppressed or expressed at very low levels in β-cells have instead been found upregulated [155]; This genes are called β-cell disallowed or forbidden genes [156-158].

Moreover, the presence of progenitor cell markers has been observed in the dedifferentiated islets of diabetic animals [159-161] and emerging evidence also suggests that elevated blood glucose can be responsible for alterations of β-cells identity in both human and mouse models promoting trans-differentiating to other pancreatic endocrine cell types such as α cells and δ cells [162-164]. Jonas JC et al. [165] was the first who, in the early nineties, identified the major role of hyperglycaemia in the loss of β-cell identity and dedifferentiation. After him, many others have and are still working on the identification of the precise molecular mechanisms involved [130, 159, 166]. This will certainly help in the development of new therapeutic strategies that will preserve or restore the functionality of β-cell in diabetes.

1.4.4 Approaches for the treatment of diabetes

While in T1D insulin treatment is necessary, T2D patients are only administered insulin if oral hypoglycaemic drugs are not effective. In the past years, multiple antidiabetic drugs able to maintain

a durable glycaemic control have become available, which are often combined, in order to reverse the manifold pathophysiological changes present in T2D [167-169]. Metformin is currently the most prescribed antidiabetic drug [170, 171], and it specifically reduces hepatic gluconeogenesis without inducing hypoglycaemia or weight gain [172]. The mechanisms by which metformin suppresses hepatic glucose production remain unclear. It was originally believed that the activation of AMPK [173, 174] and suppression of mitochondrial glycerophosphate dehydrogenase in liver [175] were important mediators of metformin action. However, more recent studies indicate that, although metformin activates AMPK through decline in cellular energy status by reducing intracellular ATP levels [176, 177], the metformin- induced acute suppression of glucose production appears to be independent of the LKB1–AMPK signalling pathway [178].

Sulfonylureas are another group of antidiabetic drugs, which work by triggering endogenous insulin secretion from pancreatic β-cells targeting mainly the ATP-sensitive potassium channels on β-cells [179]. Given their mechanism of action, sulfonylureas are effective only in the presence of residual pancreatic β-cells and can cause hypoglycaemia and weight gain [180].

Finally, Glucagon-like peptide 1 (GLP-1) analogues have been suggested for the treatment of young obese patients with recent diagnosis of T2D. These drugs cause a pharmacological increase in plasma GLP-1 that results in increased insulin secretion and inhibition of glucagon secretion [181]. These molecules are not the only ones used to compensate for the severe GLP-1 resistance in β-cells observed in T2D. In fact, Dipeptidyl peptidase 4 (DPP4) inhibitors are also used as antidiabetic drugs since they prolong the half-life of endogenously secreted GLP-1 and gastric inhibitory polypeptide [182], which can stimulate insulin synthesis from pancreatic β-cells [183].

1.5 AMPK

1.5.1 AMPK structure and expression

AMPK is a phylogenetically conserved serine/threonine protein kinase that plays an important role in regulating energy metabolism [184]. As major cellular energy sensor and master regulator of metabolic homeostasis, AMPK has recently emerged as an interesting potential drug target for

treating diabetes [185] . AMPK is a heterotrimeric enzyme characterized by a catalytic (α1 or α2) subunit and two regulatory (β1 or β2 and γ1, γ2, or γ3) subunits, all of which are encoded by separate genes, for a total of 12 possible complexes [186-188] (Figure 1.5).
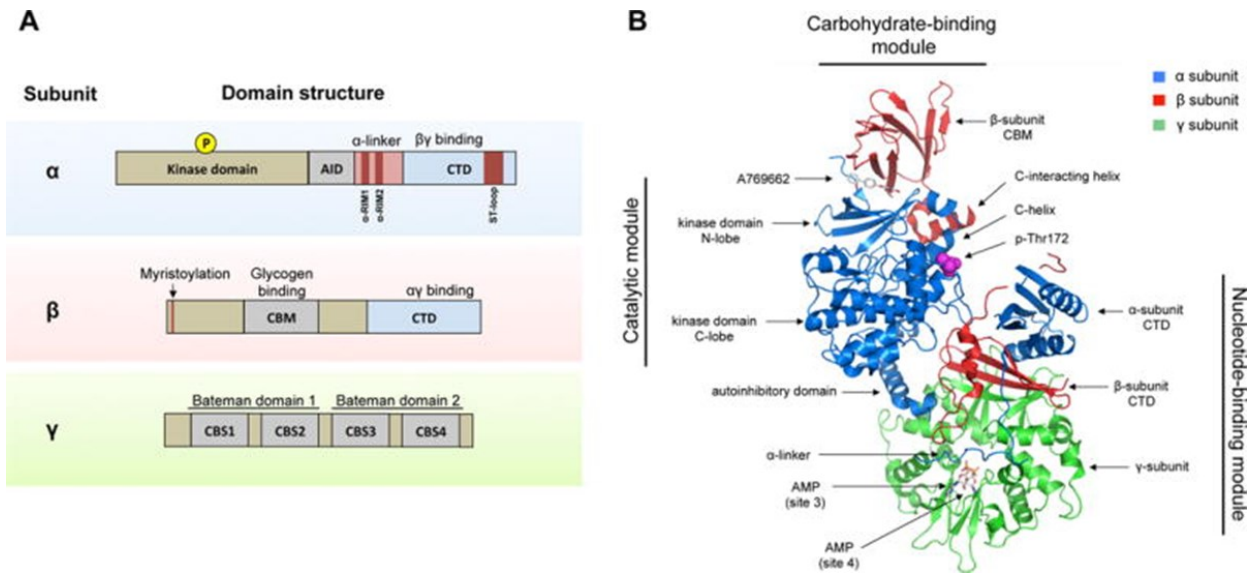
Figure 1.5: Domains and structure of the AMPK complex
A. Schematic illustration of AMPK main protein domains characterized by a catalytic subunit (α) and two regulatory subunits (β and γ).
B. Ribbon representation of AMPK α2β1γ1 trimeric complex crystal structure. The image is from Garcia et al. [189]

The α-subunit is characterized by a N-terminal Ser/Thr kinase domain, an auto-inhibitory sequence (AIS) and a β subunit binding domain at the C-terminal [184]. The β-subunit contains two conserved domains: a carbohydrate-binding module (CBM), responsible for AMPK ability to sense glycogen, and a C-terminus domain that binds the α- and γ-subunits. The β-subunit also contains a myristoylation site at its N-terminal, which facilitates the targeting of AMPK to cellular membranes [190, 191]. The γ-subunit is comprised of two Bateman domains, each containing two cystathionine β-synthase repeats (CBS). The resulting four CBS repeats form the four sites in AMPK that can potentially bind AMP, ADP or ATP in a competitive manner [184].

Interestingly, following AMPK knockout in animals it has been established that AMPKα1 is the major catalytic subunit found in both mouse insulinoma cells (MIN6) and islets, and it is also responsible for the vast majority of AMPK catalytic activity [187, 192]. It has been suggested that in islets the differential expression and localization of the multiple isoforms that compose AMPK subunits may contribute to this differences in AMPKα catalytic activity [193].

1.5.2 Regulation of AMPK activity

The activity of AMPK is mainly regulated by nutrients, such as glucose and amino acids [186, 194], but also calcium, hormones, and cellular stress can trigger the phosphorylation of different targets involved in protein synthesis, ATP and glucose metabolism [195]. AMPK uses these signals of energy availability to regulate cell function, metabolism (anabolism and autophagy) and survival (apoptosis)[196].

The phosphorylation of Thr172 in the α-subunit is the principal event required for full activation of AMPK [197, 198]. Thr172 phosphorylation is mediated by upstream activating kinases including LKB1 (Liver kinase B1), CAMKK2 (Ca2+/Calmodulin (CaM)-dependent serine–threonine protein kinase) and TAK1 (TGF-β activated kinase 1) [199, 200] and by allosteric activation mediated by adenine nucleotides that bind to the Ɣ subunit [161].

Cells normally maintain their ATP:ADP ratio at about 10:1. If this ratio declines because of higher demand of ATP consumed by energy-requiring processes, more AMP or ADP is available to bind AMPK and therefore stimulates its activation. The rise in AMP availability and its binding to the Ɣ-subunit of AMPK triggers three complementary mechanisms [185, 198]:

- Increased LKB1 mediated phosphorylation of Thr172 on AMPK α subunit
- Protection of Thr172 from dephosphorylation by protein phosphatases
- increased activity by allosteric activation of already phosphorylated AMPK

LKB1 appears to be constantly active and is responsible for the activation of AMPK in most of the tissues [201]. Indeed, LKB1 is crucial for the activation of AMPK in β-cells since its loss has been shown to abolish both AMPK and its target proteins phosphorylation [193]. It has been estimated that the combination of LKB1 mediated phosphorylation and protection from dephosphorylation

of Thr172 is able to increase AMPK activity by 100-fold while the allosteric activation is responsible for a further activation of 10-fold, for a total of 1000-fold [202].

On the contrary, the phosphorylation of AMPK by CAMKK2 does not require changes in AMP and ADP concentrations and is therefore described as a nucleotide-independent regulation [184]. CAMKK2 is activated by increases in intracellular Ca^{2+} concentration, which occur in response to different hormones such as vascular endothelial cell growth factor (VEGF) in endothelial cells where it promotes angiogenesis [203] and ghrelin in the hypothalamus [204, 205] where it triggers feeding behaviour. Once activated, CAMKK2 can phosphorylate Thr172 on the AMPK α subunit [199, 206]. The activation of AMPK by CAMKK2 has been already implicated in angiogenesis and in the response of neurons to the sensation of hunger[203, 204], but it was only recently that CAMKK2 was suggested as an important component of the Ca^{2+}-mediated response to glucose signalling in β-cells [207]. The exact mechanism of this regulation and the potential role of AMPK activation remains to be explored.

1.5.3 AMPK role in energy homeostasis

The activation of AMPK has two main roles in the restoration of energy balance: activate catabolic enzymes involved in the generation of ATP and deactivate anabolic enzymes that are involved in processes driving the consumption of ATP such as synthesis and storage of macromolecules [202].

Together with insulin, activated AMPK promotes muscle glucose uptake through the translocation and fusion of glucose transporter type 4 (GLUT4; also known as SLC2A4) intracellular storage vesicles to the plasma membrane. AMPK, in fact, phosphorylates and triggers the dissociation of TBC1 domain family member 1 (TBC1D1) from the GLUT4 vesicles, which otherwise acts as an inhibitor, promoting the fusion of the vesicles to the plasma membrane [208, 209]. Similarly, AMPK promotes glucose uptake by glucose transporter type 1 (GLUT1) in cells that are only characterized by this type of receptor (such as adipose tissue and brain) [210].

AMPK is also involved in the first steps of glycolysis where it phosphorylates and activates two isoforms of phosphofructokinase (PFK), the enzyme responsible for the conversion of fructose-6-phosphate to fructose-2,6-bisphosphate (F26BP) [211].

AMPK regulates glucose homeostasis also by inhibiting glucose production by the liver. In fact, the activation of AMPK was showed to suppress the transcription of genes encoding for important enzymes involved in the gluconeogenesis pathway such as phosphoenolpyruvate carboxykinase (PEPCK), glucose-6-phosphatase (G6Pase) [212] and, not surprisingly, AMPK α2 liver-specific knockout mice show glucose intolerance and have high fasting glycemia [213].

In addition to gluconeogenesis, AMPK has been implicated in regulation of liver lipogenesis and lipid oxidation. On one hand, AMPK has been shown to suppress glucose-induced expression of lipogenesis-associated genes such as fatty acid synthase [174], whereas through the activation of the fatty acid transporter CD36 it promotes the uptake of fatty acids into cells [214] to facilitate the fatty acid oxidation mediated by the mitochondria. To help this process AMPK also enhances the expression of oxidative enzymes that are part of the TCA (Krebs) cycle [215]. Interestingly, the uptake of fatty acid into mitochondria mediated by AMPK can also be promoted through the phosphorylation of Acetyl-CoA carboxylase 2 (ACC2) and therefore inhibition of malonyl-CoA [216].

In conditions where glucose levels are low, such as during starvation, AMPK has been shown to promote autophagy. This mechanism allows the cells to survive by recycling nutrients from organelles, proteins, and different macromolecules that are delivered to the lysosomes for degradation [217]. Kim et al., [218] have proposed a mechanism by which AMPK is able to promote autophagy by direct phosphorylation and activation of the autophagy-initiating kinase (ULK1). However, they also show that this mechanism is antagonised by mTOR activity under nutrient sufficiency. mTOR, which role in autophagy inhibition has already been widely established [219, 220] phosphorylates ULK1 on a different site from AMPK and by doing so disrupt the interaction between the two proteins [218].

Interestingly, in mammals, AMPK can also influence metabolism and energy balance at the whole-body level through its actions in the hypothalamus. The hypothalamic area of the brain in fact represents a key compartment for energy sensing and satiety regulation: stimulation of neurones that express neuropeptide Y (NPY) and agouti-related protein-expressing neurons (AgRP) induces feeding, while stimulation of neurons expressing pro-opiomelanocortin (POMC neurons) [G] inhibits it. It has been demonstrated that, in rodent, hormones inhibiting feeding such as leptin are

able to inhibit AMPK α2 isoform and, similarly, those promoting feeding such as ghrelin can acti-vate AMPK [221] . In line with this, activation of AMPK mediated by injections of 5-aminoimidazole-4-carboxamide 1-β-D-ribofuranoside (AICAR) into the hypothalamus promoted feeding [205] whereas expression of dominant negative AMPK in the hypothalamus was enough to reduce food intake [221].

1.5.4 AMPK molecular mechanism of action in pancreatic β-cell

In the past few years, AMPK has been shown to regulate many proteins in a disparate number of cell types [222-226]. Some of these regulations have been also confirmed in in β-cells [222, 227-229], however, this area of research still remains poorly understood.

In line with its important role in energy homeostasis, the activation of AMPK in pancreatic β-cells targets several proteins such as enzymes, transporters and transcription factors involved in im-portant metabolic pathways [193]. For example, one of the best characterized AMPK targets is Acetyl-CoA carboxylase 1 (ACC1). ACC1 catalyses a rate-limiting step in fatty acid biosynthesis and AMPK-mediated phosphorylation has been shown to inactivate this protein leading to inhibition of cholesterol and de novo fatty acid synthesis [228]. For these reasons, targeting AMPK and/or ACC1 could be a strategy for treating obesity and insulin resistance [196]. Moreover, it has been demonstrated that silencing of *Acc1* using RNA interference impairs GSIS alongside with a de-crease in glucose oxidation ATP:ADP in rat insulinoma cell line INS-1. Furthermore, chronic, but not acute, treatment of *Acc1* inhibitors impairs GSIS [228].

Another well characterised AMPK target is mTOR. AMPK inhibits mTOR and protein synthesis through phosphorylation and activation of Tuberous Sclerosis Complex 2 (TSC2) protein [226]. In-terestingly, mice lacking p70S6K1, a downstream target of mTOR, are glucose intolerant and pre-sent hypoinsulinemia and reduced cell size demonstrating that the regulation of this pathway is crucial for pancreatic β-cells [230].

AMPK also regulates β-cell gene expression to modulate lipogenesis and insulin production [229, 231]. The role of Sterol regulatory element-binding transcription factor 1 (SREBP1c) in the up-reg-ulation of lipogenic genes such as FAS (fatty acid synthase) is well studied [232, 233]. Interestingly,

Diraison et al. [229] found that, while SREBP1c-transduced islets displayed increased levels of FAS alongside with triacylglycerol, the treatment with AMPK activators was able to decrease the expression of the endogenous SREBP1c and FAS and, more importantly, to reverse the effect on increased glucose oxidation, ATP levels, and decreased glucose-induced insulin secretion deriving from overexpressing SREBP1c [229].

Interestingly, by dephosphorylating and destabilizing Hepatocyte Nuclear Factor 4α (HNF4α), AMPK may also have a role in the β-cell insulin synthesis [231, 234]. Expression of active AMPK α1 in INS-1 cells [231] or treatment of mouse islets with AICAR [234] has been associated with decreased levels of HNF4α protein. Interestingly, HNF4α transcription factor has previously shown to mediate the indirect [235] and direct [236] activation of the insulin gene promoter in β-cells.

1.5.5 AMPK role in pancreatic β-cells and insulin secretion

AMPK has recently emerged as a key regulator of insulin secretion and β-cell survival and proliferation [192]. In the past years many groups have tried to clarify the connection between insulin secretion and the LKB1-AMPK pathway. After more than 20 papers and counting, the topic is far from being settled since both positive and negative regulation has been reported after AMPK activation [160, 193, 237].

The first study that addressed the role of AMPK in controlling insulin secretion was published in 1998 [186]. Salt and colleagues demonstrated that INS-1 rat β-cells treated with AICAR showed inhibited GSIS whereas the basal insulin secretion was increased [186]. However, in the same study stimulatory effects on GSIS were observed after AICAR treatment on isolated rat islets. A few years later, these two conflicting results were confirmed by others. Akkan and Malaisse [238], along with other groups [160, 186, 237], demonstrated that increasing doses of AICAR administered alongside increasing concentrations of glucose caused increased insulin secretion in both isolated pancreatic rat islets or perfused pancreases, suggesting that the activation of AMPK promotes insulin secretion.

On the contrary, other studies performed with different insulinoma cell lines (INS1, MIN6 and HIT) and pharmacological activation of AMPK with AICAR [187, 194] and more consistently with the

anti-diabetic drugs thiazolidinediones (TZDs) [239-241], found AMPK activation to reduce insulin secretion. While acute activation of AMPK results in varied GSIS responses, a long-term activation has been more consistently shown to inhibit insulin secretion in both β-cell lines and islets [242-244].

Two independent groups generated an AMPKα2 β-cells and brain mice knock-out using RIP2-cre in the background of a whole-body knockout of AMPKα1 [192, 245]. In both studies the mice presented impaired glucose tolerance and reduced plasma insulin while β-cell mass was slightly increased, although the difference was not significant. Sun et al. [192] showed that AMPKα1/α2 KO β-cells had an increased cell proliferation and reduction in size. Insulin secretion in isolated islets was increased in both low glucose and high glucose conditions. Beall et al. [245] confirmed the increased insulin secretion under low glucose concentration, but the islets showed a defect in GSIS. The discrepancies in the age of the animals could explain the different phenotypes, however both studies couldn't be used to draw a conclusion on how the AMPK catalytic subunit deletion can impact insulin secretion. In fact, global loss of α1 or loss of both α1 and α2 using *Cre* recombinase could masks potential β-cell dependent defects as this protein may be expressed in neurons involved in the regulation of energy and glucose homeostasis themselves [246].

In order to avoid the complications associated with the deletion of AMPK in the brain following an *Ins2Cre*-mediated recombination, Kone et al. [247] used a knock-in mouse with the *Cre* recombinase in the *Ins1* locus. These mice, in which both catalytic subunits of AMPK were deleted, did not show an increase in postnatal body weight and had a minor phenotype compared to the previous model used by Sun et al.: glucose tolerance was slightly impaired, insulin release was diminished by only 50% and β-cell mass was unchanged [247]. Interestingly, RNA-seq data showed that deletion of AMPK in β-cells leads to an upregulation of neuronal and hepatic genes also defined as "disallowed genes".

On the contrary, mice with a β-cell-specific ablation of LKB1 generated by Fu et al. and others [247-249] have shown discordant phenotype from the previously described AMPK model. According to Fu et al. [248], LKB1-KO mice displayed improved glucose tolerance and protection against diet-induced hyperglycaemia whereas LKB1-KO β-cells were hypertrophic because of elevated mTOR

activity. The authors concluded that inhibition of LKB1 activity may facilitate β-cell expansion and glucose tolerance in vivo. Kone et al.[247] also reported that loss of LKB1 led to improved glucose homeostasis and increased β-cell mass and insulin secretion in vivo but GSIS in vitro was reported as largely unchanged. Moreover, GSEA performed on RNA-seq data from LKB1KO vs control mice displayed no significant changes in β-cells marker genes such as Slc2a2 (Glut2), Pdx1, MafA and NeuroD. However, as for AMPK-KO, the analysis identified significant enrichment for a subset of neuronal and hepatic genes. This data, together with those obtained from AMPK-KO, suggest that the LKB1-AMPK axis might maintain β-cell identity by suppressing alternative pathways.

		Cells	Drug/Model	Effect on IS/GSIS	Ref.
In vitro	Activation	INS-1	AICAR	↑ basal IS/ ↓ GSIS	[250]
		MIN6	AICAR	↓ GSIS	[251]
		MIN6	Metformin	↓ GSIS	[194]
		MIN6	Troglitazone	↓ GSIS	[239]
		MIN6	Berberine	↓ GSIS	[241]
		INS-1	Pioglitazone	↓ GSIS	[240]
		Rat islets	AICAR	↑ GSIS	[250]
		Rat islets	AICAR	↑ GSIS	[238]
		Perfused pancreas	AICAR	↑ GSIS	[238]
		Mouse islets	AICAR	↑ GSIS	[160]
		Mouse islets	AICAR	↑ GSIS	[237]
		Human islets	Metformin	↓ GSIS	[194]
		Rat islets	Troglitazone	↓ GSIS	[239]
		Rat islets	Berberine	↓ GSIS	[241]
	Inhibition	Mouse Islets	Whole body AMPKα2 KO	normal IS	[252]
		Mouse islets	β-cell AMPKa1/AMPKa2 KO	↑ basal IS and GSIS	[192]
		Mouse Islets	β-cell AMPKa1/AMPKa2 KO	↑ basal IS/↓ GSIS	[245]
		Mouse Islets	β-cell AMPKa1/AMPKa2 KO	↑ IS	[247]
		Mouse Islets	β-cell LKB1 KO	normal GSIS	[247]
		Mouse Islets	β-cell LKB1 KO	↑ IS	[248]
		Mouse Islets	β-cell LKB1 KO	↑ IS	[249]
In vivo	Inhibition	Mouse	Whole body AMPKα1 KO	normal IS	[253]
		Mouse	Whole body AMPKα2 KO	↓ IS	[253]
		Mouse	Whole body AMPKα2 KO	↓ IS	[252]
		Mouse	β-cell and hypothalamus AMPKa1/AMPKa2 KO	↓ IS	[192]
		Mouse	β-cell and hypothalamus AMPKa1/AMPKa2 KO	↓ IS	[245]
		Mouse	β-cell AMPKa1/AMPKa2 KO	↓ IS	[247]
		Mouse	β-cell LKB1 KO	↑ IS	[247]
		Mouse	β-cell LKB1 KO	↑ IS	[248]
		Mouse	β-cell LKB1 KO	↑ IS	[249]

Table 1.1: In *vivo* and in *vitro* effects of AMPK on GSIS

1.6 MicroRNAs

MicroRNAs (miRNAs) are small non-coding RNAs ~22 nucleotides long that silence gene expression post-transcriptionally.

MiRNAs were identified for the first time in 1993 by Ambros' and Ruvkun's groups while studying the *Lin-14* gene in *C. elegans* [254, 255]. At the time *Lin-14* was already known to be down-regulated during development in C. Elegance and that *Lin-4* was a negative regulator of *lin-14* essential for the transition from the first larval stage (L1) to the second (L2).

Ambros' and Ruvkun's groups were able to simultaneously elucidate the molecular mechanism that underpinned the inhibition of *Lin-14* mediated by *Lin-4* [254, 255]. In fact, they demonstrated that *Lin-4,* which encodes for two short non-coding RNAs (approximately 22 and 61 nt), was able to regulate *Lin-14* at the posttranscriptional level and that the 3'UTR of *Lin-14* was necessary and sufficient for this regulation. More specifically, this regulation occurred through seven conserved elements in the 3'UTR of *Lin-14* that are complementary to a portion of *Lin-4* transcripts. Therefore, they proposed a model by which the seven sites in the 3'UTR of *Lin-14* bind the *Lin-4* RNAs and synergically downregulate Lin-14 expression during development through the generation of a temporal gradient.

However, the importance of the role of miRNAs became clear only 7 years later when Ruvkun's and Horvitz's laboratories [256] identified a new miRNA (Let-7) that was able to regulate the heterochronic pathway in *C. elegans* and promote the transition from the last larval stage (L4) into the adult stage. Importantly, in 2002 George Adrian Calin [257] and his team described for the first time an involvement of miRNA genes in human tumours. They associated the deletion, and therefore downregulation, of miR-15 and 16 genes in chronic lymphocytic leukaemia. Since then many studies on miRNAs have been carried out and their role and function widely investigated within different species: many miRNAs are evolutionarily conserved, and this strongly suggests an important role for these little molecules in relevant regulatory pathways [258]. At the time of writing, miRbase v22, the central online repository for miRNAs, reported that, in human alone, there are

2656 annotated mature miRNAs [259]. Remarkably, it has been estimated that 60% of protein-coding genes are under their control [260].

1.6.1 Biogenesis of miRNAs

In mammalian cells, miRNAs are mostly transcribed by RNA polymerase II from both intragenic and intergenic regions. The former are mainly processed from introns of protein coding genes, while the latter are regulated by their own promoters [261, 262]. In less frequent cases, human intragenic miRNAs can be independently transcribed by RNA polymerase III if they are close to Alu repeats [263].

It is possible for different miRNAs to be transcribed together as long transcriptional units or "clusters" that contain multiple miRNAs [264]. However, the expression of miRNAs that are co-transcribed from the same cluster can be individually regulated at the post-transcriptional level [265]. Interestingly, most of the clusters are evolutionary conserved, thus implying an important biological function [266].

A schematic illustration of the genomic organization and structure of miRNA genes can be found in Figure 1.6.
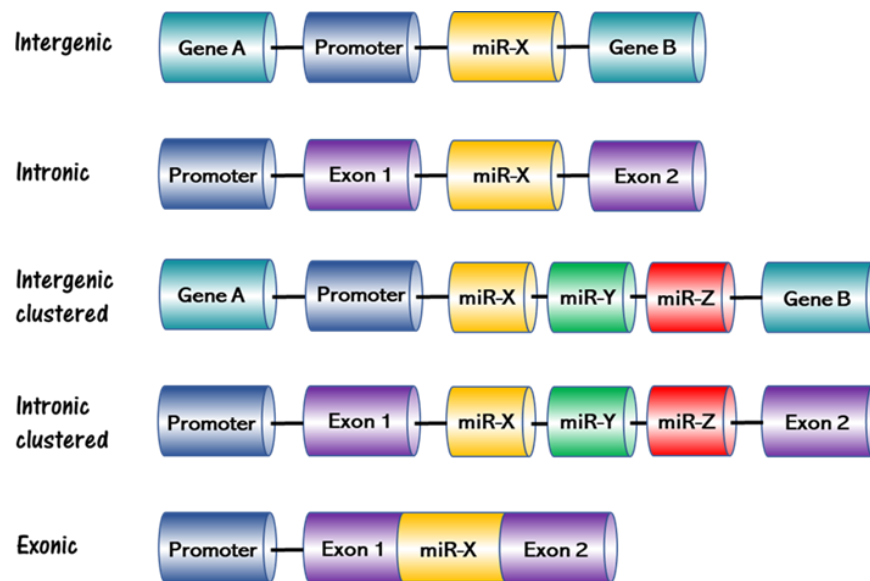
Figure 1.6: Schematic representation of genomic miRNA genes organization and structure

MiRNAs are typically transcribed as long (over 1kb) primary transcripts (pri-miRNA) that contain a double stranded stem of about 33-35 bp where the mature miRNA sequence is enclosed by a terminal loop and single-stranded RNA segments at both the 5' and 3' sides and characterized by a CAP and a poly-A tail [265]. Pri-miRNAs are quickly processed into the nucleus by the Microprocessor complex that consists of an RNAse III enzyme, DROSHA, and the DiGeorge syndrome critical region 8 (DGCR8), also called PACHA (Partner of DROSHA). The Microprocessor recognizes the pri-miRNA through DGCR8, and cleaves it (through DROSHA) at the base of the hairpin structure [267]. This results in the formation of a hairpin-shaped precursor RNA of ~70nt with a 2nt and 3' overhang called pre-miRNA (Figure 1.7) [268].
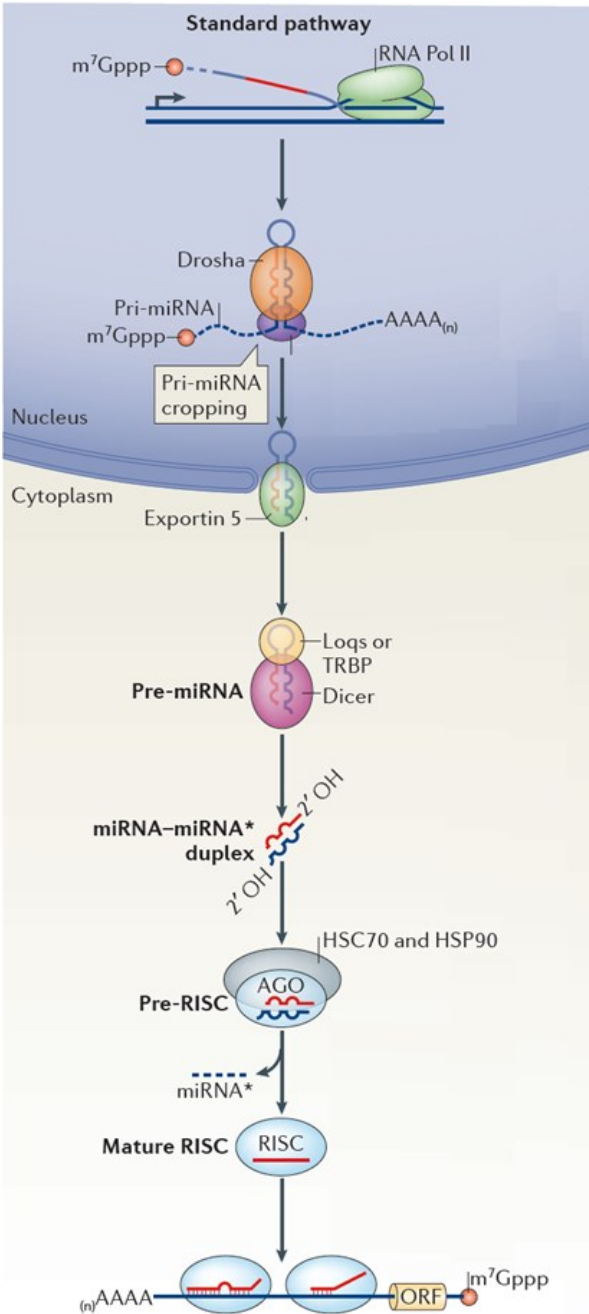
Figure 1.7: miRNA biogenesis.
MiRNAs are typically transcribed mainly by RNA polymerase II (Pol II) as long primary transcripts (pri-miRNA) consisting of multiple parts: a double strand, a terminal loop and single-stranded RNA segments at both the 5' and 3' sides that are characterized by a CAP and a poly-A tail. Pri-miRNA transcripts are processed by Drosha in the nucleus resulting in the precursor miRNA (pre-miRNA). The pre-miRNA is exported from the nucleus by exportin 5 and then further cleaved by the endonuclease Dicer together with its dsRNA-binding partner TRBP (transactivation-response RNA-binding protein; in mammals) form a miRNA–miRNA* duplex. Supported by the HSC70–HSP90 chaperone machinery, this duplex is loaded into an Argonaute (AGO) protein as a dsRNA. Subsequent maturation steps expel the miRNA*, producing a mature RNA-induced silencing complex (RISC). The image was adapted from Ameres and Zamore et al. [269].

Studies demonstrate that the depletion of both DGCR8 and DROSHA results in a reduction of both precursors of mature miRNAs in multiple organisms, including humans [268, 270-273]. Interestingly, the deletion of the genomic region that includes DGCR8 is implicated in a genetic disorder called DiGeorge syndrome [274], a condition present from birth that can cause a range of lifelong problems.

The efficiency of Drosha-mediated processing of pri-miRNAs is important to determine miRNA abundance and is regulated by the amount of Microprocessor available to the cell. Remarkably, Drosha and DGCR8 autoregulate each other post-transcriptionally [275]. In fact, while DGCR8 is able to stabilize DROSHA through protein-protein interactions [276], DROSHA destabilizes DGCR8 mRNA by cleaving it at a hairpin in the second exon [276, 277].

Following microprocessor processing, the pre-miRNA is ready to be exported to the cytoplasm. This is mediated by Exportin 5 (XPO5) that, helped by cofactor GTP-binding nuclear protein Ran-GTP, recognises and binds the pre-miRNA with high affinity [278, 279] and promotes its passage from the nucleus to the cytoplasm. Additionally, Exportin 5 plays also a role in pre-miRNAs stabilization by preventing their degradation by exonucleases [280]. In fact, knockdown of XPO5 results in the reduction of miRNA levels without an accumulation of nuclear pre-miRNA [281].

Once into the cytoplasm, the pre-miRNA is processed by DICER, an RNAse III endonuclease, which cleaves the ~70 nt double stranded RNA at the loop level leaving a small RNA duplex of 21–24 nucleotides [282, 283] characterised by a phosphate at 5' and a 2nt overhang with hydroxyl at the 3' end.

The regulation of this step is affected by different factors [284]. In human, for example, miRNAs themselves can regulate DICER. In fact, *DICER1* mRNA contains binding sites for miRNA Let-7 resulting in a loop where both DICER and Let-7 regulate each other [285]. Moreover, DICER activity can be regulated by many protein interactions: for example, when the trans-activation response RNA-binding protein (TRBP) is bound to DICER, it enhances the stability of Dicer–substrate complexes and therefore the cleavage of miRNA precursors [286]. Moreover, it has been demonstrated that some proteins like LIN28 act selectively on specific miRNAs. In mammalian cells,

LIN28A blocks DICER processing of Let-7 and simultaneously recruits Terminal uridylyltransferase 4 (TUT4) that adds uridines to the 3' ends of the RNA and promotes its degradation [287].

After DICER produces the 21-24nt miRNA duplex, it is ready to be loaded onto Argonaute proteins and form the miRNA inducing silencing complex (miRISC). TRBP is able to sense the strand with the less stable 5' end (guide strand) in the duplex and loads the double stranded miRNA in the right orientation in order to allow the Argonautes to unwind the duplex, remove the "passenger" strand and leave the mature miRNA molecule [278, 288]. The sequence of the mature miRNAs drives the miRISC to the target mRNAs by partial complementarity.

1.6.2 Role of AGO proteins

Argonaute proteins are evolutionarily conserved and can be divided into two categories, Ago and Piwi [289]. Ago proteins have a role in short interfering RNAs (siRNA) and miRNA pathways, while Piwi proteins interact with PIWI-interacting RNAs (piRNAs) to regulate gene expression in the germ-line [290]. In mice and humans, Argonautes are encoded by 8 different genes [291] and, more specifically in humans, AGO1, AGO2, AGO3 and AGO4 form the AGO subfamily [292]. *AGO1, 3* and *4* are clustered on chromosome 1, whereas the *AGO2* gene is located on chromosome 8. The human PIWI subfamily consists of HIWI1, HIWI2, HIWI3 and HILI proteins, all encoded from different chromosomes[289] (Figure 1.8).
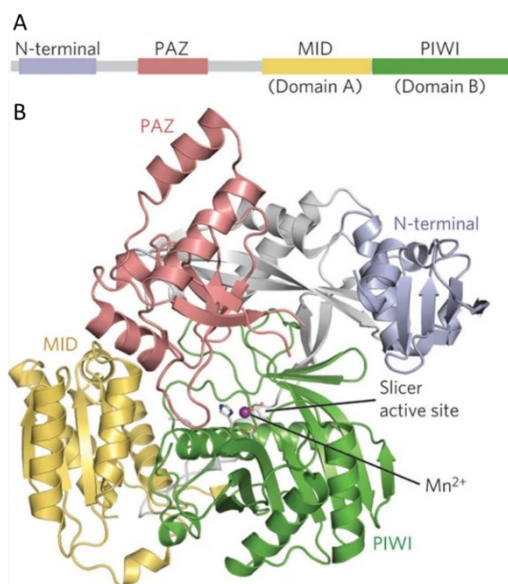
Figure 1.8: Domains and structure of Argonaute proteins
A. Schematic illustration of Argonaute main protein domains: the N-terminal, PAZ, MID and PIWI domains.
B. Ribbon representation of eukaryotic Argonaute crystal structure. The image was adapted from Jinek et al. [293]

AGO2 is the most studied argonaute [294] and is the only AGO to function as an endonuclease in mammals [295, 296]. While the role of the remaining ones has not been fully identified yet, it has been suggested that the action of each AGO is functionally equivalent when operating with bulged miRNA duplexes [297]. In humans, the four AGO proteins (AGO1–4) have also been associated with almost indistinguishable sets of miRNAs [297-299].

Like most proteins, Argonaute proteins are regulated post-translationally through various types of modifications that can affect their function [292]. For example, human AGO proteins under cellular stress conditions can be modified by poly (ADP)-ribose, a major nuclear regulatory macromolecule, that in the cytoplasm can mitigate miRNA-guided repression [300]. Argonaute can also undergo ubiquitination or phosphorylation: like for many other proteins the former is responsible for the Argonaute degradation, while the latter can dictate its intracellular localization [301] or regulate the efficiency of small RNA binding [302].

AGO proteins have been localized both in the nucleus and the cytoplasm suggesting that they could have a role also in nuclear processes. In fact, it is well established that, in plants, AGO4 can be associated with nascent transcripts and induce DNA methylation directed by double-stranded RNA through short interfering RNA (siRNA) [303, 304]. RNAi has also been linked to transcriptional silencing in mammals, but evidence supporting a link with AGO proteins is lacking [292]. Interestingly, human AGO1 and AGO2 have been shown to mediate transcriptional silencing by interacting with the transcription machinery [305, 306] and to affect alternative splicing facilitating spliceosome recruitment and modulating RNA polymerase II elongation rate [307].

1.6.3 Mechanism of action of miRNAs

As mentioned earlier, the mechanisms by which miRISC recognizes complementarity sequences on the mRNA targets is by Watson–Crick base pairing (Figure 1.9). In plants, miRNAs recognize nearly or fully complementary binding sites promoting target endonucleolytic cleavage in the middle of the miRNA–mRNA duplex (between nucleotide 10 and 11 opposite to the guide strand) and causing efficient degradation of the RNA [308]. In animals, the canonical miRNA-target recognition consists in a 7-8nt seed-matched sites interaction that occurs between the miRNA seed region

(nucleotides 2–8 of the miRNA) and the 3' UTRs of the target mRNAs. However, binding sites in the 5'UTRs and coding regions of target mRNAs have also been reported, with the latter occurring much more often than the former [309, 310]. Using both experimental and bioinformatics analyses [311-314] scientists have identified a set of constraints that are essential for miRNA-mRNA imperfect binding. Firstly, the seed region must bind to the target with perfect base paring to have an efficient repression. Furthermore, it has been demonstrated that an "A" residue opposite the nucleotide 1 of the miRNA and an "A or U" across nucleotide 9 can improve the efficiency even though they don't bind the miRNA sequence [311-314].

Secondly, in order to prevent AGO-mediated endonucleolytic cleavage of the target, a bulge or mismatch has to be present in the central region of the miRNA-mRNA duplex [315]. However, although mismatches and bulges are generally tolerated in this region, a good base pairing between residues 13–16 of the miRNA appears to be important when the matching in the seed region is not optimal [311, 313].
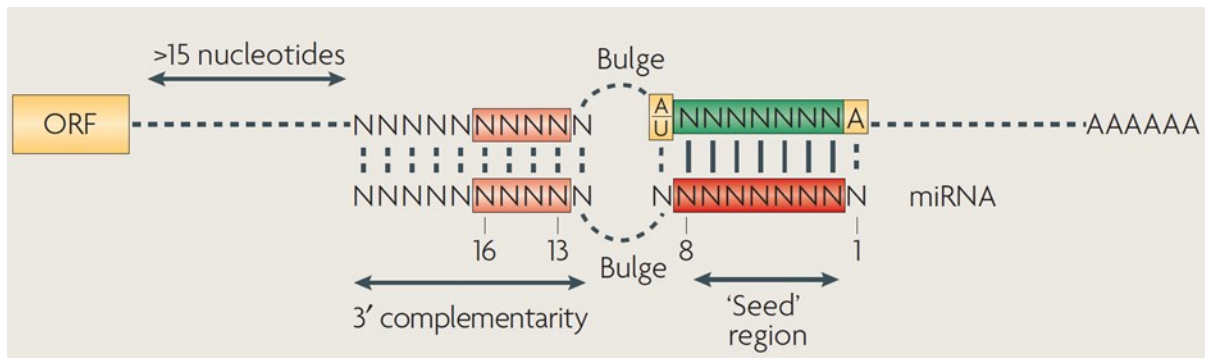
Figure 1.9: Schematic representation of miRNA-mRNA target interaction.
The miRNA-target canonical recognition in animals consists in 7-8nt seed- matched sites interaction that occur between the miRNA seed region (nucleotides 2–8 of the miRNA), and the 3' UTRs of target mRNAs. The seed region must bind to the target with perfect base paring in or-der to have an efficient repression. An "A" residue opposite the nucleotide 1 of the miRNA and an "A or U" across nucleotide 9, can improve the efficiency of the binding between miRNA and mRNA target. Figure from Filipowicz et al. [316]

In animals, miRNA-mediated repression of gene expression can occur through different mechanisms. Indeed, miRISC can recruit the CAF1-CCR4–NOT deadenylase complex to the target mRNAs [317], which results in their deadenylation and decapping. Destabilized, decapped mRNAs are then degraded by exonucleases [318]. Additionally, in the past few years it has been suggested that miRNAs in mammals can repress gene expression through different mechanisms (Figure 1.10) and inhibition of translation initiation is one of them [319, 320]. More specifically, Kiriakidou's study [321] showed that Argonaute proteins exhibit sequence similarities to the cytoplasmic cap-binding protein eIF4E (eukaryotic translation initiation factor 4E), which is essential for cap-dependent translation initiation, thus indicating that Argonaute can compete with eIF4E to bind to the cap structure [321]. AGO2 has also been shown to recruit eIF6 (Eukaryotic translation initiation factor 6), a protein that, amongst other functions, prevents the large ribosomal subunit from prematurely joining with the small ribosomal subunit, suggesting that its recruitment by AGO2 could have a role in repressing translation [322]. Another proposed mechanism of translational repression is the inhibition of translation elongation [323]. This is supported by the fact that in some studies, although protein expression was inhibited after miRNA action, clusters of ribosomes (polysomes) were detected bound to mRNAs suggesting that the repression occurred after translation had been initiated [324, 325]. Petersen et al. further demonstrated that inhibition of translation elongation might also occur through ribosome dissociation or ribosome drop-off [325].

Finally, it has also been recently suggested that, along with translation repression, co-translational protein degradation may occur. In this model, after premature translation termination, the incomplete aberrant protein might be degraded by protease activities [326]. At present, although there is some evidence suggesting this mechanism [323, 326], we lack any conclusive proof.
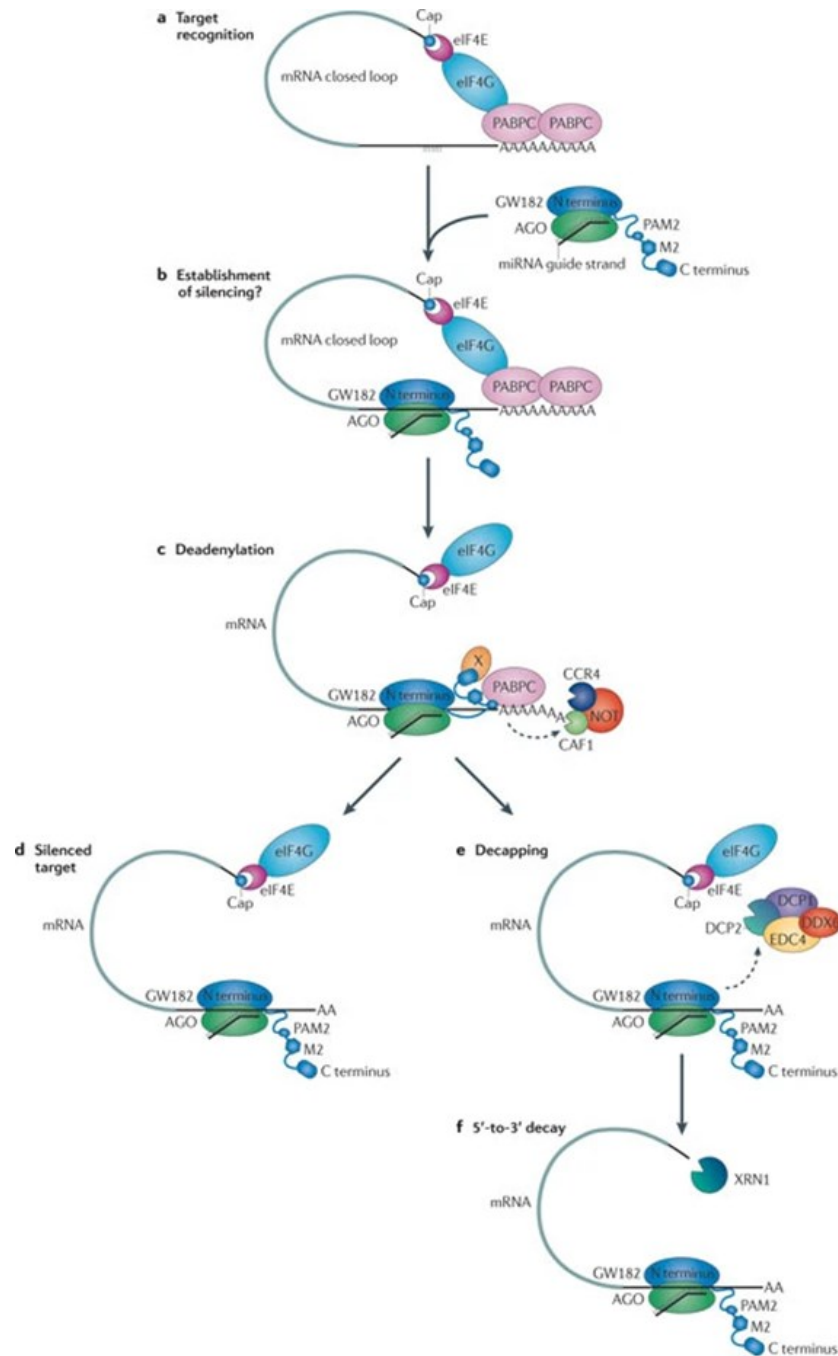
Figure 1.10: Mechanisms of miRNA gene repression in mammals

After miRNA-target recognition **(a)**, which in animals occurs through base-pairing to partially complementary binding sites, the AGO–GW182 complex directs the mRNA to deadenylation **(c)**. **(b)** represents the possibility of the translation being inhibited before deadenylation, but this step is still unclear. Depending on the cell type and/or specific target, deadenylated mRNAs can be stored in a translationally repressed state **(d)** or can be de-capped **(e)** and rapidly degraded by the major 5′-to-3′ exonuclease **(f)**. Figure from Huntzinger et al. [327]

1.6.4 MiRNAs biological function

MiRNAs build a complex regulatory network that finely regulates gene expression [328] with tissue specificity [329]. While each single miRNA can simultaneously target hundreds of genes, they usually have a moderate effect in the expression of any given one [330]. However, a specific miRNA can often target multiple genes on the same pathway with the aim to reinforce cell identity or respond to physiological conditions [330]. In order to orchestrate this complex regulatory network, miRNAs often take part together with transcription factors in several circuit motifs and build a so-called "coordinated repression" where both work together to repress a target. This type of repression has been found to be prevalent in mammalian genomes [331]. On the other side, it is also possible for miRNAs and transcription factors to carry out opposing functions and work as incoherent feedforward [328, 332], a process that is mostly used to reduce noise in gene expression [332].

It has been demonstrated that miRNAs regulate a broad range of biological processes, including proliferation, differentiation, and apoptosis [328, 333]. One of the first functions attributed to miRNAs was the ability to suppress the residual transcripts remaining from previous stages of cell development [334]. By doing this, miRNAs reinforce the cell-specific transcriptional gene expression program during development and this could at least partially explain why interactions between miRNAs and their mRNA targets are often characterized by tissue specificity [329, 335]. The importance of miRNAs in early development is supported by the lethal effect observed at a very early development phase upon *Ago2* and *Dicer-1* deletion in mouse [295, 336]. Notably, day 9.5 embryos of AGO2-deficient mice show a predominant defect in neural tube closure and cardiac failure with enlarged hearts that Liu et al. have mainly associated to yolk sac and placental defects [295]. Similarly, Bernstein et al. found numerous empty and necrotic decidua in embryos collected from *Dicer-1* mutant mice at embryonic day (E) 11.5 [336].

MiRNAs have also been studied in the context of cell cycle regulation. In fact, the cell entrance into the cell cycle and the transition between different phases is tightly regulated at multiple levels and miRNAs participate in many of these crucial pathways [337, 338]. In fact, it has been demon-

strated that they can both positively and negatively regulate cell cycle, mostly acting on G1–S transition [339], as first demonstrated by Hatfield et al. in *Drosophila*. In their study, the authors showed that the deletion of *Dicer-1* in germline stem cells resulted in cells blocked at the G1/S transition [340].

MiRNAs also play a role in the regulation of cell survival. Both pro-apoptotic and anti-apoptotic genes can be regulated by miRNAs and it is not to exclude that a single miRNA might have both pro-apoptotic and anti-apoptotic properties [341]. Not surprisingly, the role of miRNAs in apoptosis has been predominantly studied in the context of cancer [339]. Some miRNAs are defined pro-apoptotic because they can suppress anti-apoptotic factors [342]. These include miR-24, miR-195 and miR-365, which in human breast cancer cells MCF-7 have been shown to negatively regulate B-cell lymphoma 2 (*Bcl2*), an important anti-apoptotic protein. Singh et al. [343] also demonstrated that overexpression of these miRNAs causes a decrease in the mitochondrial membrane potential, increase in the release of cytochrome c into the cytosol, and therefore induction of apoptosis. On the contrary, miRNAs that mainly target pro-apoptotic genes are considered anti-apoptotic [341]. Studies carried out in MCF-7 cells showed how miR-21 down-regulates tumour suppressor tropomyosin 1 (*Tpm1*) and acts as an oncogene by promoting tumour growth [344].

MiRNAs were first linked to cancer in 2002 by Calin et al. In this study, two specific miRNAs, miR-15 and miR-16, were found deleted or down-regulated in 68% of chronic lymphocytic leukaemia (CLL) patients [345]. Since then, many studies have explored the role of miRNAs in cancer. They can in fact act oncogenes, either by down-regulating tumour-suppressor genes and/or promoting cell growth [346]. Interestingly, it has been demonstrated that miRNAs are generally downregulated in tumour conditions compared to normal tissues[347] and that defects in the miRNA biogenesis machinery is responsible for miRNAs deregulation and promotion of tumour formation [348]. Downregulation of miRNA processing regulators DROSHA, DGCR8 and DICER-1 in mouse lung adenocarcinoma cells caused, as expected, reduced miRNAs expression levels that led to increased proliferation. Moreover, a higher percentage of miRNA processing–impaired cells were entering S phase and grew faster than controls [348].

Finally, miRNAs have been implicated in the development of different pathologies including liver and kidney disorders [349], neurological, neurodevelopmental [350, 351] and cardiovascular diseases [352, 353], inflammation [354], and diabetes [355].

1.6.5 MiRNAs in pancreas, β-cell development and β-cell identity

Pancreas development is a complex process that involves a multistep stringent regulation of the expression of transcription factors, hormones and cell specific markers [356].

Pancreatic progenitor cells are characterized by the presence of several transcription regulators such as PDX1, NGN3, NEUROD1, NKX2.2, and PAX6 [357-359]. These have different expression patterns during the different stages of pancreas development, with some being expressed during very specific developmental timeframes and others, for example PDX1, being present from early development to adult β-cells [356].

Growing evidence supports the importance of miRNAs in processes such as pancreas development and β-cell differentiation [360, 361]. Lynn and colleagues demonstrated that deletion of *Dicer1* in mice during early pancreas development (using *Pdx1-Cre*, e8.5) has an impact on all pancreatic lineages but, more importantly, they observed a reduction of β-cells numbers that was associated with a decrease of neurogenin3, the initiator of endocrine cells differentiation. This study demonstrated the importance of miRNAs in β-cells formation [360].

Contrarily, β-cell specific *Dicer1* deletion from mid gestation (E9–11.5) using Cre-lox System showed that, although the mice had normal foetal and neonatal β-cell development, they displayed a progressive hyperglycaemia preceded by reduced insulin gene transcription and secretion as a consequence of reduced β-cell mass and reduced insulin granules [362].

Interestingly, when *Dicer1* was deleted in endocrine precursor cells, the animals showed loss of insulin and glucagon expression that was not associated with altered expression of characteristic β-cells transcription factors like PDX1, NKX6.1 and MAFA, but rather with an upregulation in the expression of neuronal genes [363].

Besides, in adult β-cells, the disruption of DICER caused a strong decrease in β-cell mass and impaired glucose-stimulated insulin secretion. In Melkman-Zehavi et al., the deletion of *Dicer* in adult mice using an inducible RIP-CreER system caused a dramatic decrease in insulin content and insulin mRNA level after only 3 weeks and a striking diabetic phenotype [364]. Martinez-Sanchez et al found that some genes, usually poorly expressed or "disallowed" in normal β-cells, were upregulated when DICER was inactivated. Further analysis demonstrated that some of those genes were miRNAs direct targets [365].

Thus, taken together these studies firmly indicate that miRNAs are very important for early pancreas and β-cell development, as well as the function and identity of the mature β-cell [366].

The role in cell development and function of a few specific, highly expressed miRNAs in pancreatic islets, such as miR-7 and miR-375, has been explored in detail in the past few years, both in *vitro* and in *vivo* [361, 367-369]. MiR-375 is a highly conserved and enriched miRNA in the adult endocrine pancreas and its increased expression has been associated with the generation of pancreatic islet cells while the maturation and function of β-cells is associated with its decrease [361, 367]. In fact, miR-375 represses genes associated with cellular growth and proliferations, including *Pdx1* and *NeuroD1* [370, 371]. It has also been suggested that NEUROD1 and PDX1 regulate miR-375 transcription themselves to form a β-cell specific regulatory network [370].

Kloosterman et al. [371] also demonstrated that specific knockdown of miR-375 in the first 4 days of development in zebrafish altered pancreatic islets morphology with a specific malformation of the endocrine pancreas. Interestingly, in these embryos the insulin-positive cells showed an unusual scattered pattern. In mouse, a genetic deletion of miR-375 has a similar effect. Poy et al. [368]demonstrated that 375KO mice had a reduced proliferative capacity of endocrine pancreas, exhibited a reduction in β-cell mass and an increase in alpha cell numbers.

MiR-7, like miR-375, is a highly conserved miRNA and is the most abundant in rat and human islets [372]. It is expressed at high level in both human endocrine cells during development and adult islets of Langerhans [369, 373]. Importantly, at stage E17.5 of development in mouse, an increased apoptosis throughout the pancreas followed by a strong reduction in insulin content followed the

administration of a miR-7 antisense *in vivo*. These animals also showed lower insulin content, impaired glucose intolerance and a decreased number of β-cells [374].

1.6.6 MiRNAs, insulin production and secretion

Insulin production and secretion are essential for the maintenance of blood glucose homeostasis and failure of β-cells to secrete enough insulin results in T2D [375]. Insulin transcription and the stability of its transcript are processes mainly regulated by glucose [56, 376], which acts through a network of transcriptional activators and repressors [54] of which miRNAs are part of. As described before, it has already been demonstrated that the specific deletion of *Dicer1* results in reduced expression of the insulin gene [360, 362, 364, 365] and, in the past few years, several studies have described many specific miRNAs able to regulate this important process acting at different levels [375]. For example, Melkman-Zehavi et al. [364] found that the downregulation of several specific miRNAs such as miR-24, miR-26, miR-148a and miR-182 following *Dicer1* KO in mouse islets was directly responsible for alteration observed in insulin expression. These miRNAs were described as positive regulators of insulin transcription via directly targeting the transcriptional repressors *Sox6* (miR-24, miR-26 and miR-148a) and *Bhlhe22* (miR-182) [364]. On the other hand, in the same study miR-7 was described as a negative regulator of insulin transcription [364]. Although Melkman-Zehavi and his group didn't investigate the mechanism, others demonstrated that miR-7 can repress insulin gene transcription acting on *Pax6* mRNA [377, 378].

Abdelfattah El Ouaamari et al [379] also provided evidence for a role of miR-375 in the glucose regulation of insulin gene expression. They demonstrated that, in rat, miR-375 targets *Pdk1*, a protein part of the PI3-kinase/PDK1/PKB signalling, a pathway that had already been implicated in glucose-induced upregulation of insulin gene expression. In 2004, miR-375 was the first miRNA to be associated with the regulation of glucose stimulated insulin secretion [380]. In their work, Poy et al.[368] suggested that miR-375 acts as a negative regulator of insulin secretion by targeting myotrophin and interfering with the exocytosis of the hormone. Following that study, many other miRNAs have been described as regulators of the exocytotic process of insulin secretion [381-383]. For example, in rat miR-124a has been suggested to regulate insulin secretion through the suppression of *Foxa2* (Forkhead box protein A2) and thus inhibiting K_{ATP} channel subunits Kir6.2 and

Sur-1 [384, 385]. This role was confirmed by Regazzi's group [383] in mouse MIN6 cells, where miR-124a and miR-96 are shown to act as key regulators of insulin secretion by targeting proteins such as RAB27A and NOC2, which are critical for insulin exocytosis [386, 387].

Moreover, miRNAs have also been found to target important proteins of the SNARE complex. The miR-29 family (miR-29a/b/c) has for example been shown to regulate β-cell insulin secretion by targeting Syntaxin-1 in rat β-cells [388], whereas in the same cells miR-335 directly binds and downregulates *Stxbp1* mRNA (syntaxin-binding protein 1 or Munc18-1) [389], a protein involved in the formation of the SNARE complex [390] and the recruitment of insulin granules to the membrane [391].

miRNA	Function	Ref.
let-7	Negatively regulates GSIS	[377]
miR-7	Negatively regulates insulin transcription	[364]
miR-9	Negatively regulates insulin secretion by targeting *Onecut-2*	[381]
miR-19a	Negatively regulates insulin expression by targeting *NeuroD1*	[392]
miR-24	Positively regulates insulin transcription targeting *Sox6*	[364]
miR-26	Positively regulates insulin transcription targeting *Sox6*	[364]
miR-29	Negatively regulates GSIS by targeting *Mct1*	[393]
miR-34a	Negatively regulates insulin secretion by targeting *Vamp2*	[383]
miR-96	Negatively regulates insulin secretion by targeting *Noc2*	[383]
miR-124a	Negatively regulates insulin secretion by targeting *Foxa2*	[384, 385]
miR-148a	Positively regulates insulin transcription targeting *Sox6*	[364]
miR-182	Positively regulates insulin transcription by targeting *Bhlhe22*	[364]
miR-204	Negatively regulates insulin transcription by targeting *MafA*	[384]
miR-335	Negatively regulates GSIS by targeting *Stxbp1*	[389]
miR-375	Negatively regulates insulin transcription and secretion by targeting *Mtpn* and *Pdk1*	[368, 379, 380]
miR-410	Positively regulates GSIS	[394]

Table 1.2: miRNAs involved in insulin production and secretion.

1.6.7 MiRNAs in Diabetes

MiRNAs are important regulators of multiple fundamental processes for the functioning of the human body, including metabolic homeostasis [395]. As described before, miRNAs participate in the regulation of glucose homeostasis and insulin secretion and defects in the miRNA processing machinery result in the development of diabetes [395].

Many rodent models of diabetes have been used in the past few years to identify changes in miR-NAs expression that coincide with the onset of this disease and potentially contribute to its development [396]. Interestingly, genetic mouse models of diabetes such as Goto-Kakizaki rats, diabetic db/db and obese ob/ob mice have been used by independent research teams [397-401] to identify a group of commonly misregulated miRNAs including miR-34a, miR-132, miR-184, miR-199a-5p, miR-210, miR-212, miR-338-3p and miR-383. Some of these miRNAs, including miR-34a, miR-210 and miR-383, are involved in processes such as apoptosis of β-cells and/or glucose-induced insulin secretion inhibition [397, 398] or, as it is the case for miR-184 and miR-338-3p, their down-regulation was found to induce β-cell proliferation [400, 401]. Interestingly, miR-184 has also been identified as an important modulator of compensatory β-cell expansion during insulin resistance in obesity [397, 400, 402] and pregnancy [400].

However, these studies also identified dozens of non-overlapping miRNAs between different animal models. This, together with the fact that often studies carried on humans and rodents have also led to the identification of distinct profiles of microRNAs associated with diabetes, suggests that, perhaps, animal studies should be used more to investigate the functional effect of the changes in miRNAs expression rather than the profiling itself.

In humans, comparative high throughput sequencing analysis identified multiple miRNAs differentially expressed between T2D and nondiabetic donors' islets [403], some of them been previously implicated in β-cell function and diabetes, such as miR-7. Surprisingly, the most downregulated miRNAs in T2D islets, comprising miR-432, miR-136, miR-469-3p and miR-656, were transcribed from the DLK1-MEG3 locus on human chromosome 14q32. This interesting finding captured the attention of the authors because the decreased expression of the gene MEG3, which is part of a

maternally expressed cluster of ncRNAs and implicated in numerous diseases, had never been associated with T2D before. Further investigation indicated that this cluster of miRNAs is primarily transcribed in the insulin secreting β-cells and its repression correlates with hypermethylation of the MEG3 locus [403].

As mentioned in previous chapters, genome-wide association studies (GWAS) data have identified more than 80 T2D susceptible loci. Although many of these polymorphisms are located in the coding region of susceptible genes [404-406], some were found in the 3'UTR [407] meaning that there is a possibility that they could have an impact on the functionality of miRNAs binding sites and therefore create potential susceptibility to T2D. Interestingly, when Bunt et al. [408] screened human islets for miRNAs, they identified 40 miRNAs that were predominantly expressed in islets compared to other tissues, and remarkably, the potential targets of these β-cell-specific miRNAs, such as *NOTCH2* and *SLC30A8*, were among those previously identified as type 2 diabetes susceptibility genes in GWAS.

SNPs can also occur in the pre-miRNAs sequences, but they appear to be relatively rare. In fact, it has been reported that only 10% of human pre-miRNAs are subject to mutation and only <1% of the SNPs are in the seed region [409]. However, when SNPs occur in the pre-miRNA sequence, they can affect the secondary structure of the pre-miRNA and therefore the miRNA processing efficiency [410]. An example is miR-146a rs2910164 C>G: this SNP is located within the seed sequence of the pre-miR-146a [411] and it has been suggested that it is responsible for the reduced levels of mature miR146a and associated with Diabetic Nephropathy and Retinopathy [412].

MiRNAs have also been detected in blood and other body fluids and it has been suggested that the profiling of these miRNAs profiling could reflect the physio-pathological state of an individual and therefore be used as potential biomarkers for many diseases including diabetes [413]. Many studies reporting the dysregulation of miRNAs in the serum or plasma from diabetic patients have emerged in the past few years [414]. Zhang et al. [415] have for example identified miR-126 expression to be down regulated in susceptible and T2D groups compared to non-diabetic individuals. These findings have also been confirmed by others [416-418], suggesting that this miRNA could be used as a biomarker for the identification of T2D susceptible individuals [415].

In a global profile study focusing on T2D circulating miRNAs, several miRNAs showed dysregulated levels (up or down regulation) in blood samples of T2D patients [419]. Moreover, a meta-analysis confirmed 40 significantly dysregulated miRNAs in T2D patients, and highlighted that circulating miR-29a, miR-34a, miR-103, miR-107, miR-132, miR-142-3p, miR-144 and miR-375 levels may serve as potential biomarkers for T2D [420]. While blood level of miRNAs seems to be a promising strategy to prevent diabetes development, in order to use the miRNA signature as a tool to diagnose T2D more independent studies with larger number of patients are needed.

As mentioned before, high blood sugar has a negative impact on pancreatic β-cells and contributes to β-cell failure [165] but the mechanisms behind this regulation are not well understood. Interestingly, miRNAs seem to have a role in this process. miR-375-null Mice, for example, develop hyperglycaemia at only 4 weeks of age and this has been linked to the role that this miRNA has in the regulation of a cluster of genes involved in cellular growth, proliferation and apoptosis [368]. Likewise, miR-155 and miR-26a are also possibly required to maintain normal glucose levels. In fact, it was reported that, while the overexpression of this miRNAs in mice enhanced glycolysis and insulin secretion, their deficiency caused hyperglycaemia [421, 422].

Besides, miRNAs expression can change in response to fluctuation of glucose levels [389] indicating that hyperglycaemia itself has an active role in the regulation of miRNA expression in islets. In fact, it has been shown that, in a spontaneous rat model of type 2 diabetes, hyperglycaemia can alter the expression patter of critical miRNAs such as miR-222 and miR-27a, and that this alteration not only depends on the glucose concentration but it is also tissue-dependent [423]. Moreover, circulating miRNA profiles of patient with T1D confirmed that more than 50 miRNAs, amongst the 2,083 screened, correlated with HbA1c [424] with miR-125b-5p showing the strongest positive correlation.

Insulin production and secretion are essential for the maintenance of blood glucose homeostasis and failure of β-cells to secrete enough insulin results in T2D [375]. Insulin transcription and the stability of its transcript are processes mainly regulated by glucose [56, 376], which acts through a network of transcriptional activators and repressors [54] of which miRNAs are part of. As described before, it has already been demonstrated that the specific deletion of *Dicer1* results in

reduced expression of the insulin gene [360, 362, 364, 365] and, in the past few years, several studies have described many specific miRNAs able to regulate this important process acting at different levels [375]. For example, Melkman-Zehavi et al. [364] found that the downregulation of several specific miRNAs such as miR-24, miR-26, miR-148a and miR-182 following *Dicer1* KO in mouse islets was directly responsible for alteration observed in insulin expression. These miRNAs were described as positive regulators of insulin transcription via directly targeting the transcriptional repressors *Sox6* (miR-24, miR-26 and miR-148a) and *Bhlhe22* (miR-182) [364]. On the other hand, in the same study miR-7 was described as a negative regulator of insulin transcription [364]. Although Melkman-Zehavi and his group didn't investigate the mechanism, others demonstrated that miR-7 can repress insulin gene transcription acting on *Pax6* mRNA [377, 378].

Abdelfattah El Ouaamari et al [379] also provided evidence for a role of miR-375 in the glucose regulation of insulin gene expression. They demonstrated that, in rat, miR-375 targets *Pdk1*, a protein part of the PI3-kinase/PDK1/PKB signalling, a pathway that had already been implicated in glucose-induced upregulation of insulin gene expression. In 2004, miR-375 was the first miRNA to be associated with the regulation of glucose stimulated insulin secretion [380]. In their work, Poy et al.[368] suggested that miR-375 acts as a negative regulator of insulin secretion by targeting myotrophin and interfering with the exocytosis of the hormone. Following that study, many other miRNAs have been described as regulators of the exocytotic process of insulin secretion [381-383]. For example, in rat miR-124a has been suggested to regulate insulin secretion through the suppression of *Foxa2* (Forkhead box protein A2) and thus inhibiting K_{ATP} channel subunits Kir6.2 and Sur-1 [384, 385]. This role was confirmed by Regazzi's group [383] in mouse MIN6 cells, where miR-124a and miR-96 are shown to act as key regulators of insulin secretion by targeting proteins such as RAB27A and NOC2, which are critical for insulin exocytosis [386, 387].

Moreover, miRNAs have also been found to target important proteins of the SNARE complex. The miR-29 family (miR-29a/b/c) has for example been shown to regulate β-cell insulin secretion by targeting Syntaxin-1 in rat β-cells [388], whereas in the same cells miR-335 directly binds and downregulates *Stxbp1* mRNA (syntaxin-binding protein 1 or Munc18-1) [389], a protein involved

in the formation of the SNARE complex [390] and the recruitment of insulin granules to the membrane [391].

1.6.8 MicroRNAs as therapeutic targets

Nowadays, it is possible to deliver oligonucleotides that can mimic or inhibit the activity of specific miRNAs in *vivo*, and it has been demonstrated that they can have therapeutic effects in animal models and humans [425]. Development of miRNAs therapeutics is an area of intense interest for pharmaceutical companies that are trying to develop both anti-miRNAs and oligonucleotides able to mimic the miRNA for replacement therapies [220]. This is due to the fact that, given their molecular characteristics, miRNA-targeting oligonucleotides have advantages over traditional small-molecule drugs. Fundamentally, they can be easily modified in order to enhance their pharmacokinetic/pharmacodynamic profile [426] and, more importantly, can regulate multiple transcripts simultaneously allowing for a global control of gene networks that could provide an advantage to cure complex diseases such as Diabetes.

One of the main approaches used in this field includes the use of antisense oligonucleotides (ASOs) [426], which are molecules that can bind to the miRNA in the miRISCs with high complementarity and block their binding to the mRNA target. They are very sensitive to the RNAses present into the blood and therefore need to be chemically modified to be delivered into the cells. Modifications are usually also made to enhance their cellular affinity. In addition, since the cell membrane is negatively charged, oligonucleotides are unable to enter the cells without modifications or encapsulation. Huynh et al., for example, showed that intraperitoneal injection of 2' modified phosphorothioated miR-182 anti-sense oligonucleotides could reduce liver micro-metastases of melanoma cells in mouse [427]. However, the best example of therapeutic that targets and inhibits a specific miRNA is Miravirsen. This drug acts as an antisense inhibitor of miR-122 and has been developed to reduce the levels of hepatitis C viral RNA in patients affected by hepatitis C infection [221, 222]. Data coming from the Phase II trial suggested that patients with chronic HCV did not show serious side effects after the weekly subcutaneous injections and showed a dose-dependent reduction in HCV RNA levels [221].

As mentionend before, it is also possible to use oligonucleotides that mimic the miRNA and function as replacement therapies. An example of this approach comes from Trang et al., where they demonstrated that in vivo delivery of Let-7 mimics reduces lung tumour development in mouse [428]. Remlarsen (MRG-201), designed to mimic the activity of miR-29 and currently in Phase 2 clinical trial, could potentially be a first-in-class miRNA mimic used to limit the formation of fibrous scar tissue in certain diseases idiopathic pulmonary fibrosis treatment [429].

Several studies [430-433] conducted in murine diabetes models have demonstrated that it is possible to modulate the level of specific miRNAs using oligonucleotides in order to improve insulin sensitivity. Evidence indicates that restoring the expression of specific miRNAs can induce insulin secretion or action and reverse the disease. Belgardt et al. demonstrated that, while overexpression of miR-200/miR-141 family in mice was responsible for increased levels of glucose in the blood and the development of diabetes in the first weeks of life, the specific deletion of these miRNAs in β-cells had a protective role towards the development of diabetes [430]. Moreover, manipulation of different miRNAs had the same outcome in mice deficient for the miR-143-145 cluster. These mice, which showed impaired insulin activation and glucose homeostasis when subject to induced transgenic overexpression of miR-143, were reported to be protected from development of obesity-associated insulin resistance when miR-143-145 cluster was knockdown [434]. Trajkovski et al. used antisense oligonucleotides in liver and adipocytes to silence miR-103 and miR-107 and this led to an improved glucose homeostasis and enhanced insulin sensitivity [431]. Moreover, when oligonucleotides against let-7 where directed to liver and muscle, the obese mice used in the study showed improved insulin sensitivity and better control of glucose homeostasis. The authors concluded that anti miR-induced knockdown of let-7 could provide an approach to treating T2D [432]. Indeed, Tuskita et al. [433] have recently demonstrated that intravenously administered miR-106b/222 mimics promotes β-cell proliferation and improves hyperglycaemia in mice with insulin-deficient diabetes. This opens new possibilities to the discovery of novel therapeutic strategies for diabetes but first we certainly need to understand how to overcome the delivery limitations regarding cell specificity when using miRNAs as therapeutics [435].

Therapeutic molecule	Treatment	Target miRNA	Biotechnology or Biopharmaceutical Company	Stage
Miravirsen (SPC3649)	Treatment of hepatitis C virus infection	miR-122	Santaris Pharma	Phase-II
MRX34	Treatment of different types of cancers	miR-34a	miRNA Therapeutics	Phase-I
RG-101	Treatment of viral effect	miR-122	Regulus Therapeutics	Phase IB
RG-012	Prevent nephropathy	miRNA-21	Regulus Therapeutics	Preclinical
RGLS4326	Treatment of Polycystic kidney disease	miR-17	Regulus Therapeutics	Phase-I
MGN-1374	Treatment of post-myocardial infarction	miRNA-15 / miR-195	miRagen Therapeutics	Preclinical
MGN-2677	Treatment of vascular disease	miR-143 /miR-145	miRagen Therapeutics	Preclinical
MGN-4220	Treatment of cardiac fibrosis	miR-29	miRagen Therapeutics	Preclinical
MGN-4893	Treatment of abnormal red blood cell production	miR-451	miRagen Therapeutics.	Preclinical
MGN-5804	Treatment of cardiometabolic disease	miR-378	miRagen Therapeutics	Preclinical
MGN-6114	Treatment of peripheral arterial disease	miR-92	miRagen Therapeutics	Preclinical
MGN-9103	Treatment of chronic heart failure	miR-208	miRagen Therapeutics	Preclinical
Cobomarsen (MRG-106)	Treatment of cutaneous T-cell lymphoma (CTCL)	miR-155	miRagen Therapeutics	Phase-I
MRG-107	Treatment of amyotrophic lateral sclerosis (ALS)	miR-155	miRagen Therapeutics	Phase-I
MRG-110	Target blood vessel growth and control of ischemia	miR-92a	miRagen Therapeutics	Phase-I
Remlarsen (MRG-201)	Treatment of different type of fibrosis	miR-29	miRagen Therapeutics	Phase-I

Table 1.3: Major miRNA-based therapeutics which are in the development phase.

1.6.9 Approaches for miRNAs targets identification

As described before, miRNAs are important posttranscriptional regulators of a large genetic networks and, in order to understand their molecular function in specific cells or tissues, it is important to identify their targetome, which is defined as the group of mRNA molecules targeted by a specific miRNA [436]. Several algorithms able to predict the miRNA binding site to the mRNA

targets have been developed in the past few years [436] and most of them rely on the canonical rule of miRNAs binding to the 3' untranslated region (3'UTR) of their mRNA targets *via* the seed region [437]. Some tools, such as for example TargetScan and MiRanda, predict miRNAs targets by searching for the presence of conserved 6mer, 7mer and 8mer sites which match the seed of each miRNA while taking into account target site accessibility and thermodynamic properties of the seed region [438].

Site accessibility is the measurement of a miRNAs ability to find and hybridise to its mRNA target. mRNAs are normally characterized by a long and complex secondary structure that can obstruct miRNA binding, and therefore miRNAs normally first bind to an accessible short region on the mRNA sequence and induce a secondary structure unfold that allows their complete hybridization to the target. MiRanda uses the hybridization energy and amount of energy required to make an mRNA site accessible to the miRNA to predict the targets.

On the other side, the thermodynamic properties refer to the measure of changes in the free energy (ΔG) that characterize a specific miRNA-mRNA duplex. Free energy can be used as a measure of the stability of a biological system and since reactions with a negative ΔG have less energy available to react in the future, miRNA-mRNA duplex with low free energy is considered stable. Therefore, tools like MiRanda can help to identify candidate mRNA targets by predicting how strongly the miRNA can bind to it [439].

However, it has been shown that only 25% of miRNA-mRNA interactions occur through canonical binding [440]. In fact, functional miRNA-target interactions can also occur through G:U wobbles, bulges and with the gene coding sequence [441] and this, together with the substantially high false positive (46-63%) and negative (44-82%) rate given by prediction programmes [442, 443], suggest that computational approaches for miRNAs binding site prediction still represent a big challenge. For this reason, it is widely accepted that only a combination of experimental approaches and prediction algorithms can represent a winning strategy to study a specific miRNA regulatory network [444].

Profiling-based strategies are widespread experimental methods used to study the molecular function of the miRNA of interest in a specific tissue/cell line. These methods generally involve

overexpression or inhibition of the miRNA of interest, which can be achieved by transfecting a double-stranded RNA that mimics or inhibits the mature endogenous miRNA, coupled with a downstream gene-expression analysis, such as RNA microarray or sequencing, and/or a proteome profiling (Table 1.4). Although these methods are very sensitive and characterized by a high genome profile coverage, they are unable to experimentally discriminate between direct and indirect miRNA targets. This distinction is particularly important when the specific mechanism of action of the miRNAs needs to be understood to develop new therapeutic agents without off-target effects [445].

		SENSITIVITY/ RESOLUTION	COVERAGE	SAMPLE TYPE	DATA ANALYSIS	PERFORMANCE NOTES
PROTEIN ARRAYS	Reverse-phase protein arrays (RPPA)	●	Candidate (antibody specific)	Unrestricted	Signal intensity	+ parallel sample analysis (diagnostics) − cross-reactivity − signal/noise ratio
	Antibody arrays	●	Low coverage (antibody arrays)	Unrestricted	Signal intensity	− cross-reactivity − depends on quality of antibody array
RNA PROFILING	Microarray	●●	High coverage (RNA array)	Unrestricted	Standard bioinformatics	+ cost effective − background interference (signal/noise) − predefined gene complement
	RNAseq	●●●	Genome-wide	Unrestricted	Standard NGS bioinformatics	+ unbiased + broad dynamic range + deep coverage
PROTEOME PROFILING	2D-DIGE	●	Limited coverage (visual detection)	Unrestricted	Signal intensity mass spec.	− reduced proteome-wide resolution − qualitative assesment − influenced by protein characteristics
	iTRAQ	●●	Whole proteome	Unrestricted	Mass spec. bioinformatics	+ multiplex mass-spec. analysis − sample complexity − biased detection
	ICAT	●●	Whole proteome	Unrestricted	Mass spec. bioinformatics	− non-linear recovery from beads − requires Cys for detection
	SRM, RIP-chip-SRM	●●●	Selective coverage (candidate list)	Unrestricted	Mass spec. bioinformatics	+ high resolution + dynamic range − limited coverage (predefined peptide set)
	SILAC, pSILAC	●●	Whole proteome	Cell cultures *C. elegans*	Mass spec. bioinformatics	+ accurate parallel analysis + temporal resolution − *in vivo* delivery of labelled aminoacids
	Label-free mass spectrometry	●●	Whole proteome	Unrestricted	Mass spec. bioinformatics	+ high coverage + cost effective − lower quantitative precission
RIBOSOME MAPPING	Polysome profiling	●●	High coverage (RNA array readout)	Unrestricted	Standard bioinformatics	+ miRNA-mediated translational effects − microarray readout (signal/noise ratio)
	Ribosome profiling	●●●	Genome-wide (RNAseq readout)	Unrestricted	Standard bioinformatics pipeline (NGS)	+ miRNA-mediated translational effects + RNAseq readout + high resolution footprint

Table 1.4: Profiling-Based Strategies for miRNA target identification
From Steinkraus et al. [446]

New methods that rely on the immunoprecipitation of AGO proteins to directly capture miRISC-bound mRNAs have recently been developed to overcome previous limitations (Table 1.5). RNA

immunoprecipitation (RIP) is one of the first methods of this kind used to identify miRNAs direct targets. RIP is typically performed following overexpression/inhibition of the miRNA of interest using an antibody against members of the miRISC complex, often AGO2. Subsequently, immunoprecipitated RNAs are identified by qRT-PCR, microarray (RIP-Chip), or next-generation sequencing (RIP-seq). Overexpression or inhibition of the miRNA of interest results in an enrichment or reduction, respectively, in the amount of target mRNAs present in the immunoprecipitates. Although straightforward and useful, RIP-based approaches can be characterized by a significant false positive discovery rate due to the immunoprecipitation of RNAs that can unspecifically interact with proteins and are not eliminated because of the low-stringency purification protocols [446]. Conversely, loose interactions can be lost during the immunoprecipitation and important targets remain undetected [446].

		FIDELITY	ENDO. CONTEXT	X-LINK	COVERAGE	SAMPLE TYPE	DATA ANALYSIS	PERFORMANCE NOTES
RIP	RIP-chip/seq	●	✓	✗	High coverage/ genome-wide	Cells, tissue	Signal intensity/ standard NGS bioinformatics	+ straightforward protocol − unspecific interactions
CLIP APPROACHES	Ago-HITS-CLIP	●●●	✓	254 nm	Genome-wide	Cells, tissue, *C. elegans*	Extensive bioinformatics (ternary maps)	+ high *in vivo* specificity/coverage − high-affinity antibody − amount of input material
	PAR-CLIP	●●●	✓	365 nm	Genome-wide	Cells	Extensive bioinformatics, (nt. footprint)	+ high *in vivo* resolution − high-affinity antibody − nucleoside analogue delivery
	iCLIP	●●●	✓	254 nm	Genome-wide	Cells, tissue, *C. elegans*	Extensive bioinformatics, (nt. resolution)	+ high *in vivo* coverage − high-affinity antibody − amount of input material
	CLASH	●●	✓	254 nm	Partial coverage	Cells	Extensive bioinformatics (chimeras)	+ precise miRNA-target pairs − proximity ligations − low *in vivo* coverage
	iPAR-CLIP	●●●	✓	254 nm	Genome-wide	Cells, *C. elegans*	Extensive bioinformatics (chimera/nt. res.)	+ high *in vivo* resolution + precise miRNA-target pairs − nucleoside analogue delivery

Table 1.5: AGO2-Immunoprecipitation based methods for miRNA-mRNA interaction identification
From Steinkraus et al. [446]

For these reasons, stabilization of RNA-protein interaction through covalent binding have become increasingly adopted. Crosslinking and immunoprecipitation methods (CLIP) use UV to induce protein–RNA crosslinks preceding immunoprecipitation of miRISC [446]. Following stringent washes, the precipitated mRNA is released from the protein complexes using Proteinase K and, normally,

submitted to High-Throughput sequencing (HITS-CLIP)[447]. Ago-HITS-CLIP was the first CLIP approach used for miRNAs target identification [448] and several variations and refinements of the technique have been implemented during the past few years to improve efficiency and achieve single nucleotide resolution in the binding site of miRISC or the RBP (Rna binding protein) of interest. PAR-CLIP (Photoactivatable Ribonucleoside Enhanced Crosslinking and Immunoprecipitation), for example, uses 4-thiouridine (4SU) or 6-thioguanosine (6SG) photoactivatable nucleosides in the cell culture media that are then incorporated in the cellular RNA and enhance both crosslinking and immunoprecipitation efficiency [449-451]. Another example is iCLIP (individual-nucleotide resolution CLIP), which utilizes the truncation of the cDNA synthesis at the crosslinking site and a combination of adapters and enzymatic reactions that, following deep-sequencing, allows for the determination of the protein location and the miRNA binding to the target RNAs with single nucleotide resolution. At the time of writing, iCLIP for miRNA target identification has however only been successfully used so far in *C. elegans* with an antibody against the AGO2 homologous ALG-1 [452] and further investigation is required to successfully apply this methodology to the identification of miRNA direct targets in mammalian cells.

1.6.10 AMPK and miRNAs

Various miRNAs have been shown to reduce AMPK activity by targeting AMPK mRNA, both directly and indirectly [453-455]. Different studies have been focused on the regulation of AMPK α subunit, which translational suppression has already been linked to tumorigenesis and formation of metastasis [455, 456]. Zhao et al. found that miR-148b was downregulated in human pancreatic cancer tissues and its expression inversely correlated with AMPKα1 subunit expression. Moreover, in vitro experiments further showed that miR-148b directly targets AMPKα1 3'UTR and its overexpression in PANC-1 cells decreases both mRNA and protein levels of AMPK α1 [455]. Finally, Zhang et al. demonstrated that AMPKα1 is a miR-301a direct target in osteosarcoma cells and suggested that miR-301a contributes to chemoresistance in osteosarcoma by inhibiting AMPKα1 [457].

While AMPK regulation by miRNAs has been widely studied, albeit mostly in cancer, the role of AMPK in the regulation of miRNAs has not been exhaustively described, especially in the β-cell. Kone et al. has previously shown that, in pancreatic β-cells, loss of AMPK impairs the expression

of more than 4000 genes [247], and interestingly, our group demonstrated that AMPKdKO mouse islets also showed an altered expression of several miRNAs. Moreover, ~850 predicted targets of these miRNAs identified using TargetScan were also differentially expressed in AMPKdKO mouse islets and GO analysis revealed that they could be involved in processes important for β-cell function [458]. Specifically, genes up-regulated upon AMPKdKO that were predicted targets of down-regulated miRNAs were enriched in pathways such as vesicle fusion, calcium dependent exocytosis, and nervous system development, suggesting that the newly identified miRNAs could contribute to the altered insulin secretory phenotype observed in islets depleted of AMPK and/or be responsible for the up-regulation of genes involved in neuronal function characteristic of AMPKdKO islets.

1.6.10.1 MiR-184

MiR-184 is the most down-regulated in AMPKdKO islets [458]. This is a highly conserved microRNA from fly to human [459] and has been shown to be involved in various biological processes such as germline development in *Drosophila*, where loss of miR-184 leads to multiple severe defects during oogenesis and early embryogenesis up to complete loss of egg production [460].

MiR-184 is also important for the regulation of cell proliferation and migration in different types of cells [461-464]. In the epidermis, for example, miR-184 represses proliferation and activates the NOTCH pathway suggesting a role as tumour suppressor [465]. Nagosa et al. demonstrated that miR-184 controls the balance between epidermal cell proliferation and differentiation: by targeting *K15* and *Fih1*, miR-184 regulates the transition from proliferation to early differentiation, while mutation in miR-184 results in impaired homeostasis. In addition, miR-184 knockout mice develop epidermal hyperplasia, while overexpression of miR-184 in stem/progenitor cells induced epidermal hypoplasia [459].

Different groups have investigated the function of miR-184 in pancreatic β-cells in the past few years [389, 397, 400, 458, 466]. The expression of miR-184 in human pancreatic islets is considerably high in comparison with other tissues such as liver and skeletal muscle [467] and its expression

positively correlates with insulin mRNA expression in human islets [389]. It has also been suggested that miR-184 could target proteins important for optimal insulin secretion such as L-type and T-type voltage-gated Ca^{2+} channel (*CACNA1C* and *CACNA1H*, respectively) [389]. Indeed, in MIN6 cells, mir-184 appears to inhibit insulin secretion through the repression of *Slc25a22*, a glutamate transporter that plays a role in the control of GSIS [468].

Moreover, miR-184 expression is downregulated in prediabetic and diabetic mice [397] and it has been demonstrated to play an important role in the regulation of β-cell proliferation during compensatory expansion in pregnancy and obesity [397, 401]. Tattikota et al. [400] demonstrated that miR-184 targets *Ago2* in obese mouse and this regulation is essential for the compensatory response to balanced proliferation according to insulin sensitivity. However, in Kone's work *Ago2* expression was found unaltered upon AMPK β-cell specific deletion [247].

Both Nesca et al. and Tattikota et al. [397, 402] showed that miR-184 expression is regulated according to glucose metabolism both in *vivo* and in *vitro*. These researchers found that miR-184 is slightly down-regulated in mouse islets cultured at high glucose concentrations and is strongly increased in islets from mice fed a low-sugar/ketogenic diet. These data were confirmed by our group [458] and we also demonstrated that AMPK is required for the glucose-dependent regulation of miR-184 expression. In *vitro*, treatment of isolated mouse islets with increasing concentrations of glucose for 48h led to inactivation of AMPK and a slight but significant reduction of miR-184 expression. However, in AMPKdKO islets, miR-184 expression remained unchanged. In *vivo,* while mice fed a ketogenic diet showed an important increase in miR-184 expression, in both AMPKdKO and LKB1KO the miRNA remained unchanged, confirming that AMPK is required for glucose-mediated regulation miR-184 in pancreatic islets.

1.6.10.2 MiR-125b

MiR-125b, conversely, is one of the most up-regulated miRNAs in AMPKdKO islets. Interestingly, in the past few years high levels of circulating miR-125b have been associated with hyperglycaemia in prediabetic [469], T2D [469] and T1D subjects [424] suggesting that this miRNA may be a potential contributor to the detrimental effect of hyperglycaemia and the development of diabetes.

Indeed, according to our data, the expression of miR-125b positively correlates with BMI in human islets (Unpublished).

MiR-125b is ubiquitously expressed and its role in β-cells in still unknown. Preliminary data generated by our group indicated that overexpression of miR-125b in MIN6 cells impaired both insulin content and secretion. Insulin content of MIN6 cells transfected with miR-125b mimics was significantly reduced by 40% when compared with cells transfected with control mimic and GSIS was 30% lower. Overexpression of miR-125b has also an effect on MIN6 cell morphology, with cells appearing less rounded and more sprawl (Figure 1.11).
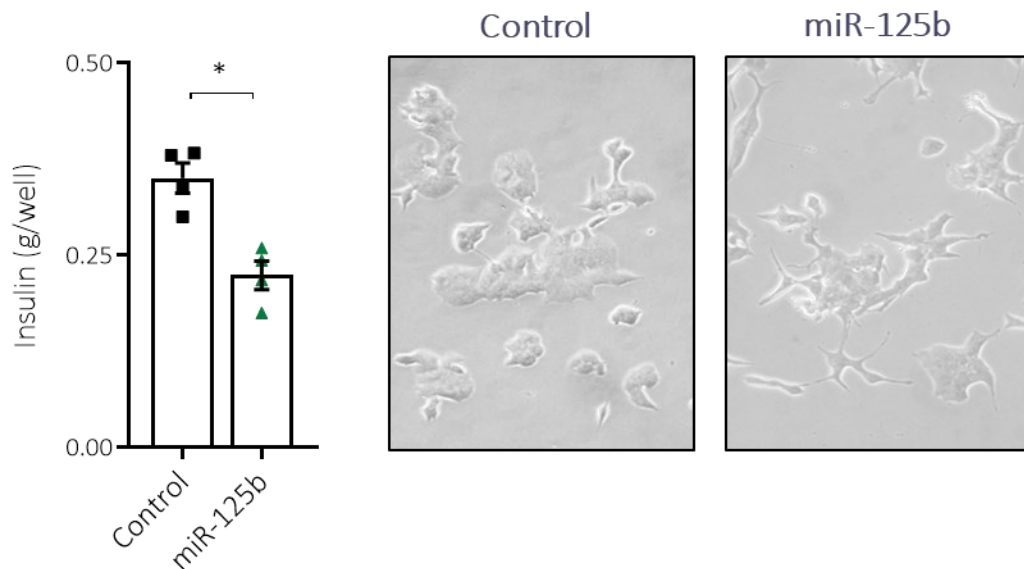
Figure 1.11: miR-125b overexpression reduces insulin content and affects cell morphology in MIN6 cells.
Transfection of miR-125b mimic in MIN6 resulted in a significant 40% reduction of insulin content (left) and affected morphology of the cells that appeared less rounded and more sprawling (Right).

In other cell types, miR-125b controls proliferation, apoptosis and differentiation [470]. MiR-125b expression has been found to be aberrant in a great variety of tumours where it can act as an oncogene or a tumour suppressor. This dual and opposing function depends on the role of the genes that this miRNA regulates [471].

Mir-125b elevated expression has been detected in several tumour types including acute myeloid leukaemia and B-cell acute lymphoid leukaemia [472, 473]. Bousquet et al. demonstrated that overexpression of miR-125b, using retroviral vectors, was the main cause of leukaemia development in mouse [473]. In acute myeloid leukaemia , where its overexpression has been identified as the main cause of tumour development [473] miR-125b directly targets core-binding factor subunit β (*Cbfβ*), a suggested tumour suppressor [474].

Furthermore, miR-125b has been shown to be upregulated also in pancreatic [475] gastric [476] and follicular cancers [477] but the molecular mechanisms that lead to this upregulation are not very well understood.

As mentioned earlier, miR-125b can also act as tumour suppressor by inhibiting cancer formation and proliferation [478-481]. Interestingly, in breast, ovary and cervical cancer miR-125b is subject to loss of function as both *MIR125B-1* and *MIR125B-2* loci are located in the so-called fragile sites that are commonly deleted in these types of tumours [482]. Indeed, in human ovarian cancer cells, miR-125b suppresses the growth by targeting *BCL3* [480], while in breast cancer the down regulation of miR-125b and the subsequently activation of ERBB 2 and 3 proteins, two confirmed targets of miR-125b, has been associated with cell proliferation and inhibition of programmed cell death [478, 479].

1.7 Aims of the thesis

This thesis aims to widen the understanding of the mechanisms underlying regulation and function of AMPK-regulated miRNAs in pancreatic β-cells.

Accordingly, the chapters of this thesis will aim to:

1. Determine whether AMPK controls CTCF activity to modulate miR-184 transcription and to identify *MIR184* TSS in mouse islets.

2. Determine whether SMAD2/3 promotes miR-125b expression in β-cells and if this regulation is inhibited by AMPK.

3. Identify miR-125b direct gene targets in β-cells in a high-throughput manner to elucidate its function and its potential as a target for the treatment of diabetes.

4. Develop a high-throughput method (AGO2 eiCLIP-seq) to identify all specific miRNA-mRNA interactions in human β-cells.

Chapter II: Materials and Methods

2.1 Mammalian cell culture

MIN6 cell line (mouse insulinoma) [483] was cultured in Dulbecco's modified Eagle's medium at 25mM D-glucose (DMEM; Sigma Aldrich), containing 15% [vol./vol.] foetal bovine serum (FBS), 100 U/mL streptomycin and 100 U/mL penicillin and supplemented with 50 μM 2-mercaptoethanol (Sigma Aldrich), at 37°C with 5% CO_2.

Human EndoCβH1 cell line [484] was cultured in DMEM 1g/L (Sigma) medium, supplemented with 2% [wt/vol] Albumin from Bovine Serum Fraction V Fatty Acids Free (Roche), 100 units/ml Penicillin-Streptomycin (Thermo Fisher Scientific), 100 μg/ml Streptomycin (Thermo Fisher Scientific), 10 mM Nicotinamide (Sigma), 5.5 μg/ml transferrin, 6.7 ng/ml sodium selenite and 50 μM 2-mercaptoethanol (Sigma Aldrich), at 37°C with 5% CO_2.

EndoCBH1 lacking miR-125b (EndoCBH1 miR125BKO) were generated using CRISPR-Cas9 by Rebecca Cheung in our lab. Briefly, EndoCBH1 cells were infected by a lentiviral vector expressing Cas9 under the control of a RIP promoter (a kind gift from Ferrer's Lab at Imperial) and two gRNAs against MIR125B2. These cells consist in a heterogenous population with a ~80% reduction in mature miR-125b levels.

2.2 MIN6B1 clone 14

Stable MIN6B1-SNAP-GLP-1R cells were provided by Dr Alejandra Tomas. They were generated by transfecting pSNAP-GLP-1R (Cisbio) into wild-type MIN6B1, G418 (1mg/ml) selection, and single-cell sorting by fluorescence-activated cell sorting (FACS) following SNAP-Surface-488 (New England Biolabs) labelling [485].

2.3 Mouse islet isolation and culture

Mice were euthanised by cervical dislocation of the neck followed by decapitation prior injection of collagenase NB8 from Clostridium histolyticum (1mg/mL in RPMI 1640) (Nordmark Biochemicals) into the bile duct until the pancreas was fully inflated. The pancreas was then removed and the pancreata were incubated for 12 min at 37°C in a water bath for digestion and islets separation from remaining pancreatic tissue. Digested pancreas were washed 3X with RPMI 1640 before separation of the exocrine tissue by centrifugation through a histopaque gradient at densities 1.119 and 1.087 g/mL and RPMI-1640 medium (Sigma Aldrich) at 2500 rpm for 20 min. Islets were picked from the interphase and incubated in RPMI-1640 medium supplemented with 10% heat-inactivated foetal bovine serum (FBS) (Sigma) and 100U/ml penicillin and streptomycin (Thermo Fisher Scientific) and incubated at 37°C with 5% CO_2 for 24 hours before use.

2.4 Human islets isolation and culture

Human islets were isolated from deceased heart-beating donors in the isolation centres in Oxford, Pisa or Milan, respecting the relevant national and local ethical permissions. Islets were cultured in RPMI-1640 medium supplemented with 5.5 mM D-glucose, 10% FBS, 1% penicillin/streptomycin and 0.25 µg/ml fungizone.

2.5 Electron microscopy

MIN6 cells transfected with control or miR-125b mimics were fixed with 2% PFA + 2% glutaraldehyde in 0.1 M cacodylate buffer for 30 minutes at room temperature, washed with phosphate buffer and prepared for conventional EM. Cells were then post-fixed in 3% potassium ferricyanide mixed with 2% osmium tetroxide, treated with tannic acid and then followed by dehydration with increasing concentrations of ethanol and propylene oxide. Finally, they were embedded in Epon resin and mounted onto Epon stubs by Dr. Alejandra Thomas. Epon was polymerised overnight at

60°C and coverslips removed by surface heating. Ultrathin 70-nm sections were cut with a diamond knife on a Leica ultramicrotome, stained with 2% uranyl acetate and visualised in a Tecnai T12 Spirit transmission electron microscope with a CCD camera.

2.6 RNA extraction

Total RNA from all samples was extracted using TRIzol reagent (Thermo Fisher Scientific) and chloroform (Sigma) in a 5:1 ratio, respectively, following manufacturer's instructions. After centrifugation the clear upper aqueous phase was transferred to a new Eppendorf tube and the RNA precipitated using 0.8 volume of isopropanol, washed twice with 70% ethanol, resuspended in RNAse free water and quantified using a Nanodrop spectrophotometer.

2.7 Reverse transcription (RT)

For mRNA, 200-300 ng of total RNA was reverse transcribed using a High Capacity cDNA RT kit (Thermo Fisher Scientifics) according to manufacturers' instructions. Briefly, 5 µL of 2X RT reaction (10X RT buffer, 4 mM dNTPs, 10X random primers and 1 µL of Multiscribe Reverse Transcriptase) were combined with an equal volume of RNA and incubated at 25 °C for 10 minutes, 37 °C for 120 minutes and 85 °C for 5 minutes.

For miRNAs, 20-30ng of total RNA was retrotranscribed using the miRCURY LNA RT Kit (Qiagen) according to manufacturers' instructions. Briefly, 3µL of RT reaction (5X RT Buffer, 0.5 µL of 10x miRCURY RT Enzyme Mix, 0.25µL of Synthetic RNA spike in) was combined with 2µL of RNA and incubate at 42 °C for 60 min, 95 °C for 5 min.

2.8 Quantitative Real-Time PCR (qRT-PCR)

To assess the level of gene expression qRT-PCR using Fast SYBR Green Master Mix (Thermo Fisher Scientific) was used. A final volume of 10 µL containing 6 µL SYBR green master mix and 0.35 µl of 10 µM primers was mixed to 2 µL of 1:4 dilution of cDNA.

MiRNAs expression was instead assessed using miRCURY LNA SYBR® Green PCR Kit (Qiagen). 6µL of 2X miRCURY SYBR Green Master Mix, 0.05µL of ROX Reference Dye and 1µL of primers mix were combined to 2µL of 1:10 dilution of cDNA. Primer sequences are shown in Table 2.1

In both cases, the amplification curves for each gene, derived from fluorescent signals, were generated using a 7500 HT qPCR machine (Thermo Fisher Scientific). A threshold was manually set to determine the number of cycles (Cts) required to reach the set level of fluorescence (threshold was set to cut the exponential part of the curve, above background signals). The relative expression was determined by normalising to either the housekeeping genes cyclophilin or β-Actin for mRNA and let-7 or miR-576 for miRNAs.

Name	Forward Primer	Reverse Primer
Smad2	TCCGGCTGAACTGTCTCCTACT	ACTGGCTGCAAATCCAAGCT
Smad3	CCAATGTCAACCGGAATGC	CGCACACCTCTCCCAATGT
Hnf4g	AAAAGAAGCGGTCAAAATGA	GCCCTCGTAGGTACTTCTTCTTGT
Bmf	CCAGAAAGCTTCATTATTGCA	TGCTGGTGTTGTTGCGTATGA
Them6	CCTTCGCGATGGTTTCGT	CTCCACCCTCCGTTTGCA
β-actin	CGAGTCGCGTCCACCC	CATCCATGGCGAACTGGTG
Cyclophilin	TATCTGCACTGCCAAGACTGA	CCACAATGCTCATGCCTTCTTTCA
Pri-miR-125b-1	GTCCCTGAGACCCTAACTTGTG	CAAGAGCCTAACCCGTGGAT
Pri-miR-125b-2	CCTAGTCCCTGGACCCTAACTT	TAGGTCCCAAGAGCCTGACTT
MIR184-Peak1	CAGACCGCCCCACAAGAG	TAAGGCTCGTGGTGGGTATAGG
MIR184-NEG	GAGGCTGCCTGGTCAACATC	ATTGCTCTTGATCACCCATCAGA
MIR125B-2 peak	AGAGGGTGGATGCCAGACTGT	GGCGGCGGCTGGTATT
MIR125B-2 NEG	TCCAGAAGAGTAGATTCCCATCAAT	TTGCCTCTGTTCCCCTATCTGT

Table 2.1: Validated qPCR primers sequences

2.9 Transfections

2.9.1 miRNAs transfection

Both EndoC-βH1 and MIN6 cells were transfected using Lipofectamine 2000 (Thermo Fisher Scientific) following manufacturer's instructions. Briefly, Lipofectamine and 5 nM of miR-125b or control mimic (Qiagen) were gently mixed after 5 mins incubation in transfection media OptiMEM (Thermo Fisher Scientific) and, after a further 20 mins incubation, the Lipo/RNA mimic mix was added in each flask/well. After a 4-5 hours incubation period at 37 °C with 5% CO_2, the transfection media was replaced with normal growth media and the flask left in the incubator for 24-48 hours prior to RNA/Protein extraction.

2.9.2 siRNAs transfections

MIN6 cells split at 50% confluency overnight were transfected with 50 nM of a mixture of four ON-TARGETplus Mouse Smad2, ON-TARGETplus Mouse Smad3 (Horizon Discovery) or ON-TARGETplus Non-targeting Pool (Horizon Discovery) siRNAs using Lipofectamine 2000 as described above in Chapter 0.

2.10 Luciferase assay

2.10.1 Plasmid generation and cloning

2.10.1.1 PCR

3'UTR or CDS sequences were downloaded using UCSC genome browser and two primer pairs were designed using NCBI Primer Blast (see Table 2.2 for primers list). Each sequence was amplified by PCR (polymerase chain reaction). Appropriate restriction sites were incorporated in the primer sequence and are indicated in bold in Table 2.2. The regions were amplified using Q5 polymerase (New England Biolabs). The products for these DNA fragments were run on a 1% agarose gel to assess specificity.

Name	Forward Primer	Reverse Primer	Restriction Enzymes
Taz 3'UTR	**CATGAGCTC**AGTCAGCTGTGGA-GATGCG	**CGTCTAGA**TATGGTGG-CACAACACAAATG	SacI/XbaI
M6pr 3'UTR	**CATGAGCTC**AGCTGGGTGAAGAG-TCGGAA	**CGTCTAGA**TCATTGAAAACAC-GGGTTAGGATG	SacI/XbaI
Tor2a 3'UTR	**CATGAGCTC**GGCCTCCCGACTCAC-ATTT	**CGGTCG**ACTTTTTTTCTAATAC-CAAGGCTAAGTT	SacI/SalI
Mtfp1 3'UTR	**CATGAGCTC**GTACCCGTCAGTGGA-GAAGC	**CGTCTAGA**AGGTTAGCACTTGGGTT-GTGA	SacI/XbaI
Tnks1bp1 CDS	**CATGCTAGC**CTTCTT-GCTCCGCTCAGAG	**CGGTCG**ACACACTTGCTGCTCCAG-TCTC	NheI/SalI
Gnpat 3'UTR	**CATGAGCTC**CAGTGTGCTCCTC-TATGCGAA	**CGTCTAGA**GTGGCTGGTTTTCC-TATTGGC	NheI/SalI
Gnpat CDS	**CATGAGCTC**CATGGAC-GTTCCTAGCTCCTC	**CGGTCGA**CGGTCCGGGG-TACTGAGGT	SacI/XbaI
Trnp1 3'UTR	**CATGAGCTC**CAGCGCTGACCTCAG-TACC	**CGTCTA-GA**TTGCCTCTCGTTATTTCGATTT	NheI/SalI
M6pr mutated 3'UTR	TGGGTCAGACACTTC**CAG**-TAGCCAGCAGGAGAGACAGG	CCTGTCTCTCCTGCTGGCTA**CTG**-GAAGTGTCTGACCCA	

Table 2.2: PCR primers sequences used to for Luciferase assay plasmid generation.

2.10.1.2 Restriction digestion and dephosphorylation.

Both PCR products and pmiRGlo plasmid DNA (Promega) were digested at 37 °C for 1-3 hours using 10 U of the two appropriate restriction enzymes in a total volume of 50 µL and then samples were run on a 1% agarose gel to check for successful digestion. The enzymes were inactivated by purification using QIAQuick PCR purification kit (Qiagen) and, to prevent re-ligation of the digested plasmid, the DNA was dephosphorylated using Shrimp Alkaline Phosphatase (rSAP) (New England

Biolabs) at 37°C for 30 minutes and the reaction stopped by incubation at 65°C for 5 minutes. The plasmid and fragments concentrations were quantified using Nanodrop.

2.10.1.3 Ligation

Resulting PCR products were then ligated into the plasmid with a vector with an insert ratio of 1:3/ 1:5 using T4 DNA ligase (New England Biolabs). This reaction was incubated at room temperature for 1 hour, heat-inactivated at 65°C for 10 minutes and then placed on ice to cool. A control sample containing just the plasmid without the insert was also generated.

2.10.1.4 Transformation of chemically competent cells

5-100 ng DNA was added to 10-Beta Competent E. Coli cells (New England Biolabs, Hitchin, UK), mixed and incubated on ice for 30 minutes. Cells were then heat shocked at 42 °C for 30 sec exactly and placed on ice for 5 minutes before being resuspended in 200μL of broth and incubated at 37 °C for 60 min with vigorous shaking (250 rpm). After incubation, 100 μL of cell suspension was plated onto pre-warmed agar plates containing ampicillin (AMP), since the plasmid contains an AMP resistance gene, and incubated overnight at 37 °C.

2.10.1.5 DNA extraction & Sanger sequencing

Individual colonies were picked and amplified in 4 ml of broth containing ampicillin (AMP) for DNA extraction using Qiagen Miniprep kit.

The DNA samples were digested with the same restriction enzymes described earlier and run on a 1% agarose gel to assess if the plasmid contained the inserts of interest. Two bands, one at approximately 7 kb indicating the plasmid and one between 1 and 2.5 kb indicating the insert, were expected. The samples that resulted successfully cloned were grown in 200 ml of liquid broth containing AMP, their DNA then extracted using a Qiagen Maxiprep kit and sent for Sanger se-quencing for verification.

2.10.1.6 M6pr 3'UTR mutation

Three-point mutations in the predicted miR-125b seed binding site were introduced in the pmiR-Glo-M6pr-3'UTR plasmid using site-directed mutagenesis. The sequence of primers designed using the Primer Design Program by Agilent (https://www.agilent.com/store/primerDesignProgram.jsp) is in Table 2.2. The mutation introduced by the primers are indicated in bold.

Briefly, 5ng of pmiRGlo-M6pr-3'UTR plasmid were mixed to 5X Phusion Buffer, 10mM dNTPs, 10 μM of reverse and forward primers and 0.2 μL of Phusion DNA Polymerase (Fisher) and incubated at 98 °C for 30 sec, then 98 °C for 10 min, 55 °C for 30 sec, 72 °C for 6 min for 12 cycles, and finally 72 °C for 5 min.

PCR product was added to 10-Beta Competent E. coli cells and plasmid was generated as described previously. Mutations were confirmed by DNA sequencing.

2.10.2 Dual-luciferase quantification

MIN6 cells at 50-60% confluency were co-transfected with 5ng of pmiRGlo plasmid DNA and 5 nM miRNA mimic (control or miR-125b) in a 48-well plate using Lipofectamine 2000 as described in section 2.9.1 . After 24 hours from transfection, a dual luciferase assay (Promega, #E1910) was used to measure the luciferase and Renilla luciferase activity within the transfected cells. The growth media was removed, and the cells were washed with 100 μl of once PBS. Then, cells were incubated for 15 min at room temperature with 50 μl passive lysis buffer for cell lysis and 10 μl of the lysate were transferred to a luminometer tube. First, 50 μl of Luciferase Reagent II was mixed with the lysate and luciferase activity was measured using a Lumat LB 9507 luminometer. Subsequently, 50 μl of Stop and Glo® reagent was added to the lysate to stop the luciferase activity and catalyse the Renilla luciferase reaction. Two blanks were made using 10 μl of passive lysis buffer following the same steps. The value generated from the blanks was subtracted from the firefly and Renilla luciferase values of the samples and a ratio between firefly and Renilla luciferase was generated and data presented as fold change.

2.11 SDS-PAGE

2.11.1 Western Blot

Cells or mouse islets (50-100) were lysed on ice in RIPA buffer (150mM NaCl, 1% Triton, 0.1% SDS (sodium dodecyl sulphate), 0.5% sodium deoxycholate and 50mM Tris, pH 8.0) supplemented with proteinase (Roche) and phosphatase inhibitors (Sigma Aldrich). Lysates were centrifuged at 16,000g for 15 minutes at 4 °C. Proteins were quantified by Bicinchoninic acid assay (BCA) and equal amounts of protein for each sample were analysed by SDS-PAGE under reducing conditions at constant voltage in polyacrylamide gels. The acrylamide used was adjusted according to the desired gel pore size depending on the size of protein analysed. The gels consisted of two parts: the stacking gel (5% polyacrylamide, 250 mM Tris-HCl, 0.1% SDS, 0.1% ammonium persulfate and 0.025% N,N,N',N'-tetramethylethylenediamine (TEMED), pH 6.8) and the resolving gel (8% poly-acrylamide, 375 mM Tris-HCl, 0.1% SDS, 0.1% ammonium persulfate and 0.025% TEMED, pH 8.8). Samples were then transferred onto Immobilion polyvinylidene difluoride (PVDF) membranes (Fisher). After blotting, the membranes were blocked in 5% milk or Bovine Serum Albumin (BSA), dissolved in PBS-Tween (TPBS) or Tris-buffered saline and 0.1% Tween (TBST) depending on the primary antibody used, and incubated overnight with primary antibody in blocking solution at 4°C (antibodies are listed in Table 2.3). After three washes with TBST, membranes were incubated in horseradish peroxidase (HRP)-conjugated secondary antibody for 1 hour at room temperature. Bands were detected (GE Healthcare) and imaged using chemiluminescence with ECL and x-ray films.

For cell membrane proteins such as GLP-1R, a lysis buffer containing 20 mM Tris, 150 mM NaCl, 1 mM EDTA, 1% NP-40, pH 7.4 plus complete mini EDTA-free protease (Roche) and phosphatase inhibitors (Sigma Aldrich) was used for the lysis. Urea loading Buffer (100 mM Tris-HCl [pH 6.8], 2.5% SDS, 4 mM urea, 50 mM dithiothreitol, 0.05% bromophenol blue) followed by 10 minutes at 37°C incubation was used for the SDS-PAGE

Name	Species	Source	Dilution
Anti-AMPK	Rabbit	Cell Signaling Technology	1/1000
Phospho-AMPKα (Thr172)	Rabbit	Cell Signaling Technology (40H9)	1/1000
Anti-SMAD2/3	Mouse	Santa Cruz Biotechnology (sc-133098)	1/1000
Anti-pSMAD3	Rabbit	Abcam (ab52903)	1/1000
Anti-M6PR	Mouse	Santa Cruz Biotechnology (sc-365196)	1/500
Anti-CTCF	Rabbit	Millipore (07-729)	1/1000
Anti-Tubulin	Mouse	Sigma Aldrich (T5168)	1/1000
Anti-GAPDH	Rabbit	Santa Cruz Biotechnology (2118S)	1/1000

Table 2.3: List of Antibodies used for Western Blot analysis

2.11.2 Phos-tag

Equal amounts of proteins/sample were diluted in 1X sample-loading dye solution, boiled for 5 min and separated in Running Buffer (Table 2.5) at constant voltage (25 V) at 4°C using gels prepared as described in Table 2.4.

After run completion, the gel was soaked in EDTA-Transfer buffer (Table 2.5) for 10 min before being transferred to PVDF membrane using Transfer buffer prepared as described in Table 2.5. Transfer was performed under constant-voltage conditions (25 V) for 16 h (O/N). Immunoblotting analysis was carried out following normal condition as described earlier (see section2.11.1)

Resolving Gel Solution	Stacking Gel Solution
30% Acrylamide/Bis Mixed Solution	30% Acrylamide/Bis Mixed Solution
1.5 mol/L Tris/HCl Solution, pH 8.8	1 mol/L Tris/HCl Solution, pH 6.8
5.0 mmol/L Phos-tag™ Solution (25µmol/l)	10 mmol/L MnCl2 Solution
10 mmol/L $MnCl_2$ Solution	10% SDS Solution
10% SDS Solution	TEMED (tetramethylethylenediamine)
10% Ammonium Persulfate Solution	
TEMED (tetramethylethylenediamine)	

Table 2.4: Phos-Tag running gel composition

Running Buffer	Transfer buffer	EDTA-Transfer buffer	3 × Sample-loading dye solution
25 mM Tris base	25 mM Tris	25 mM Tris	195 mM Tris-HCl (pH 6.8)
192 mM glycine	192 mM glycine	192 mM glycine	3.0% SDS
0.1% SDS	10% methanol	1 M EDTA (pH 8)	15% 2-mercaptoethanol
			30% glycerol

Table 2.5: Phos-Tag Buffers

2.12 ChIP-qPCR

2.12.1 Crosslink

Media from MIN6 cells grown at 80% confluency on 15cm adherent dish or mouse isletes isolated as described earlier (see section 2.3) was removed and replaced with 10 ml of cold PBS. Cells were fixed with 1% formaldehyde (Agar Scientific) for 10 minutes at room temperature and the cross-linking reaction stopped by incubating the cells for 5 minutes at room temperature with 125 mM glycine. Cells/islets were washed once with cold PBS and cells were harvested through the gentle use of a cell scraper and collected in Falcon tubes.

Cells/islets were pelletted at 2000g for 2 min at 4℃ and snap-freezed on dry ice. The pellets were stored at -80 4℃ until use.

Each pellet was lysated in 140μL of lysis buffer (see Table 2.6) using a tight pestle (VWR). 20 strokes repeated 3 times were performed for each sample and between each set the samples were incubated on ice for 5 min. Mouse islets and MIN6 cells were sonicated for 16 and 12 min, respectively, using Covaris 220 default settings.
The sonicated samples were collected in a clean tube and the lysates cleaned from debris by centrifugation at 21000 g for 5 min at 4℃. The Supernatant was then collected and stored at -80℃ until use. DNA fragmentation was assessed on 10% of the sample where DNA was extracted and run on 1% agarose gel.

2.12.2 Immunoprecipitation

Samples were incubated with 30 μL of Dynabeads Protein G and A in a 1:1 ratio (Thermo Fisher Scientific) in a total volume of 1ml of lysis buffer at 4℃ for 2 hours with end-over-end rotation for precleaning.

In parallel, 40μL of Dynabeads Protein G and A were coated with 3μg/ml of Rabbit anti-CTCF (07-729, Millipore), 5μg/ml of Mouse anti-SMAD2/3 (sc-133098, Santa Cruz Biotechnology) antibody or same concentration of negative control anti-Rabbit IgG (ab205718, Abcam) or anti-Mouse IgG

(NXA931, Merck), respectively, in a total volume of 250 µL of pre-blocking buffer (Table 2.6) for 2 hours at 4℃ while rotating.

After the incubation, the beads conjugated with the antibodies were collected on a Dynamagnet (30sec, on ice) and gently washed 3 times with working buffer and resuspended in adequate volume of the same buffer. Simultaneously, the pre-cleaned lysates were separated from the beads, transferred in a new collection tube and split in equal volumes between anti-CTCF and IgG coated beads for CTCF ChIP-qPCR and anti-SMAD2/3 and IgG for SMAD2/3 ChIP-qPCR in a total volume of 1ml of working buffer supplemented with complete EDTA-free proteinase inhibitors (Roche). Samples were incubated O/N at 4℃ with end-over-end rotation. 10% of the volume was stored at -20℃ and used as INPUT sample (half for qPCR and half for western Blot).

After the incubation, the supernatant was separated from the beads and stored at -20℃ to be used as flow-through sample.

The beads were washed once in cold low salt buffer for 5 min, once in cold high salt buffer and once in cold LiCl buffer for another 5 min before being washed 3X in TE buffer (see Table 2.7).

Sample were incubated at 37℃ for 15 min while shaking in 150µL of freshly prepared elution buffer (1% SDS, 0.1M NaHCO₃) to elute the Protein-DNA complexes from the beads. The step was repeated twice and samples were eluted in a total volume of 300 µL. 250 µL were used for DNA extraction to be used as immunoprecipitated (IP) sample whereas 50 µL were kept for Western Blot analysis.

2.12.3 Reverse cross-link and DNA extraction

Both Input sample and Immunoprecipitated sample were incubated with 5 µL RNAse A (10 mg/ml) (Thermo Fisher Scientific) at 65℃ for a minimum of 30 min to remove the RNA contaminant and then incubated with 4.5 µL of Proteinase K (20 mg/ml)(Sigma) and 12 µL of 5M NaCl (Sigma) and incubated at 37℃ for 5 h to reverse the cross-link reaction.

The DNA was extracted using first 1.25X Phenol:Cholorofom (Sigma). Phases were separated by centrifuging the sample for 5 min at 21,000x g and the aqueous layer was transferred to a new tube. 2.25X chloroform was then added and the phases were separated again by centrifuging the

samples for 5 min at 21,000x g. The aqueous layer was transferred to a new tube and the samples were precipitated O/N adding 3 µL of Glycogen (5mg/ml) (Thermo Fisher Scientific), 35 µL of Sodium Acetate (3 M), pH 5.5 (Thermo Fisher Scientific) and 875µL of 100% EtOH molecular grade.

2.12.4 qPCR analysis

DNA was used for quantitative real-time PCR as described in section 2.8 .

Ct values were used to perform the calculation which consists of evaluating the fold change difference between experimental sample and normalized input using this formula:

$$\Delta Ct \text{ [normalized IP]} = (Ct \text{ [IP]} - (Ct \text{ [Input]} - Log2 \text{ (Input Dilution Factor)}$$

where Input Dilution Factor is the fraction of the input chromatin saved. For these experiments, 10 µl of input were saved from a total volume of 110 µl recovered after sonication, thus obtaining an IP fraction that was 10 X the input fraction. Moreover, the Input was further diluted 4.5 times for the qPCR run, making the final dilution factor of the Input fraction (Input Dilution Factor)=10×4.5=45.

The percentage (% Input) value for each sample was calculated as follows:

$$Input \% = 100/2 \, \Delta Ct \text{ [normalized ChIP]}.$$

where "Input %" represents the enrichment of the DNA region of interest.

Lysis Buffer	Dilution buffer	Working buffer	Pre-blocking buffer
2% Triton X-100	50 mM Hepes pH8.0	1X Dilution buffer	Working buffer
1% SDS	140 mM NaCl	4X Lysis buffer	1% SDS
100 mM NaCl	1 mM EDTA	1X protease inhibitor	0.5% BSA
10 mM Tris-HCl pH 8.0	0.75% Triton X-100		
1 mM EDTA	0.1% Na-deoxycholate		
1X protease inhibitor	1X protease inhibitor		

Table 2.6: ChIP-qPCR buffers (1/2)

Low Salt Wash Buffer	High Salt Wash Buffer	LiCl wash buffer	1X TE buffer
1% Triton X-100	500 mM NaCl	0.25 M LiCl	10 mM Tris-HCl, pH 8.0
150 mM NaCl	1% Triton X-100	1% deoxycholate sodium	1 mM EDTA
20 mM Tris-HCl, pH 8.0	20 mM Tris-HCl, pH 8.0	10 mM Tris-HCl, pH 8.0	
0.1 % SDS	0.1 % SDS	1% NP40	
2mM EDTA	2 mM EDTA	1 mM EDTA	

Table 2.7: ChIP-qPCR buffers (2/2)

2.13 RNA immunoprecipitation and sequencing (RIP-seq)

Lysis Buffer	Lysis buffer with additives	Washing buffer
20mM Tris-HCL pH7.5	160u/ml of RNAsin (Promega)	50mM Tris-HCL pH7.5
150mM KCL	0.5mM DTT	300mM NaCl
0.5% NP40 (Igepal)	1X Protease inhibitors (Roche)	5mM MgCl2
2mM EDTA	1X Phosphatase inhibitors (Sigma Aldrich)	NP40 (Igepal) 0.05%
1mM NaF		

Table 2.8: RIP-seq buffers

2.13.1 Cell lysis

MIN6 cells were transfected with control or miR-125b mimic as described in section 2.9.1 . After 24 h the media from cells grown at 80% confluency on 15cm adherent dish was removed and cells washed twice with cold PBS. Cells were harvested through the gentle use of a cell scraper and collected in falcon tubes. Cells were then pelleted at 1100 RPM for 3 min at room temperature and lysated in cold 500μL of lysis buffer (see Table 2.8). The lysis was enhanced by pumping the samples 3X with a 25G syringe needle and once with a 27G syringe needle. The lysates were then spun at 16000g at 4℃ for 10 min and the supernatant collected in a clean tube. The pellets were stored at -80℃ until use.

2.13.2 Immunoprecipitation

The samples were incubated with 20 μL of Dynabeads Protein G (Thermo Fisher Scientific) in a total volume of 1ml of lysis buffer at 4℃ for 2 hours with end-over-end rotation for precleaning.

In parallel, 40μL of Dynabeads Protein G were coated with 6ug of mouse-anti-AGO2 (clone E12-1C9, Abnova) or mouse IgG (NXA931, Merck) and 1mg/ml of heparin for 2 hours at 4℃.

After incubation, the beads were collected on a Dynamagnet (30sec, on ice) and gently washed 3X with lysis buffer with additives (Table 2.8). After the last washing step, the conjugated beads were resuspended in lysis buffer with additives in a total volume of 500μL.

Simultaneously, the pre-cleaned lysates were separated from the beads, transferred in a new collection tube and split in equal volume between mouse-anti-AGO2 and mouse IgG coated Dynabeads Protein G where they were incubated at 4℃ for 16 hours with end-over-end rotation. 5% of each sample was saved at -20℃ as INPUT samples.

The following day the flow-through was transferred into a new collection tube and the beads washed once with lysis buffer and 3X with washing buffer (Table 2.8) for 5 min at 4℃.

2.13.3 DNAse treatment and RNA extraction

The samples conjugated to Dynabeads-Protein G were treated with 1μL of TURBO™ DNase (2 U/μL) (Thermo Fisher Scientific) for 15 min at room temperature to remove any DNA contaminant and subsequently treated with 3μL proteinase K (20mg/ml) (Thermo Fisher Scientific) at room temperature for 20 min while shaking to release the RNA bound to the proteins.

The RNA was extracted using acid Phenol:Chloroform, pH 4.5 (with IAA, 125:24:1) (Thermo Fisher Scientific) and precipitated O/N adding 3μL of Glycogen (5mg/ml) (Thermo Fisher Scientific), 35 μL of Sodium Acetate (3 M), pH 5.5 (Thermo Fisher Scientific) and 875μL of 100% EtOH molecular grade.

2.14 RNA sequencing

2.14.1 mRNA-seq library construction

2.14.1.1 Total RNA-seq

Library preparation was performed by WTCHG facility in Oxford using 1 μg of RNA per sample from 6 independent experiments in which control or miR-125b mimic were transfected into MIN6 cells as described in section 2.9.1

2.14.1.2 RIP-seq

Quantity and integrity of RNA samples from 6 independent experiments (described in section 2.13) were assessed using an RNA 6000 Nano Kit (Agilent) and an Agilent 2100 Bioanalyzer. mRNA enrichment was achieved from 1 µg of total RNA using a NEBNext Poly(A) mRNA Magnetic Isolation Kit (NEB). Generation of double stranded cDNA and library construction were performed using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB). NeBNext Multiplex Adapters (NEB) was used to perform ligation of the adapters. Each library was subsequently size-selected with SPRIselect Beads (Beckman Coulter). The adaptor-ligated DNA was PCR amplified using NEB-Next Ultra II Q5 Master Mix and Universal i5 and i7 primers provided in the NEBNext Kits.

2.14.2 Sequencing

2.14.2.1 Total RNA-seq

Sequencing was performed by WTCHG facility in Oxford as 75bp paired end reads on a HiSeq4000 according to Illumina specifications.

2.14.2.2 RIP-seq

Sequencing was performed by the Imperial BRC Genomics Facility as 75bp paired end reads on a HiSeq4000 according to Illumina specifications.

2.14.3 Sequencing analysis

For both Total RNA seq and RIP-seq, reads produced for each sample by the sequencing were provided in the FASTQ format. FastQC [486] was used to assess the quality of the obtained samples, and discard reads not passing filtering with default parameters.

```
fastqc -d tmp -o fastqc_out -f fastq <input_directory>/ *.fastq.gz
```

Remaining reads were then mapped to the mouse transcriptome (GRCm38, cDNA and ncRNA) using Salmon [487]:

```
salmon quant -i $HOME/libraries/< Mus_musculus.GRCm38.cdna_ncrna_index > -l A \

    -1 ${dir}/${samp}_1.fastq.gz \

    -2 ${dir}/${samp}_2.fastq.gz \

    -p 8 -o $HOME/quants/${samp}_quant
```

For Total RNA-seq, ~31 million reads were obtained and 91% were uniquely mapped. For RIP-seq, ~12 million reads were obtained and 70% were uniquely mapped. Aligned sequences where further processed to obtain differential gene expression data between control and miR-125b overexpressed samples using the DESeq2 (v1.20.0) [488].

Briefly, DESeq2 default normalization method and adjusted p-value threshold <0.1 was used for differential expression analysis in R using relevant BioConductor packages [489].

The complete script used for the analysis can be found in Supplementary Table 3 in Appendix A.

2.15 eiCLIP

Lysis buffer	High salt wash buffer	PNK buffer	5x PNK pH 6.5 buffer
50 mM Tris-HCl pH 7.4	50 mM Tris-HCl pH 7.4	20 mM Tris-HCl pH 6.5	350 mM Tris-HCl pH 6.5
100 mM NaCl	1 M NaCl	10 mM MgCl2	50 mM MgCl2
1 % Igepal CA-630 (Sigma Aldrich)	1 % Igepal CA-630 (Sigma Aldrich)	0.2 % Tween-20	5 mM dithiothreitol
0.1 % SDS	0.1 % SDS		
0.5 % sodium deoxycholate	0.5 % sodium deoxycholate		

Table 2.9: eiCLIP buffers (1/4)

4x Ligation buffer	PK-SDS buffer	2X Bind and Wash buffer
200 nM Tris-HCl pH 7.8	10 mM Tris-HCl, pH 7.4	10 mM Tris-HCl pH 7.4
100 mM NaCl	100 mM NaCl	2M NaCl
4 mM dithiothreitol	1 mM EDTA	1 mM EDTA
	0.2% SDS	

Table 2.10: eiCLIP buffers (2/4)

Tween Wash Buffer	No Tween Wash Buffer
5 mM Tris-HCl pH 8	10 mM Tris-HCl pH 8
1M NaCl	2M NaCl
0.5 mM EDTA	1mM EDTA
0.05 % Tween-20	

Table 2.11: eiCLIP buffers (3/4)

2X Bind and Wash buffer	Tween Wash Buffer	Tween Wash Buffer
10mM Tris-HCl pH 7.4	5 mM Tris-HCl pH 8	10 mM Tris-HCl pH 8
1mM EDTA	0.5 mM EDTA	1 mM EDTA
2M NaCl	1M NaCl	1M NaCl
	0.05% tween	

Table 2.12: eiCLIP buffers (4/4)

3' end RNA dephosphorylation mix	µL	3' adaptor ligation mix	µL	Recj- deadenylase mix	µL
5X PNK pH 6.5 buffer (Table 2.9)	4	4X ligation buffer (Table 2.10)	5	Buffer 2 (NEB)	2
T4 PNK (NEB)	0.5	0.1M ATP	0.2	Recjf endonuclease (NEB)	0.5
RNasin (Promega)	0.5	T4 RNA ligase (NEB)	0.5	RNasin (Promega)	0.5
ddH2O	15	DMSO	0.6	PEG 400	4
		RNasin (Promega)	0.4		
		ddH2O	3.3		
		PEG 8000	7		
		L7-Cy5 adapter 10µM	3		

Table 2.13: eiCLIP reaction mixes (1/3)

Reverse transcription mix	µL	cDNA Adapter Ligation mix	µL
5X SSIV buffer	3	DMSO	1
0.1M DTT	0.5	10x RNA ligase buffer (NEB)	2
RNasin (Promega)	0.25	0.1M ATP	0.2
Superscript IV (Thermo Fisher Scientific)	0.25	RNA ligase (NEB)	0.5
1 pmol biotenilated RT-primers + 10mM dNTPs mix	2	5'end adapter	1
		ddH2O	5.8

Table 2.14: eiCLIP reaction mixes (2/3)

Preparatory PCR amplification mix	µL	Final-PCR mix	µL
5X Phusion Buffer (Fisher)	2	Accuprime Supermix 1	21
dNTPs	0.25	10 µM P5/P7 indexed primer mix	1
10 µM P5/P3 solexa primer mix	0.25		
Phusion DNA Polymerase (Fisher)	0.1		
ddH2O	5.4		

Table 2.15: eiCLIP reaction mixes (3/3)

2.15.1 UV Cross-link

Media from cells grown at 90% confluency on 15cm adherent dish was removed and replaced with 10 ml of cold PBS. Plates were placed on ice-filled tray and were crosslinked at 150mJ/cm2 at 254 nm using s Stratalinker UV crosslinker. Cells were immediatelly harvested through the gentle use of a cell scraper and aliquoted accordingly. From each plate 2 aliquots of 1 ml each were generated.

Finally, cells were pelleted at 376g for 1 min at 4°C and snap freezed on dry ice and pellets stored at -80 C until use.

2.15.2 Beads preparation

Immunoprecipitation was performed using antibodies conjugated to magnetic beads.

100μL/sample of Dynabeads Protein G (Thermo Fisher Scientific) was washed twice with 900μL of lysis buffer (Table 2.9), resuspended in the required volume conjugated with 100ng of anti-hnRNPC (sc32308, Santa Cruz Biotechnology) or 1.5 ug of anti-AGO2 (MABE253, Sigma) antibodies used per sample. 100μL of beads were not conjugated to the antibody and kept to be used as a negative control (No Ab). The tubes were rotated for 60 min at 4°C while proceeding with sample preparation. The beads were washed once in high salt buffer and twice in lysis buffer and resuspendend in 100 μL of lysis buffer per sample.

2.15.3 Sample Preparation and immunoprecipitation

Cell pellets were resuspend in 1 ml of ice-cold lysis buffer supplemented with 1X proteinase inhibitor cocktail.

1:20 and 1:1000 RNAse I (10U/L) (Thermo Fisher Scientific) dilution were prepared accordingly in ice-cold lysis buffer and added to the appropriate samples together with 2μL of DNA Turbo DNAse (Thermo Fisher Scientific). The sample treated with the lowest dilution of RNAse I was used as a negative control. The samples were incubated for exactly 3 min at 37°C whilst shaking at 1100 rpm and immediately transferred on ice for another 3 min.

The samples were then centrifuged at 4°C at 20,000x g for 10 min and the lysate transferred to a new microcentrifuge tube. 5% of volume was removed and kept in order to be used as size matched input control (see next section). These lysates were then mixed with the antibody-conjugated beads (or no-antibody beads used as negative control) at 4 °C for 16 hours. The beads were then washed twice in high-salt wash buffer (Table 2.9) for 5 mins and another two times with PNK wash buffer (Table 2.9) before being resuspended in 1ml PNK wash buffer.

2.15.4 Size matched input preparation

1X volume of Sera-Mag SP3 beads (Sigma) and 2X volume of 100% ethanol was added to each input sample recovered from the previous step in order to initiate binding between protein-RNA

complexes and magnetic beads. The tubes were rotated at 4 °C for the duration of the immunoprecipitation step (16 hours).

The beads were then washed twice with high salt wash buffer for 5 mins and two more times with PNK wash buffer before being resuspended in 1ml PNK wash buffer.

2.15.5 Adapter ligation

3'-end dephosphorylation mix (Table 2.12) was added to the samples that were incubated at 37 °C for 20 min while shaking at 1100 rpm. By performing this step, the RNA attached to the immunoprecipitated protein was dephosphorylated at the 3' end. After incubation the samples were washed once in PNK buffer, then twice in high salt wash buffer. Samples were then resuspended in 10 μl 3' adaptor ligation mix (Table 2.12) to ligate the adaptor to the dephosphorylated 3'-end of the RNA molecules. The samples were incubated at 16 °C ON while shaking at 1,100 rpm.

After incubation, the samples were washed once with PNK buffer, twice with high salt wash buffer and once again with PNK buffer. 20μL of the Recj-deadenylase mix (Table 2.12) was then added to the samples and tubes incubated for 30 mins at 37 ℃ while shaking. This step removes the free adaptors that are not ligated to the RNA molecules.

After incubation, the samples were washed twice in high salt buffer and once in PNK buffer before being resuspended in 20μL of NuPAGE sample loading buffer and incubated at 80 °C for 5 mins.

2.15.6 Protein-RNA complex visualization

Samples were loaded onto a 4-12 % NuPAGE Bis-Tris 1.5 mm 10-well gel (Thermo Fisher Scientific) alongside 5μl of protein size marker. Gels were used in conjunction with the XCell SureLock Mini-Cell Electrophoresis System (Thermo Fisher Scientific) filled with 0.5 l of NuPAGE MOPS-SDS running buffer (Thermo Fisher Scientific) for 60 min at 180 V. Once the run was completed, the dye front was removed, and the samples transferred to 0.2μm Nitrocellulose Membrane (Whatman) using the XCell II Blot Module (Thermo Fisher Scientific) filled with 1X NuPAGE transfer buffer and 10% methanol for 1.5 h at 30 V. After transfer, the membrane was washed with PBS buffer,

wrapped in saran wrap and the protein-RNA complexes were visualised using fluorescent imager Odyssey LI-COR CLx. Images at full resolution from the instrument were saved as JPEG.

2.15.7 RNA isolation

The image generated by the LI-COR was used as a guide to produce a mask to cut the desired RNAs fragments from the nitrocellulose membrane and the high-RNase sample was used as a guide to examine the specificity of the protein-RNA complex. The membrane fragments were then placed into LoBind tubes and 200µl PK-SDS buffer (table 2.10) additioned with 10µl proteinase K (Roche) was added. The mix was incubated at 50 °C for 1 hour while shaking at 1,100 rpm.

The supernatant was then collected and added to a 2 ml Phase Lock Gel Heavy tube (VWR), along with 400µl of neutral Phenol:Chloroform:Isoamyl Alcohol (Sigma Aldrich). Tubes were incubated at 30 °C for 5 min while shaking at 1,100 rpm, and the phases were separated by centrifuging for 5 min at >18,000 rpm. The aqueous layer was then transferred to a new tube and 340 µL of oligo binding buffer (Zymogen) alongside 1360 µL of 100% of EtOH were added and the resulting Mix placed on ice for 5 min before proceeding with a centrifugation step using the oligo clean and concentrator spin column. The RNA was then eluted using water in a new microcentrifuge tube.

2.15.8 Reverse transcription

RT biotinylated primers were added to the samples and their annealing to the RNA was performed at 65 °C for 5 min. Samples were then cooled to 25 °C until the next step where the reverse transcription mix (Table 2.13) was added to the samples for the reverse transcription. Samples were incubated at 25 °C for 5 min, 50 °C for 5 min, 55 °C for 5 min and then cooled to 4 °C. Free adaptors were removed by incubating the sample with Exonuclease III enzyme for 15min at 37 °C.

2.15.9 Streptavidin isolation

The samples containing the biotinylated group attached to the adapter ligated in previous step were separated using MyOne Streptavidin beads (Thermo Fisher Scientific). First, 20 µL of beads per sample were washed twice with Bind and Wash buffer and then resuspended in 100µL of the

same buffer. Then, beads were added to the sample and incubated at room temperature for 10 min.

After incubation, the samples were washed twice with 1X Bind and Wash buffer, twice with 0.05% tween wash buffer and twice with no tween wash buffer (Table 2.11) before being resuspended in PNK buffer and transferred in a new microcentrifuge tube.

2.15.10 cDNA 5' end adapter ligation

PNK buffer was removed and samples were resuspended in 11µL of adapter ligation mix (Table 2.13) containing 9µL of PEG 8000 and then incubated at 16 °C for 1 hour.

After incubation, the samples were washed twice with 1X Bind and Wash buffer, twice with 0.05% tween wash buffer, twice with no tween wash buffer (Table 2.14) and once with PNK buffer before being resuspended in 22 µL of nuclease free water.

The cDNA was eluted from the beads by incubating the samples at 80 °C for 3 min. This step was repeated twice, and the samples were eluted in a total volume of 44 µL of nuclease free water.

2.15.11 cDNA library PCR

Prior to PCR amplification of the entire cDNA, a preparatory PCR was conducted on a small portion of cDNA in order to identify the appropriate number of cycles to be used for amplification. The preparatory PCR mix (Table 2.15) was processed using the PCR programme in table 2.16.

	1 cycle	20-25 cycles			1 cycle	Hold
Temp	94 °C	94 °C	68 °C	72 °C	72 °C	25 °C
Time	2 min	20 s	30 s	45 s	3 min	hold

Table 2.16: PCR cycles for eiCLIP cDNA library amplification

12µl of PCR product were then mixed with 2 µl 6X TBE loading buffer and loaded onto a 6 % TBE gel. The XCell SureLock Mini-Cell Electrophoresis System was used for the gel electrophoresis, with

samples electrophoresed at 180 V for 30min. 2 µl SYBR safe stock solution was diluted in 10 ml TBE buffer, and the TBE gel immersed in this solution to visualize cDNA visualised by UV transillumination. If over-amplification was present in the form of smears extending above a solid band, then the cycle number was re-adjusted until over-amplification was no longer present, using reference images in [487] as a guide. Once the desired cycle number had been established, 30µL of the final-PCR mix (Table 2.15) was added to 20µL of cDNA and the samples were amplified using the program reported in Table 2.16.

Once the final library was constructed, primer-primer artefacts were removed using the select-a-size DNA clean and concentrator columns (Zymogen) and samples were eluted in 21µL of DNA elution buffer. Finally, size distribution and accurate quantification was achieved by running 1 µL of the final cDNA library on a high sensitivity DNA bioanalyzer chip (Agilent).

2.15.12 Bioinformatic analysis

The sequencing analysis was performed using NextSeq500 platform. ~4.5 million reads with an average length of 41 nt were obtained and, of those, ~2 million reads (45.7%) were uniquely mapped to the Homo sapiens (human) genome assembly GRCh38 (hg38). The results obtained from the sequencing were bioinformatically analysed by our collaborators Chris Sibley and Nejc Haberman. Briefly, sequencing data were uploaded into iMaps (https://imaps.genialis.com/iclip), a a web-based application that enables a user-friendly analysis and visualization of CLIP data, and AGO2 binding sites were identified using default settings using the Paraclu algorithm (https://zenbu-wiki.gsc.riken.jp/zenbu/wiki/index.php/Paraclu).

The most highly expressed miRNAs in Hela cells were selected using a filtering system based on the threshold of median reads obtained from single guide RNA (sgRNA) library [490] whereas 7-mer seed sequences of these miRNAs were downloaded from the TargetScan Human Release 7.2 database (http://www.targetscan.org/vert_72/).

The sequence enrichment of the 7-mer reverse completements of the most highly expressed miRNA targets were calculated and plotted as an heatmap where the strongest crosslink position, or rather the ones that had the maximum number of reads, were used as reference point.

2.16 GO and pathway enrichment analysis

GO analysis and Kyoto Encyclopedia of Genes and Genomes (http://www.gsea-msigdb.org/gsea/index.jsp) pathway analysis were conducted to identify to identify the biological function of the differentially expressed genes [491].

2.17 Statistical Analysis

For comparisons between two groups, statistical significance was assessed using two-tailed student´s t-test. For comparisons between more than two groups a one-way or two-way ANOVA tests were performed. Paired analyses were used in experiments with matched designs. Un-paired analysis on the logarithmic values was used in experiments where fold-changes were analysed.

GraphPad Prism 8.0 software was used for data analysis. Statistical significance was taken as p<0.05. Errors are represented as the standard error of the mean (SEM).

Chapter III: mir-184 transcription regulation

3.1 Introduction

MiR-184 expression is downregulated in prediabetic and diabetic mice [397] and it has been demonstrated to play an important role in the regulation of β-cell proliferation during compensatory expansion in pregnancy and obesity [397, 401]. In MIN6 cells, miR-184 inhibits insulin secretion through the repression of Slc25a22, a glutamate transporter that plays a role in the control of GSIS [468].

As described earlier, both our group and others have previously shown that the expression of miR-184 is regulated by glucose [397, 402, 458] . Recently, our group demonstrated that islet miR-184 expression is impaired upon both AMPK (AMPKdKO) and its main upstream kinase LKB1 (LKB1KO) deletion in mouse pancreatic β-cells [458]. Conversely, the activation of AMPK using specific activators [492, 493] resulted in upregulation of miR-184 in MIN6 cells and human islets. Moreover, glucose failed to regulate miR-184 expression in AMPKdKO and LKB1KO islets, demonstrating that AMPK is responsible for the effect of glucose on the expression of this miRNA [458]. Interestingly, primary miR-184 was also down-regulated in AMPKdKO islets, suggesting regulation at the transcriptional level [458]. However, the exact mechanism by which miR-184 expression is regulated by AMPK remained to be investigated.

The exact location of the miR-184 promoter has not been precisely annotated or experimentally validated. In human, miR-184 resides within region 25.1 on the q-arm of chromosome 15 [494], whereas in the mouse genome miR-184 is located in an imprinted locus on mouse chromosome 9 and it has been suggested that the transcription start site (TSS) might be located up to ~78 kb upstream of the pre-miR-184 loop [495].

ATAC-seq (Assay for Transposase-Accessible Chromatin followed by high-throughput sequencing) is a method used in molecular biology to map chromatin accessibility genome-wide. This technique targets accessible DNA with a Tn5 transposase, a member of the RNase superfamily, which cuts

and simultaneously inserts sequencing adapters into accessible regions of the chromatin. Fragments tagged with the adapters can then be sequenced and the resulting reads used to identify regions of open chromatin [496].

Interestingly, our lab has recently shown that islets from LKB1KO mice [458] contain two open chromatin peaks (ATAC-peaks) located upstream (~25 and ~72 Kb) of *MIR184* that are significantly more open (or "Upregulated") when compared to control islets (Figure 3.1) [458]. Open chromatin in the two regions upstream of *MIR184* could therefore be subjected to differential binding of regulatory molecules such as transcription factors and thus play a role in the regulation of miR-184 transcription.

It was previously shown that in CH12 cells, BHLHE40, ETS1, GCN5 and p300 bind both regions located within the two open chromatin peaks upstream of *MIR184* [497]   whereas CCCTC-binding factor (CTCF), a transcription factor with important insulator activity, has been shown to bind in liver the region located 25 Kb upstream of *MIR184* [458]. This indicates that these TFs might be involved in the regulation of miR-184 transcription and the regulation could occur also in pancreatic β-cells [458].

CTCF is an 82-kDa architectural protein which plays an essential role in the epigenetic regulation of gene expression [498]. CTCF was originally characterized as a transcription factor able to bind to a wide range of sequences and control gene expression via the activation or repression of promoters [499]. However, CTCF was later found to display properties that are characteristic of insulators [500, 501]. In fact, CTCF can, either alone or in combination with other proteins, mediate enhancer–promoter communication and arbitrate both inter-chromosomal and intra-chromosomal interactions [498, 502]. Moreover, while this TF can act as an enhancer blocker in specific loci [502], it has also been found significantly enriched alongside with histone modifications that are characteristic of active enhancers such as H3K4me1, H3K4me2, H3K27 and H3K27ac. A highly conserved DNA-binding domain with 11 zinc fingers [503] is present within this protein, and DNA binding occurs through a 52-bp sequence containing four CTCF-binding modules [504, 505] called CTCF-binding motifs, which are present at ~55,000–65,000 sites in mammalian genomes [506].
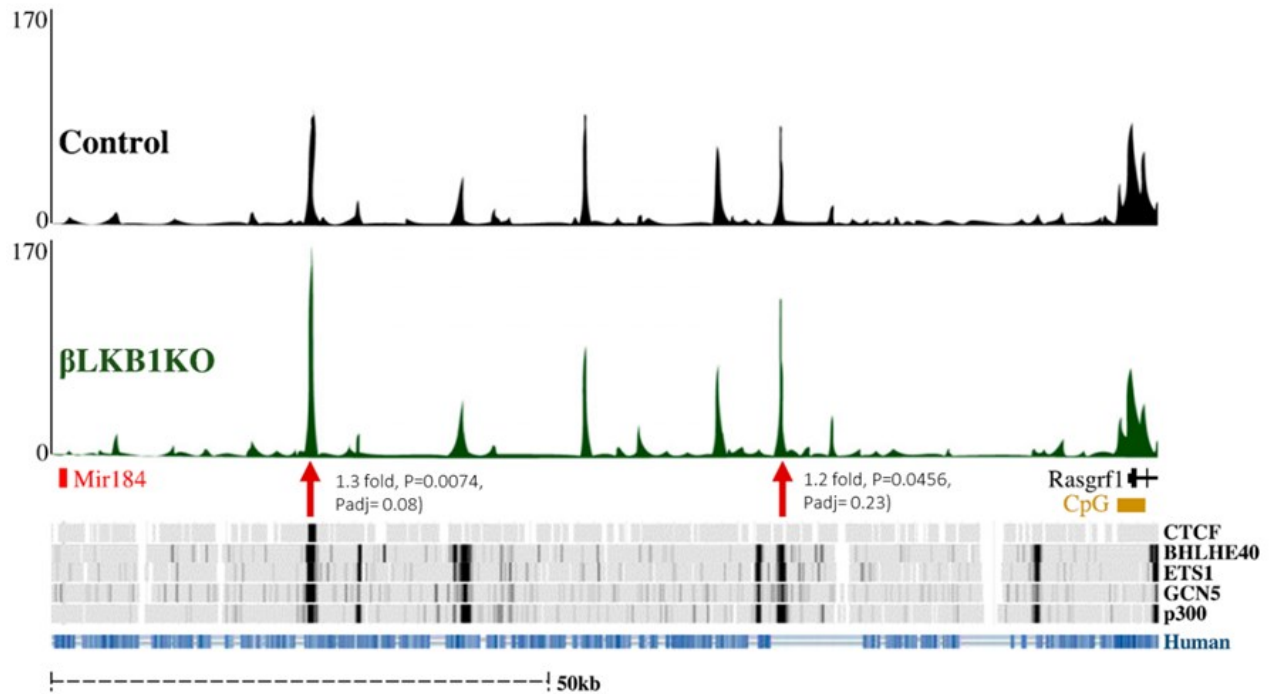
Figure 3.1: ATAC-seq enrichment profiles integrated with ENCODE CTCF and transcription factors ChIP-seq datasets.

The genomic region 100Kb upstream of (pre-) miR-184 shows 2 ATAC-seq peaks significantly more accessible/open (red arrows) in βLKB1KO (green) vs. control (black) islets. Enrichment ChIP-seq profiles for CTCF in liver and BHLHE40, ETS1, GCn5 and p300 in CH12 cells are shown. Alignment with the human DNA track is reported in blue. Modified from Martinez et al. [507]F

We thus hypothesised (Figure 3.2) that increased CTCF binding to the most proximal region upstream of *MIR184* results in miR-184 repression and reduced miR-184 levels in both AMPKdKO and LKB1KO islets. While it has been demonstrated that increased glucose and insulin concentration can upregulate CTCF in β-cells [508], the mechanism by which this happens has not been deciphered yet. Given the fact that glucose can regulate both AMPK and CTCF, we further hypothesised that AMPK is the mediator of the glucose-dependent regulation of CTCF. Therefore, this chapter aims to determine whether AMPK controls CTCF activity by regulating protein expression, phosphorylation and/or binding to *MIR184* regulatory regions in mouse islets. Finally, in order to facilitate future studies on miR-184 transcriptional regulation, I also aimed to identify *MIR184* TSS in mouse islets using a computational approach.
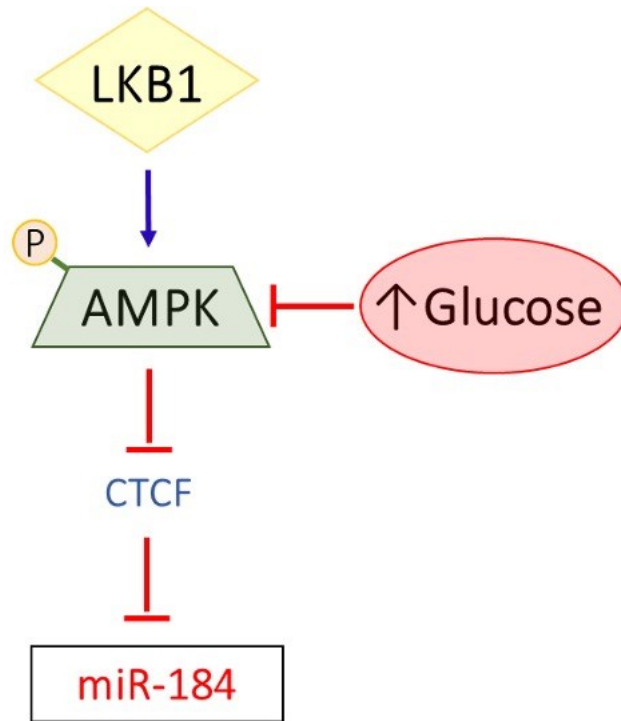
Figure 3.2: Simplified schematic representation of the hypothesis and aims of the chapter.

3.2 Results

RNA-seq data from AMPKdKO and LKB1KO mouse islets [247] demonstrated that CTCF expression is not altered by the absence of these enzymes in β-cells at the mRNA level. However, AMPK could modulate CTCF activity at different levels, such as quantity of the protein, cellular localization [509] and phosphorylation.

3.2.1 Effects of AMPK activation on CTCF protein levels

In order to investigate whether AMPK acts as an inhibitor of CTCF protein expression, I cultured MIN6 cells for 24 hours with and without the combination of two specific AMPK activators Compound 13 (C-13) and 991 (C-991) [492, 493]. These two compounds activate AMPK by acting on two different regulatory subunits [492, 493], therefore the combination of the two was used to enhance AMPK activation. Moreover, given that glucose has been shown to regulate both AMPK and CTCF [458, 508] but it is unclear whether AMPK acts as an intermediate for this regulation, MIN6 cells were also cultured at both low and high glucose concentrations (5.5 or 25 mM, respectively).

As expected, a clear increase in CTCF protein was observed in cells cultured at high glucose concentration, while, conversely, pAMPK, the active form of AMPK, was down-regulated (Figure 3.3). Surprisingly, the activation of AMPK led to increased levels of CTCF in both low and high glucose concentrations (Figure 3.3). These results suggest that glucose regulates CTCF expression independently of AMPK and that reduced miR-184 expression in MIN6 upon AMPK activation is not mediated by a reduction in CTCF protein levels.
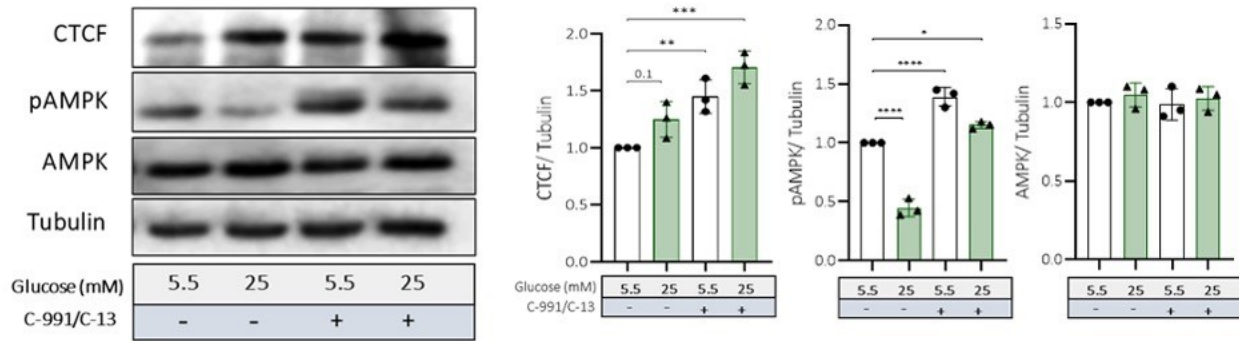
Figure 3.3: Activation of AMPK increases protein levels of CTCF
Representative western blot of AMPK, phosphorylated AMPK (pAMPK) and CTCF in MIN6 cells cultured in low (5.5 mM) and high (25 mM) glucose concentrations in the presence (+) or absence (-) of specific AMPK activators C-991 (20µM) and C-13 (50µM) for 24 hours. Tubulin is used as a loading control. CTCF, AMPK and pAMPK quantitative analysis of the immunoblots was determined by ImageJ. Each dot represents an independent experiment (n=3). Data are expressed as mean ± SEM relative to control levels. Values were normalised against the levels of Tubulin. * p<0.05 ** p<0.01, *** p<0.001, **** p<0.0001. P values were determined by two-way ANOVA with Sidak´s multiple comparisons test, with a single pooled variance on the normalized values.

3.2.2 Effect of AMPK phosphorylation activity on CTCF phosphorylation

CTCF phosphorylation has been previously shown to impair CTCF binding to its DNA target sequences [1, 2]. Interestingly, an analysis performed using Scansite4 [3] suggested that AMPK can potentially phosphorylate serine 461 in CTCF through a basophilic serine/threonine kinase group (Figure 3.4 A). Since a commercial antibody capable of recognising this specific phosphorylated site of CTCF is not available, Phos-Tag™ SDS-PAGE was used to detect this change. This method uses a synthesized chemical compound that, capturing phosphorylated Ser/Thr/Tyr and His/Asp/Lys, induces a mobility shift of the phosphorylated proteins on the gel during an SDS-PAGE. This allows to separate and therefore identify phosphorylated forms of proteins from their non-phosphorylated counterparts using an antibody against the (total) protein of interest.

To determine whether AMPK phosphorylates CTCF I performed Phos-Tag™ SDS-PAGE with protein extracts from MIN6 cells treated with AMPK activators (C-13 and C-991) or vehicle (DMSO) at high glucose concentration (25 mM). As expected, a small up-shifted migrating band (~20kDa), in samples treated with AMPK activators compared to controls was detected using an antibody against

(total) AMPK, suggesting that the Phos-Tag^{TM} compound was able to detect the phosphorylation of AMPK upon activation.

When an antibody against CTCF was used, a broad smear was detected in both samples suggesting firstly that CTCF phosphorylation might not change upon AMPK activation, and secondly that other types of post-translation modifications of CTCF might happen independently from AMPK activation. However, the addition of more controls is necessary to further investigate these aspects. For example, phosphorylation of CTCF by casein kinase 2 (CK2), which has been shown to phosphorylate different serines of this protein [510], could provide a positive control whereas the use of a phosphatase could instead add an appropriate negative control to the experiment. Also, the addition of a sample lacking Phos-Tag might be necessary to demonstrate whether the molecular weight shifts are due to the post-translation modifications or to the way the samples were handled during the experiment and/or to the Phos-Tag intrinsic background.

Figure 3.4: PhosTagTM didn't detect enhanced CTCF phosphorylation upon AMPK activation

A. Predicted motif sites shows the summary of the information obtained from the analysis performed using ScanSite and includes the candidate consensus site (Sequence). In Sequence alignment the red rectangle shows the conservation of the site amongst different species.

B. The plot illustrates the percentile rank for S461-AMPK predicted phosphorylation site in mouse compared to the distribution of all potential phosphorylation sites for AMPK motif sequences in the vertebrate subset of SWISS-PROT. A query site is tentatively scored as a "hit" if the final sequence scores (S_f) value falls within the top 0.2% (high stringency), 1% (medium stringency) or 5% (low stringency) of all scored sites within vertebrate SWISS-PROT

C. Representative PhosTag™ immunoblot analysis of CTCF and AMPK in MIN6 cells cultured in the presence or absence of specific AMPK activators C-991 (20μM) and C-13 (50μM) for 24 hours. Non-specific band is used as loading control. n=2 independent experiments

116

3.2.3 CTCF binds to an open chromatin region 25Kb upstream *MIR184* in mouse islets

As mentioned earlier, in liver cells CTCF binds the region chr9:89721201-89723471 located ~25 Kb from *MIR184* (Figure 3.1, [497]). In order to demonstrate that the binding of CTCF also occurs in pancreatic islets, I performed Chromatin Immunoprecipitation followed by quantitative PCR (ChIP-qPCR) using an antibody against CTCF on mouse islets. ChIP-qPCR allows for the identification of specific protein–DNA interactions, but in order to guarantee reliable results it needs to be carefully optimised. Figure 3.5 A details the different steps involved in this technique.

The original ChIP-qPCR protocol was provided by Ferrer's group at Imperial College and for the purpose of my thesis I have adapted it to mouse islets in order to detect CTCF binding sites.

The first stage of ChIP-qPCR protocol requires the crosslink between the proteins and the DNA strands to which they are bound. The protocol provided by Ferrer's group was previously optimized in human islets and required an initial 10 minutes incubation with 1% formaldehyde solution, so I decided to apply the same settings for the cross link to mouse islets and therefore no further optimization was required for this step.

Adequate chromatin fragmentation (200-500 bp) is essential to provide consistency when analysing different samples simultaneously and at the same time improves the efficiency of the technique, leading to higher recovery and lower background [511]. After testing different sonication methods, I decided to use a Covaris 220 shearing instrument and identified 16 minutes as the optimal running time. This generated DNA fragments of optimal length, falling within the range of 150-500 bp (Figure 3.5 B).

For the following immunoprecipitation step, I tested two different concentrations of anti-CTCF antibody (3 and 15ug/ml) and both protein G and A magnetic beads. The combination of both protein G and A in a 1:1 ratio combined with 3ug/ml of antibody proved to be optimal as in the CTCF-immunoprecipitated sample (CTCF IP) I could clearly detect by western blot a band around 130 KDa corresponding to CTCF, whereas nothing was detected in the sample immunoprecipitated using IgG antibody (IgG IP, negative control) (Figure 3.5 C).
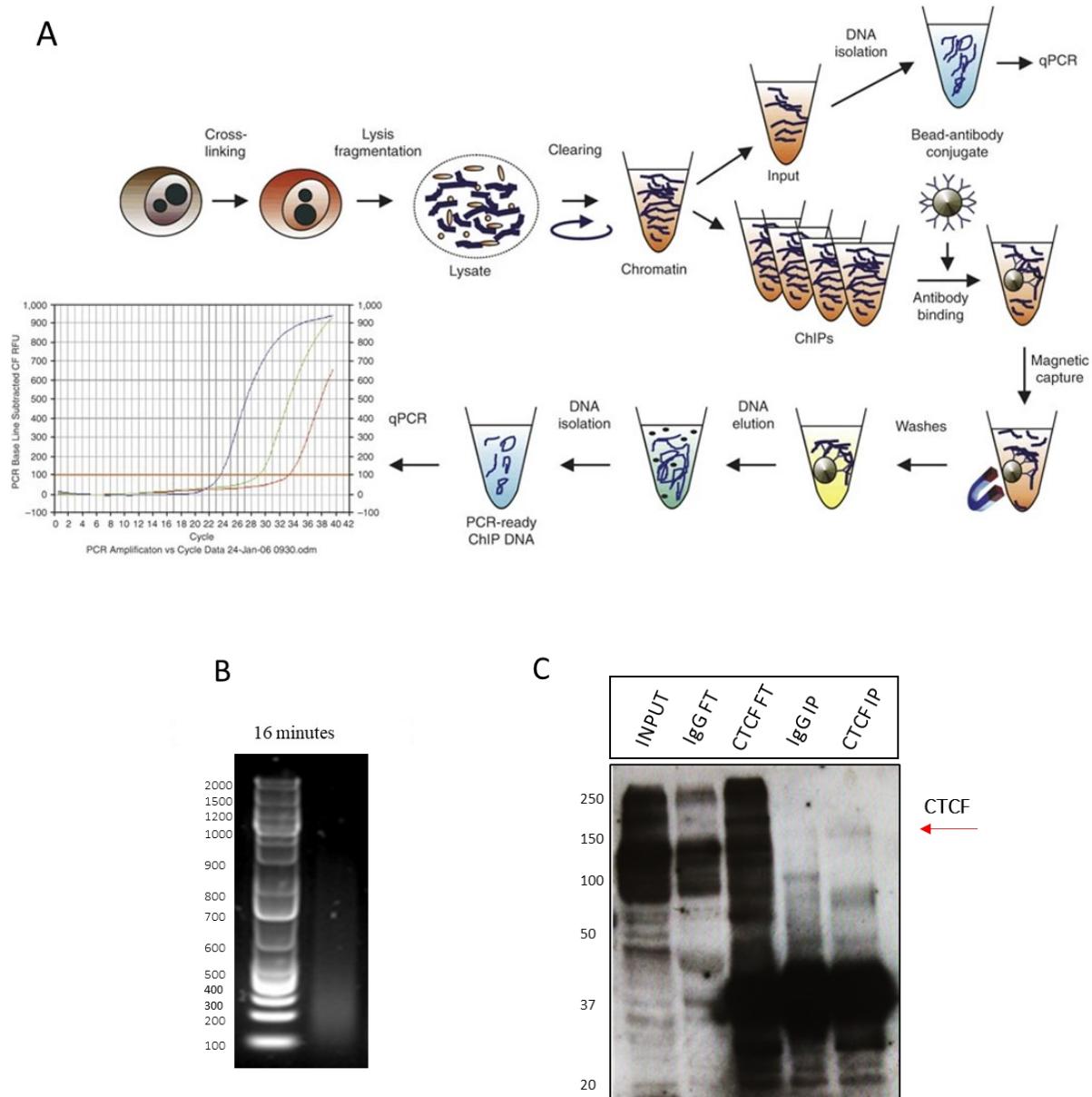
Figure 3.5: CTCF ChiP-qPCR optimization in mouse islets.
A. Schematic representation of the main steps of the ChIP-qPCR protocol.
B. Agarose gel electrophoresis assessment of chromatin fragmentation for C57BL/6 islets. The red arrow indicates the optimal average DNA fragments size (200nt) obtained with DNA sonicated for 16 minutes using Covaris 220.
C. Western Blot analysis obtained with the optimized protocol in C57BL/6 islets. CTCF protein (130 kDa) levels in samples where immunoprecipitation was performed with an anti-CTCF antibody (CTCF IP) or an IgG control (IgG IP). Input: 5% of the total lysate submitted to immunoprecipitation (IP); FT (Flow-through): 5% of the extract recovered after IP.

However, it is important to draw attention to the fact that, in Figure 3.5 C, the band that corresponds to CTCF in the input lane seems to be characterized by a broad smear. This could indicate that, as mentioned earlier, this protein undergoes post-translation modifications in β-cells under circumstances yet to be identified. Additionally, from the same blot it is also possible to notice that the CTCF antibody unfortunately recognises numerous bands with unexpected molecular weights, indicating that the specificity of this antibody could be in question. The antibody used for this western blot is the same used for the ChIP protocol and, due to the different nature of these two techniques (e.g., denaturation of the proteins during the SDS-Page), it cannot be excluded that the results obtained with the western blot are not a clear representation of the ones obtained by ChIP. Indeed, a different antibody against CTCF or a tag-based pull down of tagged CTCF should have been used as an independent test to rule out this possibility.

The following step of the protocol consists in the detection of the DNA fragments bound to the immunoprecipitated proteins. This can be achieved by qPCR, therefore, primers able to recognize and amplify the region in chromosome 9 (Chr9:89721201-89723471) where CTCF binding site was previously predicted to bind (*MIR184*-peak1) were designed. In addition, I have also generated primers for a region distant >5 Kb from *MIR184*-peak1 where no ATAC-seq peaks were detected that I used as the negative control (*MIR184*-NEG, chr9:89717577-89719847). A schematic representation of the regions amplified by the primers can be found in Figure 3.6A.

The optimized protocol was then performed on 400 islets isolated from C57BL/6 mice and the results are shown in Figure 3.6 B. Using this protocol, I was able to detect a clear binding of CTCF to *MIR184*-peak1 (P = 0.022) in mouse islets whereas no DNA corresponding to a region >5 Kb upstream from *MIR184*-peak1 used as our negative control (*MIR184*-NEG) was enriched in CTCF-precipitates. The CTCF ChIP-qPCR data are represented as % of input and the value shown in Figure 3.6 B are in the range of 0.05%. This data indicates that, although CTCF binds to the region of open chromatin 25Kb upstream *MIR184* (peak1) in mouse islets, the amount of DNA recovered after immunoprecipitation is lower than what is normally expected from this type of experiments (~0.5-1 %). This might be due to a low efficiency of the cross-link step between the proteins and the DNA

target or a lower pull down of CTCF during the immunoprecipitation step. However, it is also possible that CTCF binds to *MIR184*-peak1 with low efficiency under normal conditions and this binding can increase when the islets are subject to different stimuli.
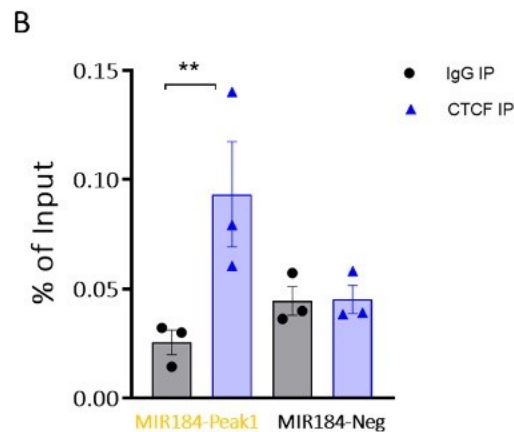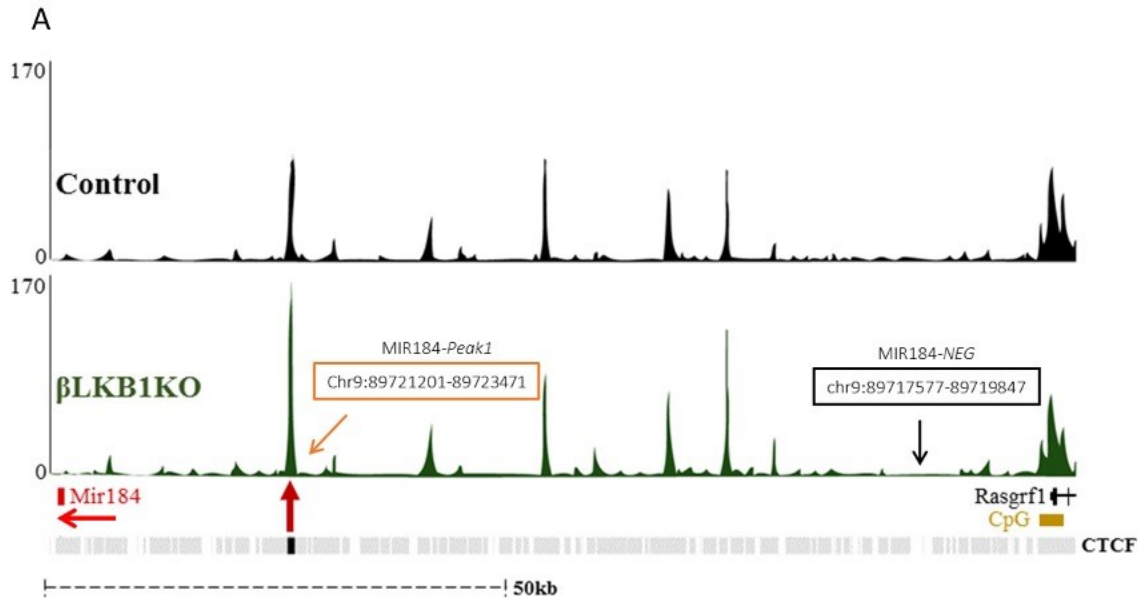
Figure 3.6: CTCF binds *MIR184*-Peak1 in MIN6 cells.
A. Schematic representation of the regions recognized by the primers designed for CTCF ChIP-qPCR. In orange the predicted CTCF binding site (*MIR184*-Peak1) and in black the region used as negative control (*MIR184*-NEG).
B: qPCR results using primers specific for the predicted CTCF binding site (*MIR184*-peak1) and negative primers (*MIR184*-Neg). Values are represented as % of input. Black dots represent samples immunoprecipitated (IP) using IgG antibody, blue triangle samples immunoprecipitated (IP) using CTCF antibody. Each dot represents an independent experiment performed using 400islets/experiment (n=3). Data are expressed as mean ± SEM of 3 independent experiments. ** p<0.01. P values were determined by one-way ANOVA with Sidak´s multiple comparisons test, with a single pooled variance.

3.2.4 Effect of LKB1 depletion and/or ketogenic diet in CTCF binding to *MIR184*

To determine whether AMPK/LKB1 promotes miR-184 expression by limiting CTCF binding to *MIR184*-peak1, I performed ChIP-qPCR in LKB1KO vs control mouse islets. Our data show (Figure 3.7 A) that binding of CTCF to *MIR184*-Peak1 is comparable between LKB1KO and control islets.

Finally, to determine whether differential binding of CTCF to *MIR184*-Peak1 was responsible for the effect of a ketogenic diet on miR-184 expression, I performed ChIP-qPCR in islets of mice fed with a keto or a control chow diet for 28 days. As shown in Figure 3.7 B, results demonstrate that CTCF biding to *MIR184*-Peak1 is not affected by the diet.

All together, these results suggest that the glucose-dependent regulation of miR-184 mediated by AMPK occurs through mechanisms that are independent from CTCF binding to the region of open chromatin 25kb upstream *MIR184*.
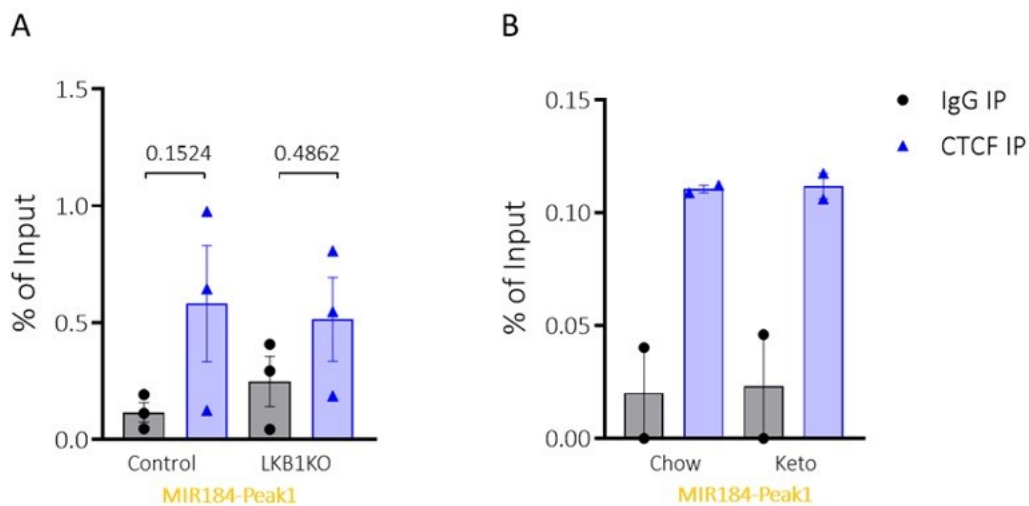
Figure 3.7: CTCF binding to MIR184-Peak1 is not affected by depletion of LKB1 or ketogenic diet.
A. qPCR results using primers specific for the predicted CTCF binding site (MIR184-peak1) in islets from LKB1KO vs Control mice
B. qPCR results using primers specific for the predicted CTCF binding site (MIR184-peak1) in islets of mice fed a Chow vs a Ketogenic (Keto) diet.
Values are represented as % of input. Black dots represent samples immunoprecipitated (IP) using IgG antibody, blue triangles samples immunoprecipitated (IP) using CTCF antibody. Each dot represents an independent experiment performed using 500 islets/experiment. Data are expressed as mean ± SEM of 3 (A) and 2 (B) independent experiments. P values were determined by one-way ANOVA with Sidak´s multiple comparisons test, with a single pooled variance (A).

3.2.5 Identification of putative miR-184 promoter and regulatory sequences

During miRNA biogenesis, pri-miRNA processing happens at such a high rate that conventional mRNA-Seq experiments fail to capture most pri-miRNAs. Consequently, conventional mRNA-Seq experiments cannot be used to identify TSSs of miRNAs directly.

Here, I have taken advantage of several published datasets of ChIP-seq histone modifications (H3K4me3, H3K4me1 and H3K27ac) [512, 513] and combined them with SLIC-CAGE data obtained by our collaborators in mouse islets to identify *MIR184* TSS in islets. Super-Low Input Carrier-Cap Analysis of Gene Expression (SLIC-CAGE) is a methodology used to obtain a genome-wide quantitative mapping of capped 5'ends of RNA polymerase II transcripts. This method can precisely and efficiently profile transcriptional start sites with single nucleotide resolution [514].

Results obtained applying this computational analysis are shown in Figure 3.8. Open chromatin regions containing H3K4me3 were identified as promoters whereas those containing H3K27Ac and/or H3K4me1 in the absence of H3K4me3 were defined as enhancers. Additionally, to identify promoters actively transcribed in mouse islets (SLIC-CAGE/ H3K4me3) a SLIC-CAGE dataset was used. Transcription starting sites transcribed from the negative (minus) strand are reported in red in Figure 3.8, whereas green ones are transcribed from the positive (plus) strand. As shown in Figure 3.8, *MIR184*-peak1 was confirmed as a promoter (H3K4me3+) although active transcription initiation was not detected in this region (Yellow rectangle). Additionally, the region 78 kb upstream *MIR184* which had been previously proposed as a potential miR-184 promoter [495], is in fact an enhancer (blue rectangle).

Interestingly, using this approach I was able to identify two actively transcribed promoters (SLIC-CAGE/H3K4me3) in the antisense strand (TSS positions chr9:89909804 and chr9:89923188-89923354, delimited by a green rectangle in Figure 3.8) ~107-121 Kb upstream of *MIR184* (chr9:89802260-89802328) and the most distal (chr9:89923188-89923354) was found located within chromatin regions significantly less accessible in LKB1KO islets.
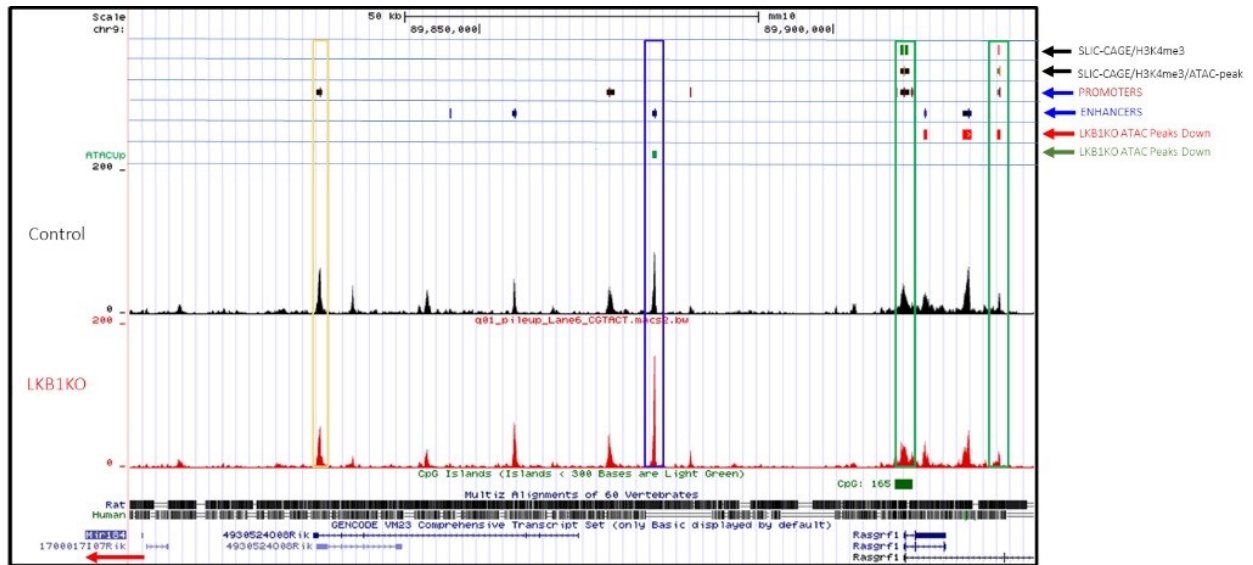
Figure 3.8: ATAC-seq enrichment profiles of MIR184 locus integrated with histone modifications and SLIC-CAGE.
Genomic region ~107-121 Kb up-stream *MIR184* is associated with less chromatin accessibility in LKB1KO (Red) vs. control (Black) islets. SLIC-CAGE/H3K4me3 track shows actively transcribed TSS from the negative strand (Red) and positive strand (Green). SLIC-CAGE/H3K4me3/ATAC-peak track shows chromatin regions containing SLIC-CAGE and H3K4me3 modifications and significantly differentially expressed ATAC-peak in LKB1KO (Red) vs. control (Black) islets. Chromatin regions containing only H3K4me1 modifications are shown as PROMOTERS (Dark red) whereas those containing H3K27Ac and/or H3K4me1 in the absence of H3K4me3 are defined as ENHANCERS (Blue). ATAC-seq significantly up-regulated peaks are shown in green (ATAC-Up) whereas the significantly down regulated are in red (ATAC-Down). *MIR184-peak1* does not show active transcription initiation (Yellow rectangle); Significantly up-regulated ATAC-peak ~75 Kb upstream of *MIR184* is an ENHANCER (Blue rectangle); Two genomic regions positioned ~107-121 Kb upstream of *MIR184* (green rectangle) transcribe for non-coding protein gene from negative strand, however only the most distal is associated with significantly differentially expressed ATAC-peak.

3.3 Discussion

Reduction of miR-184 is required for adequate β-cell proliferation during compensatory expansion in pregnancy and obesity [397, 401]. MiR-184 expression is also downregulated in prediabetic and diabetic mice [397] and concomitantly negatively regulated by glucose [402, 458]. Our group showed that this regulation is mediated by AMPK [458]. In fact, miR-184 was found downregulated upon AMPK and LKB1 knockout in mice and upregulated in both MIN6 cells and human islets where AMPK was pharmacologically activated [458]. Importantly, this regulation occurs at transcriptional level since the expression of both mature miR-184 and the precursor form (pri-miR-184) were reduced upon depletion of AMPK [458]. ATAC-seq data in LKB1KO vs control mice indicate that the most proximal regions upstream of the miR-184 predicted promoter is characterized by an increase in chromatin accessibility [458]. Therefore, we hypothesised that AMPK might regulate miR-184 transcription through a transcriptional inhibitor.

Using ENCODE ChIP-seq datasets I found that CTCF, an important epigenetic regulator with insulator functions [500, 501], binds this region in liver [497]. CTCF is an evolutionarily conserved and ubiquitously expressed zinc finger protein [515] essential for controlling many epigenetic regulatory gene expression events [498]. In pancreatic β-cell, it is positively regulated by glucose and insulin [508] and has been shown to function as a molecular mediator between insulin-induced upstream ERK signalling and down-regulation of *Pax6* contributing to β-cell proliferation stimulated by both glucose and insulin [508].

We hypothesised that AMPK might be the mediator of glucose-dependent CTCF regulation and, moreover, by inhibiting this important TF, AMPK could also prevent its binding to miR-184 promoter and promote its transcription.

According to RNA-seq data from AMPKdKO and LKB1KO mouse islets [247], regulation of CTCF by AMPK at transcriptional level is quite improbable since mRNA levels of CTCF are not altered by the absence of these enzymes in β-cells. Therefore, I investigated if this regulation occurred at the protein level. The glucose-mediated regulation of CTCF in β-cells was confirmed, but, contrary to what we previously hypothesised, also AMPK resulted to be a positive regulator of CTCF protein

level since the activation resulted in increased level of CTCF. These results suggest that glucose regulates CTCF expression independently of AMPK and therefore reduced expression of miR-184 in MIN6 cells upon AMPK activation is not mediated by a reduction in CTCF protein levels.

Increased levels of both glucose and insulin have been shown to stimulate the expression of CTCF and promote β-cell proliferation [508]. In these cells in fact CTCF acts as a molecular mediator between insulin-induced upstream ERK signalling and *Pax6* expression and may contribute to the regulation of β-cell survival and proliferation.

On the contrary, AMPK inhibits proliferation of β-cell [516] and is inactivated by glucose. Therefore, by activating CTCF and inhibiting AMPK, glucose might lead mature β-cell towards a more proliferative status.

AMPK has also been shown to play an important role in the maintenance of β-cell function and identity. It has been suggested that the reciprocal opposition between mTORC1 and AMPK determines whether β-cells adopt a more proliferative, immature phenotype, or the mature, highly insulin-responsive phenotype essential for postnatal life [535]. Indeed, β-cell specific loss of AMPK in mouse increased the expression of subsets of hepatic and neuronal genes and up-regulated "disallowed" genes indicating that AMPK maintains β-cell identity by suppressing alternate pathways [247]. While the mechanisms behind this regulation are unknown, DNA binding motifs of CTCF have interestingly been recently associated with open chromatin regions of genes that are significantly correlated to mature β-cell function and identity [536]. Therefore, it is possible that, by positively regulating CTCF expression, AMPK might promote the maintenance of β-cell function and identity. Indeed, selective targeting of these enzymes may provide a new approach to maintaining β-cell function in some forms of diabetes.

Phosphorylation of CTCF has been previously shown to impair the binding to DNA target sequences [517, 518]. AMPK is a kinase able to phosphorylates a variety of proteins, however AMPK phosphorylation of CTCF has never been reported before. An analysis performed using Scansite4 tool [519] suggested that AMPK may phosphorylate serine 461 in CTCF sequence through a basophilic serine/threonine kinase group. As mentioned before, since a commercial antibody able to recognise this specific phosphorylated site of CTCF is not available, Phos-Tag™ SDS-PAGE was used

to investigate if AMPK phosphorylates CTCF as predicted. The method was able to detect phosphorylation of AMPK upon its activation, but the antibody against CTCF detected a broad smear in both control and MIN6 cells treated with AMPK activators, thus suggesting that CTCF phosphorylation is not changed upon AMPK activation. This experiment was repeated twice, and in both cases similar results were obtained. However, as described earlier, the experiment is missing key controls therefore it cannot be excluded that the result obtained could be due to a limitation of the technique itself: the SDS-PAGE performed with this method is very sensitive to variations of protein concentration, quantity of Phos-TagTM compound added to the protein sample and salt concentration in the buffers used for the SDS-PAGE therefore more optimizations and, as discussed previously, additional controls are required. Indeed, alternative, more straightforward methods such as in vitro kinase activity assays or mass spectrometry could have also been used instead to assess CTCF phosphorylation.

It is also possible that other types of post-translation modifications that are independent from AMPK and produce a broad smear on an SDS-PAGE might have been detected by Phos-Tag[TM]. It has been demonstrated that different external stimuli, including high glucose concentrations, can mediate post-translation modifications [520]. For this Phos-Tag[TM] SDS-PAGE experiment, for example, the samples were collected from MIN6 cells grown at 25mM glucose, therefore it cannot be excluded that the high glucose concentration might be the mediator of these modifications. For these reasons, the addition of a control sample of cells cultured at low glucose concentrations might have been helpful to elucidate this aspect.

Our data from CTCF ChIP-qPCR demonstrated that this TF is able to bind the region ~25 kb upstream of *MIR184* (*MIR184*-Peak1) also in mouse islets, however the binding between LKB1KO and control islets was comparable. This implies that CTCF binding is not responsible for the regulation of miR-184 expression mediated by AMPK in mouse islets. Similarly, ChIP-qPCR in islets of mice fed with a keto or a control chow diet for 28 days showed that the binding of CTCF to *MIR184*-Peak1 was not affected by the diet, indicating that glucose-dependent regulation of miR-184 mediated by AMPK occurs through mechanisms that are independent from CTCF binding.

All these data together indicate that the regulation of miR-184 mediated by AMPK might occur through different mechanisms that do not involve CTCF.

As described before, the exact location of miR-184 promoter has not been precisely annotated. However, Marson et al., who identified the primary transcript of miR-184 by RT–PCR analysis using several primers around the mature miR-184 coding sequence, suggested that, in mouse embryonic stem cells, *MIR184* TSS might be located up to ~78 kb upstream of the pre-miR-184 loop [495]. Nomura et al. found that, in mouse brain, the binding of methyl CpG-binding protein 2 (MeCP2) to a region ~73 kb upstream of *MIR184* is responsible for a repression of miR-184 expression, whereas in neural stem cells [521] Liu et al. identified several CpG-rich sequence in the genomic region immediately surrounding miR-184 and, when they screened locations from -5 kb upstream to +2 kb downstream of the miR-184 gene, identified an enrichment for Methyl-CpG binding protein 1 (MBD1) binding to the region -4 kb upstream and +1 downstream *MIR184 [521].*

Indeed, Martinez-Sanchez et al. [458] suggested that, in addition to CTCF, also other transcription factors such as BHLHE40, ETS1, GCn5 and p300 can bind the DNA located within the two open chromatin peaks upstream of *MIR184.* However, the data reported in that work were collected from available ChIP-seq datasets produced in CH12 cells (Figure 3.1), a mouse lymphoma cell which transcription might be regulated differently from β-cells.

Moreover, according to our analysis based on the SLIC-CAGE dataset, none of these regions appear to be a promoter region in mouse islets, indicating that, even if these TFs might bind the same region in pancreatic β-cells, they would probably not be involved in the regulation of miR-184 transcription.

Using SLIC-CAGE analysis, instead, I was able to identify two other actively transcribed promoter regions from the negative strand (positions chr9:89909804 and chr9:89923188-89923354) that are located ~107-121 Kb from *MIR184* (chr9:89802260-89802328) within chromatin regions significantly less accessible in LKB1KO islets (chr9:89923188-89923354). Importantly, a more detailed analysis that included only TSS associated to non-coding protein genes suggested that the most distal region could act as miR-184 TSS. Indeed, additional studies are required to validate this newly identified region as miR-184 promoter.

For example, luciferase reporter assay can be used to understand which role this newly identified region plays in miR-184 transcription. The region of interest can be amplified by PCR using specific primers and cloned upstream of the luciferase construct present in pGL3 vector. Since the vector lacks eukaryotic promoter and enhancer sequences, if the identified putative regulatory sequence introduced into the vector shows higher luciferase activity when compared to the empty vector it might indicate that it is indeed a promoter. Moreover, the effect of AMPK activation on the luciferase activity in these vectors could also be assessed to determine if transcription is regulated by AMPK. Finally, CRISPR/Cas9 system could be used to generate a β-cell line where the computationally identified *MIR184* promoter is knocked down and assess if the deletion of this region is responsible for a reduction in miR-184 expression.

In order to identify which TF might be responsible for this regulation mediated by AMPK different strategies can be applied. For example, a computational analysis using available tools that search for TF binding site motifs, such as PROMO [522] or CIIDER [523], can be performed to identify potential TF able to bind the miR-184 promoter. It also possible to search for publicly available ChIP-seq datasets in which specific TFs binding have already been identified in pancreatic islets or other tissues and validate them as it has been done for CTCF.

Although we think that the regulation of miR-184 mediated by AMPK occurs at transcriptional level, it is still possible that other type of regulations might be involved. In the past few years, it has emerged that the Microprocessor, the complex responsible for the pri-miRNA processing, could be involved in this regulation. As described before, the Microprocessor is mainly characterized by Drosha, a protein that functions as a core complex together with its essential cofactor DGCR8. However, additional RNA-associated proteins, such as helicases or heterogeneous nuclear ribonucleoproteins (hnRNPs), have also been associated to the complex [271, 273] and have recently been shown to negatively [524] or positively [525] regulate the binding of the pri-miRNA to the Microprocessor.

Interestingly, AMPK activation mediated by metformin has been shown to induce AMPKα2 translocation into the nucleus, where it can directly phosphorylate hnRNPs and modulate metformin-mediated glucose uptake in myoblasts [526]. While such mechanism has not been identified in β-

cell yet, understanding how AMPK might be involved in this regulation could give new insights on the mechanism by which it regulates miRNAs.

Chapter IV: mir-125b transcription regulation

4.1 Introduction

MiR-125b is ubiquitously expressed and, although its function in β-cells remains unknown, it controls proliferation, apoptosis and differentiation in other cell types [470]. Interestingly, high levels of circulating miR-125b have been associated with hyperglycaemia (HbA1c) in prediabetic [469], T2D [469] and T1D subjects [424], suggesting miR-125b as a biomarker or contributor to the development of diabetes.

Importantly, our lab found that miR-125b expression is induced by glucose in both mouse and human pancreatic islets and strongly reduced in islets from mice fed a ketogenic diet. Moreover, according to our data, AMPK plays a key role in the regulation of miR-125b expression in β-cells and acts as a mediator of the glucose regulation. In fact, the effects of glucose on miR-125b expression both *in vitro* and *in* vivo were abolished upon AMPK (AMPKdKO) and its upstream kinase LKB1 (LKB1KO) deletion (Figure 4.1).
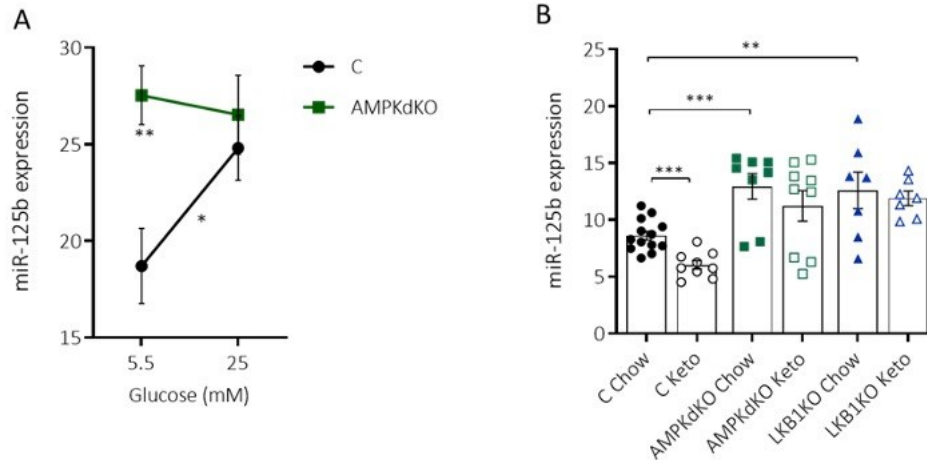
Figure 4.1: AMPK mediates glucose-dependent regulation of miR-125b.
A. miR-125b expression measured by RT-qPCR in isolated islets from AMPKdKO and control (C) mice cultured with 5.5 or 25 mM glucose for 48 h (n=6-7 animals)
B. miR-125b expression measured by RT-qPCR in AMPKdKO, LKB1KO and control (C) male and female mice fed chow or a ketogenic (Keto) diet for 28 days. Each dot represents an animal (n=7-10). Data are expressed as mean ± SEM relative to C levels. Values were normalised against the endogenous control Let-7. Data are expressed as mean ± SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

MiR-125b is highly conserved among species with mammals, nematodes and vertebrate sharing the same seed region (Figure 4.2 A) [527]. In humans, as well as in mouse, two different loci encode for miR-125b, *MIR125B-1* and *MIR125B-2* (chromosome 11 and 21 for humans and chromosome 9 and 16 for mouse, respectively). Interestingly, although the two different loci transcribe the same mature miRNA, miR-125b is clustered with different miRNAs in both loci (Figure 4.2 B). While *MIR125B-2* clusters with *MIRN99A* and *MIRLET7C* genes and is located ~50 kb downstream of them, *MIR125B-1* is in a cluster with *MIRLET7A2* and *MIR100,* two miRNAs (let-7a-2 and miR-100) that differ by only one nucleotide outside of the seed region from their family members, let-7c and miR-99a. Moreover, each locus has been annotated with several (2-3) TSS meaning that these loci have the potential to generate different pri-miR-125b on a cell-specific manner [528].
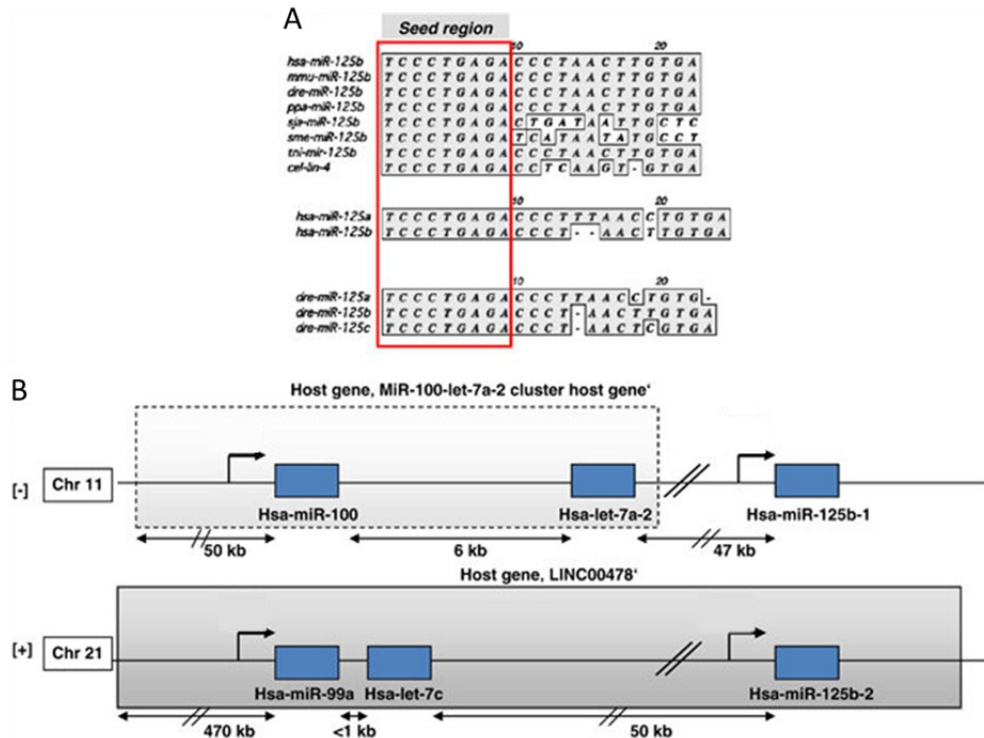
Figure 4.2: miR-125b seed region and loci.
A. The seed region of homologs and orthologs of miR-125b is highly conserved throughout different species.
B. Genomic organization of the miR-cluster in human. Figure shows chromosome 11 and 21 containing miR-125b with its host genes together with potential promoter regions (Black arrows). Modified from Shaham et al. [527]

For example, in prostate cancer cells miR-125b expression is directly stimulated by androgens through the recruitment of the androgen receptor to the 5' DNA region of *MIR125B-2* [529]. In human myeloid leukaemia, the homeobox transcription factor CDX2 positively regulates miR-125b expression through its binding to two regions 1185 and 170 bp upstream *MIR125B-1* promoter [474]. Additionally, nuclear factor kappa beta (NF-kb) regulates miR-125b expression during immune response [530-532]. Zhou et al. demonstrated that in human cholangiocytes the NF-kB p65 subunit binds to a region 1kb upstream of the putative *MIR125B-1* promoter to increase miR-125b expression and promote the miRNA-dependent regulation of epithelial anti-microbial defence [532].

MiR-125b transcription hasn't been characterized in β-cells. Our unpublished data suggest that, in EndoC-βH1 human pancreatic β-cells, *MIR125B-2* region is more transcriptionally active than *MIR125B-1* since CRISPR-mediated mutation of *MIR125B-2* leads to an important decrease in mature miR-125b, while mutation of *MIR125B-1* had a very limited effect (Cheung and Martinez-Sanchez, unpublished). Moreover, ATAC-seq data in mouse islets shows maximal chromatin accessibility around the most proximal *MIR125B-2* promoter, whereas chromatin accessibility seems limited in *MIR125B-1* (Figure 4.3 A and B). More importantly, unpublished differential analysis of ATAC-seq data between control and LKB1KO islets identified significantly higher chromatin accessibility in LKB1KO islets vs control in the proximal *MIR125B-2* promoter, suggesting that TFs binding to this region might be responsible for higher miR-125b expression in LKB1KO islets (Figure 4.3B).

In 2018, Ottaviani et al. [533] demonstrated that, in PDAC cells, TGF-β induces the transcription of miR-125b through SMAD2/3 binding to *MIR125B-1*. Moreover, David et al. [534] shows that, in the same cells, the transcription factor SMAD2/3 binds also the *MIR125B-2* locus (Figure 4.3B), suggesting that the TGF-β pathway could have a role in the regulation of transcription in this region.

Figure 4.3: ATAC-seq enrichment profiles of *MIR125B-1* and *2 loci* integrated with histone modifications and SMAD2/3 ChIP-seq dataset.
A. Genomic region 10 kb upstream of MIR125B-1 shows chromatin accessibility in LKB1KO (Red) vs. control (Black) islets.
B. Genomic region 10 kb upstream of MIR125B-2 shows an enriched ATAC-seq peak in LKB1KO (Red) vs. control (black) islets. Enrichment of SMAD2/3 TF within the ATAC-seq differential peak is shown by SMAD2/3 ChIP-seq profiles in PDAC cells (Orange).
When present, chromatin regions containing H3K4me3 modifications were identified as PROMOTERS (Dark red) whereas those containing H3K27Ac and/or H3K4me1 in the absence of H3K4me3 were defined as ENHANCERS (Blue). ATAC-seq significantly up-regulated peaks are shown in green (ATACUp) whereas the significantly down regulated are in red (ATACDown).
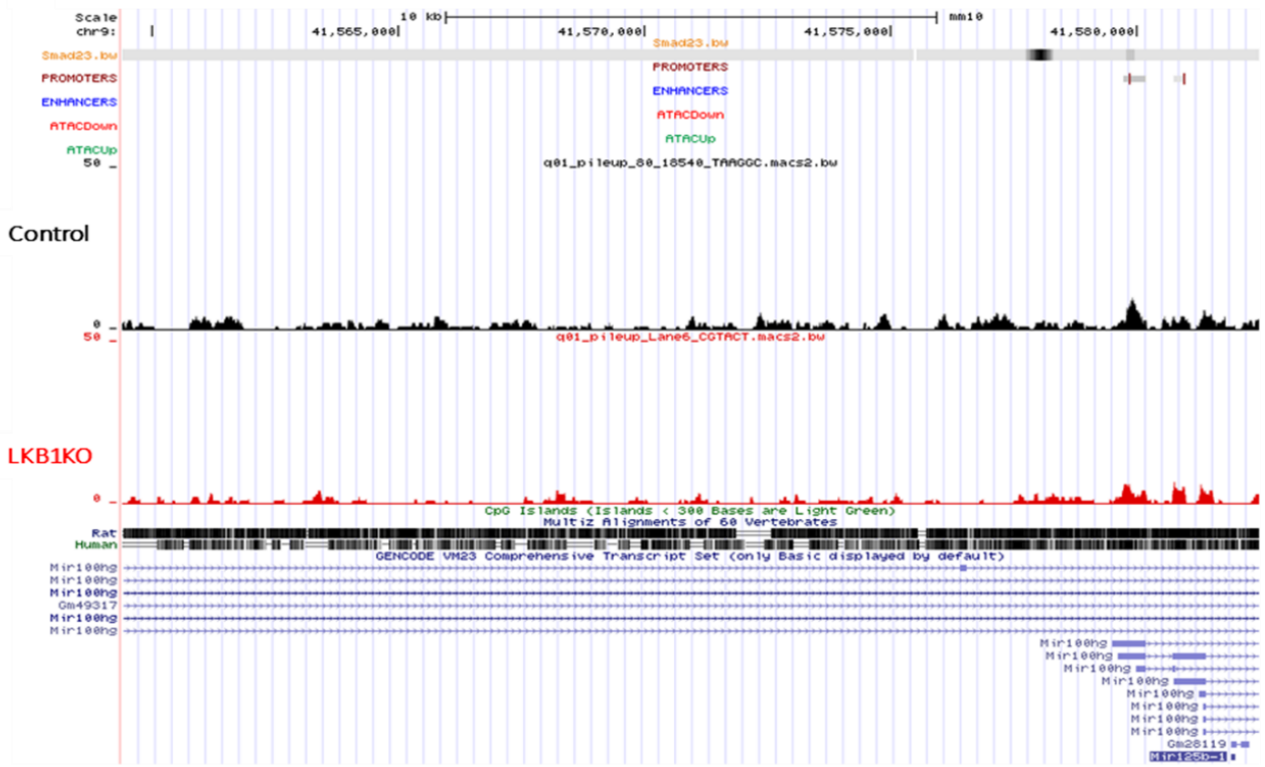
In mature pancreatic β-cells, TGF-β activation plays an important role in cell proliferation and survival [535] and has been reported to stimulate insulin secretion [536], insulin transcription, and impairs islet function [537]. AMPK has been shown to regulate the TGF-β pathway in multiple cell types [538-540]. For example, Lin et al. [538] showed that AMPK can inhibit TGF-β-induced SMAD2/3 phosphorylation in cancer cells and therefore negatively regulate cancer cell migration and EMT. It has also been demonstrated that, in hepatic stellate cells (HSCs), treatment with the AMPK agonist AICAR inhibited the interaction between SMAD3 and the transcriptional coactivator p300, whereas the interaction between AMPK and p300 was increased [539]. In addition, the authors saw a reduction of SMAD3 acetylation in HSCs treated with AICAR [539]. It has also been shown that AMPK can regulate TGF-β signalling by altering SMAD3 capacity to bind DNA. For example, a study from 2008 [540] showed that AMPK did not reduce TGF-β-stimulated SMAD3 phosphorylation and nuclear translocation but instead inhibited the binding of SMAD3 to the SMAD3-binding cis-elements during myofibroblast trans-differentiation.

However, it is not known whether AMPK regulates SMAD2/3 activity in pancreatic β-cells. Kone et al. [247] have previously found that *Smad2* and *Smad3* mRNAs were upregulated in specific β-cell AMPKdKO islets versus control (albeit only *Smad2* was significant), whereas LKB1KO islets showed the opposite: a decrease in *Smad2* and *Smad3* mRNA levels, with only *Smad3* being significant [247]. We hypothesize that SMAD2/3 may activate miR-125b transcription in β-cells and be, at least partially, responsible for the increase in miR-125b expression observed in AMPKdKO and LKB1KO islets. Thus, the focus of this chapter is the focus of this chapter is (1) to investigate whether SMAD2/3 promotes miR-125b expression in β-cells and (2) if this regulation is inhibited by AMPK (Figure 4.4).
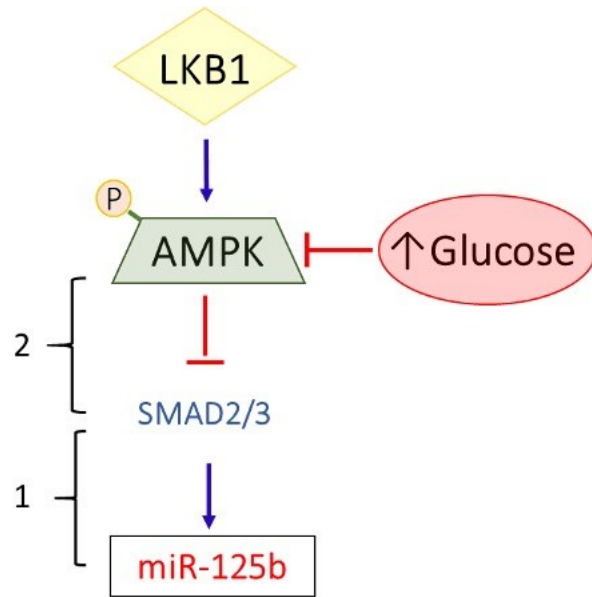
Figure 4.4: Simplified schematic representation of the hypothesis and aims of the chapter.

4.2 Results

4.2.1 Effect of AMPKdKO on pri-miR-125b expression in mouse islets

To investigate whether AMPK regulates miR-125b expression at the transcriptional level we decided to measure the expression of miR-125 precursor in AMPKdKO islets. As previously described, miRNA genes are transcribed as large poly-adenylated primary precursors, called pri-miRNAs, which are quickly processed to produce a first intermediate of ~70 nucleotides (pre-miRNA) and then the 22-25 nt mature miRNA. Full-length miR-125b pri-miRNA(s) hasn't been annotated in β-cells. Nevertheless, the ~70 nucleotide hairpin structure encoded by *MIR125B-1* and *MIR125B-2* present in both primary and precursor miR-125b is well defined. Thus, to only detect pri-miR-125b, qPCR primers complementary to this part of the sequence were designed and reverse transcription in the presence of an anchored oligo dT primer was performed. Interestingly, I found that, along with the mature form of miR-125b, also pri-miR-125b-1 expression was increased in AMPKdKO islets when compared to control (Figure 4.5), whereas pri-miR-125b-2 was undetected. Notably, low/undetectable levels of a miRNA precursor can represent both low transcription and/or fast processing.
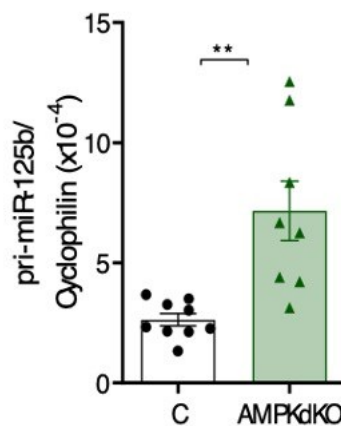
Figure 4.5: Pri-miR-125b expression is upregulated in AMPKdKO mouse islets.
Pri-miR-125b expression measured by RT-qPCR in isolated islets from AMPKdKO and littermate control (C). Each dot represents an independent animal (n=8-9). Data are expressed as mean ± SEM relative to control levels. Values were normalised against the endogenous control Cyclophilin. ** p<0.01. P values were determined by two-tailed unpaired t-test

4.2.2 Effect of TGF-β and SMAD2/3 on miR-125b expression

As mentioned above, TGF-β activates miR-125b transcription *via* SMAD2/3 in PDAC [533] and AMPK has been previously shown to inhibit SMAD2/3 activation by TGF-β in other cell types [538-540]. This led us to hypothesize that AMPK inhibits miR-125b expression by limiting SMAD2/3 transcriptional activity. Thus, I decided to investigate whether (1) TGF-β and SMAD2/3 regulate miR-125b expression and whether (2) this regulation is prevented by AMPK activation.

In order to investigate the effect of TGF-β pathway activation on miR-125b expression I first treated human islets with 30 ng/ml of TGF-β for 5 days and measured miR-125b expression by qPCR. The results illustrated in Figure 4.6 show that the treatment positively regulated miR-125b expression indicating that TGF-β activates miR-125b in human islets. Unfortunately, due to the limited availability of human islets samples, I was not able to perform a time-course treatment to identify precisely the time needed by TGF-β to induce miR-125b expression. For this reason, it cannot be excluded that short-term activation of TGF-β might have a stronger effect on miR-125b expression.

In order to further investigate if TGF- β signalling pathway regulates miR-125b expression we decided to study the role played by SMAD2/3, important proteins involved in the signalling pathways downstream of the TGF-β receptors [541]. As mentioned before, in mouse islets *MIR125B-2* locus presents high chromatin accessibility, markers associated with a promoter activity (H3K4me3) and, in PDAC, SMAD2/3 binds this region (Figure 4.3B). If SMAD2/3 is responsible for the regulation of miR-125b expression in β-cells through the binding to *MIR125B-2* locus, we expect this TF to act as an activator.
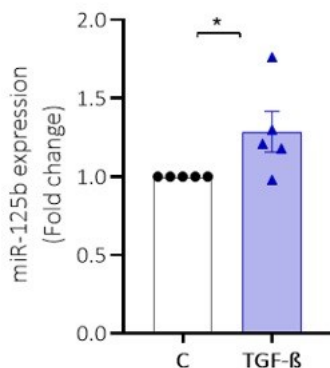
Figure 4.6: TGF-β positively regulates miR-125b expression in human islets.
miR-125b expression measured by RT-qPCR in isolated islets from human donors were treated with 30 ng/ml of TGF-β for 5 days (TGF-β) vs control (C) mice. Each dot represents an independent animal (n=5). Data are expressed as mean ± SEM relative to control levels. Values were normalised against the endogenous control let-7. * $p<0.05$. P values were determined by two-tailed unpaired t-test on the log of the fold change values.

To investigate this hypothesis, I performed ChIP-qPCR with an antibody against SMAD2/3 and primers specific for that region. Unfortunately, using mouse islets I was unable to detect immune-precipitated SMAD2/3 by Western Blot. In order to be able to use a higher amount of starting material, I tried to optimise ChIP-qPCR in MIN6 cells (Figure 4.7).
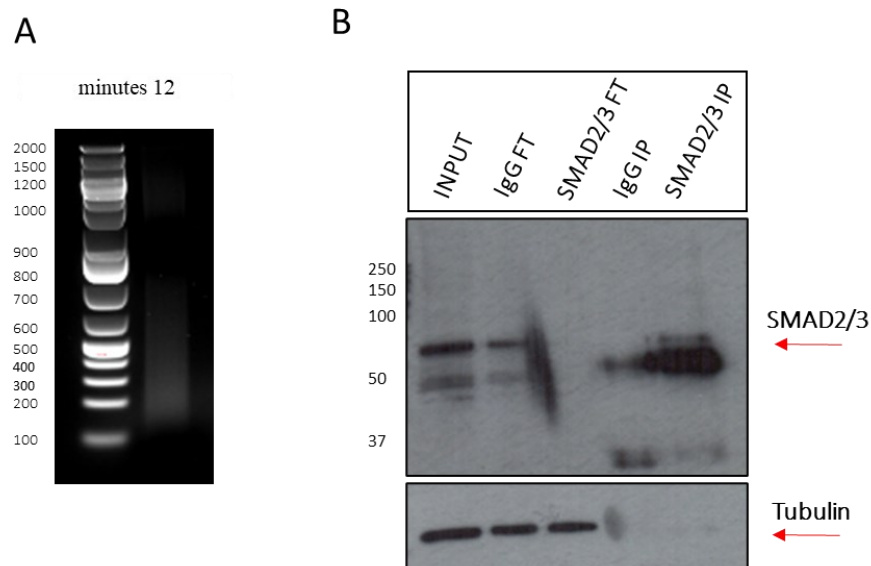
Figure 4.7: SMAD2/3 ChiP-qPCR optimization in MIN6 cells.
A. Agarose gel electrophoresis assessment of chromatin fragmentation for MIN6 cells. The red arrow indicates the optimal DNA fragments average size (200nt) obtained with DNA sonicated for 12 minutes using Covaris 220.
B. Western Blot analysis obtained with the optimized protocol in MIN6 cells. SMAD2 protein (55 kDa) levels in samples where immunoprecipitation was performed with an anti-SMAD2/3 antibody (SMAD2/3 IP) or an IgG control (IgG IP). Input: 5% of the total lysate submitted to immunoprecipitation (IP); FT (Flow-through): 5% of the extract recovered after IP.

MIN6 cells were fixed directly in tissue culture plates, and I increased the sonication time to 16 minutes to obtain DNA fragments of the optimal length of 200-500 bp (Figure 4.7 A). For the immunoprecipitation step, I used 5ug IgG and SMAD2/3 antibody and, as shown in Figure 4.7B, SMAD2/3 was specifically immunoprecipitated only in the presence of SMAD2/3 antibody. I designed specific qPCR primers to amplify the region of interest (*MIR125B-2 peak*, chr16:77644425+77644490, Figure 4.8A) as well as a region ~3 Kb upstream that did not have any detectable ATAC-seq peaks to use as a negative control (*MIR125B-2 NEG,*

chr16:77641316+77641382) (Figure 4.7 A). Unfortunately, in two independent experiments, I was unable to detect the DNA of interest in the SMAD2/3 ChIP sample (Figure 4.8 B), suggesting that SMAD2/3 does not bind to the proposed region in *MIR125B-2* under these conditions in MIN6 cells. Nevertheless, the approach could have been improved to obtain more solid results. For example, our SMAD2/3 ChIP-qPCR experiment is missing a positive control such as a DNA region were SMAD2/3 binding has been previously confirmed in β-cells. Unfortunately, no such region had at the time of our experiments been identified and reported in literature.
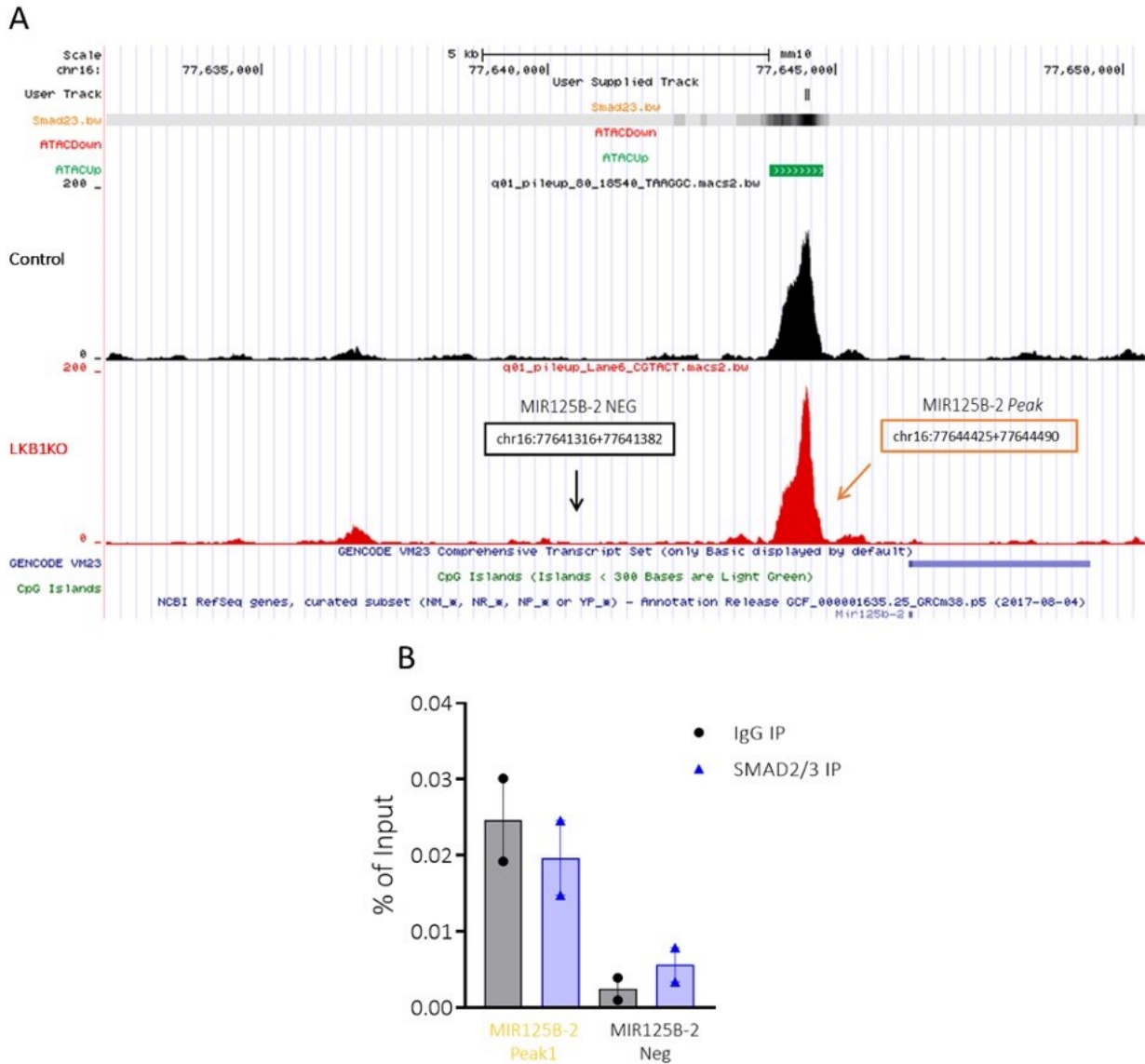
Figure 4.8: SMAD2/3 ChiP-qPCR is characterized by high background.
A: Schematic representation of the regions recognized by the primer designed for SMAD2/3 ChIP-qPCR. In orange the predicted SMAD2/3 binding site (MIR125B-2 Peak)[534] and in black the region used as negative control (MIR125B-2-NEG).
B: qPCR results using primers specific for the predicted SMAD2/3 binding site (MIR125B-2 Peak) and negative primers (MIR125B-Neg). Values are represented as % of input. Black dots represent samples immunoprecipitated (IP) using IgG antibody, blue triangles samples immunoprecipitated (IP) using SMAD2/3 antibody. Each dot represents an independent experiment performed using ~25x10[6] cells/experiment (n=2). Data are expressed as mean ± SEM of 3 independent experiments. P values were determined by one-way ANOVA with Sidak´s multiple comparisons test, with a single pooled variance.

In parallel, I attempted to determine whether SMAD2/3 (directly or indirectly) regulates miR-125b expression by measuring mature miR-125b following the silencing of both *Smad2* and Smad3 in MIN6 cells using 50nM of siRNAs. Data reported in Figure 4.8 A shows that I was able to achieve a significant decrease in *Smad2* expression upon *Smad2* siRNA transfection, whereas cells transfected with siRNAs targeting *Smad3* presented a down-regulation of *Smad3* mRNA (Figure 4.8 B). Interestingly, I also observed a decrease in *Smad3* mRNA levels in cells transfected with *Smad2* siRNAs, indicating that SMAD2 acts as a positive regulator of SMAD3 expression in MIN6 cells (Figure 4.8 B). Contrarily to what we expected, the expression of miR-125b was found significantly increased when *Smad3* was silenced, whereas the silencing of *Smad2* did not affect miR-125b expression (Figure 4.8 C). These results suggest that SMAD3 negatively regulates miR-125b expression in MIN6 cells.
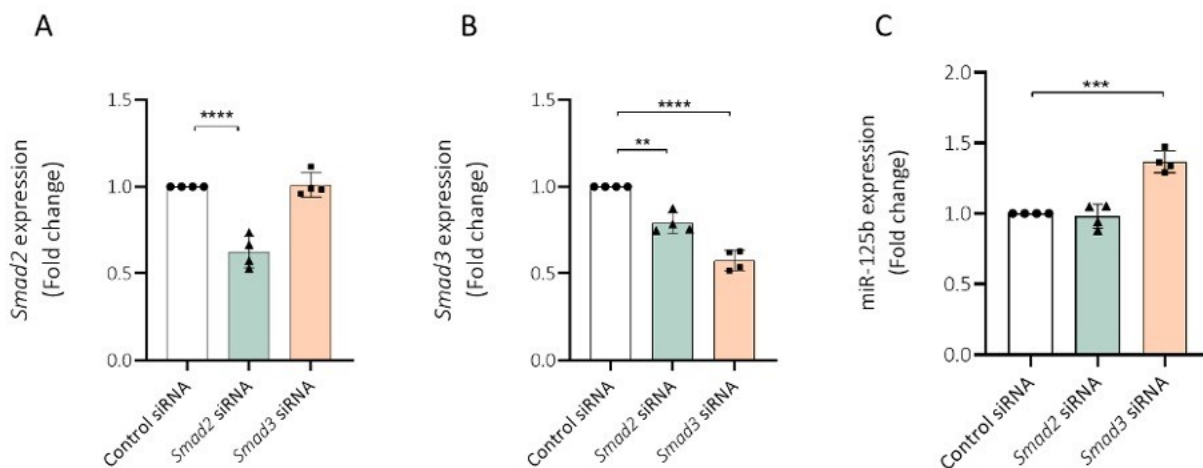
Figure 4.9: Silencing of *Smad3*, but not *Smad2*, increases miR-125b expression.
A. *Smad2* expression measured by qRT-PCR in MIN6 cells transfected with 50 nM siRNA against *Smad2, Smad3 or control for 48h.*
B. *Smad2* and *Smad3* expression measured by qRT-PCR in MIN6 cells transfected with 50 nM siRNA against *Smad2, Smad3* or control for 48h.
C. miR-125b expression measured by RT-qPCR in MIN6 cells transfected with *Smad2, Smad3* or control siRNAs.
Each dot represents an independent experiment (n=4). Data are expressed as mean ± SEM relative to Control siRNA levels. Values were normalised against the endogenous control Cyclophilin (A and B) and Let-7 (C). ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$. P values were determined by one-way ANOVA with Sidak's multiple comparisons test, with a single pooled variance on the log of the fold change values.

4.2.3 Effects on AMPK activation on SMAD2/3 and miR-125b expression

In order to determine whether (1) AMPK controls SMAD2/3 expression and (2) the effect of SMAD2/3 on miR-125b is modulated by AMPK, I transfected MIN6 cells with *Smad2* and *Smad3* siRNAs and, 24 hours later, activated AMPK with compounds C-13 and C-991 for an additional 24 hours [492, 493].

I first measured *Smad2* and *Smad3* mRNA levels by qPCR and observed that, while AMPK activation did not affect *Smad2* mRNA levels (Figure 4.9 A), cells transfected with non-targeting siRNAs and treated with the activators showed a small tendency (albeit not significant) to increased *Smad3* mRNA levels (Figure 4.9 B). As expected, I again detected a decrease in *Smad3* mRNA levels when *Smad2* siRNAs were introduced into the cells, however this effect disappeared upon AMPK activation. I then measured miR-125b expression in the same samples and as shown in Figure 4.9 C, I confirmed that, as expected, *Smad2* silencing did not affect miR-125b expression. Additionally, AMPK activation alone did not result in an alteration of miR-125b levels in MIN6 cells. Finally, *Smad3* siRNAs caused a small increase in miR-125b expression in both presence and absence of AMPK activators though only in the latter the results were statistically significant. Importantly, the silencing of *Smad3* induced by *Smad3* siRNAs in this set of experiments appears to be less effective (~30%) (Figure 4.10 B) compared to the one observed in the previous dataset (~50%) (Figure 4.9 C) and this might explain why the silencing of *Smad3* without the co-activation of AMPK was not sufficient to significantly up-regulate miR-125b expression.
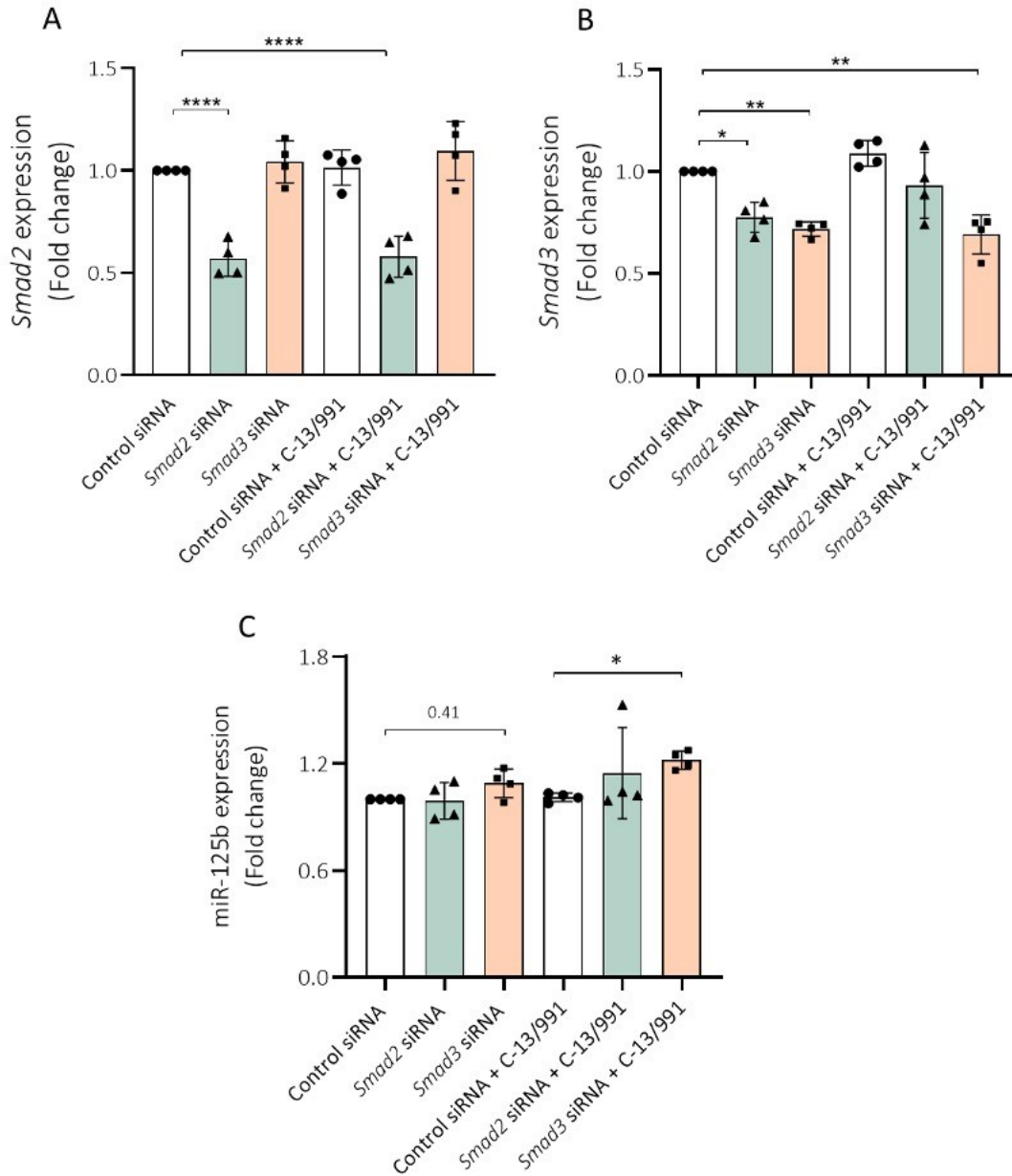
Figure 4.10: Activation of AMPK and silencing of *Smad3* increases miR-125b expression.

A. *Smad2* expression measured by qRT-PCR in MIN6 cells transfected with 50 nM siRNA against *Smad2, Smad3 or control for 48h.*

B. *Smad2* and *Smad3* expression measured by qRT-PCR in MIN6 cells transfected with 50 nM siRNA against *Smad2, Smad3* or control for 48h.

C. miR-125b expression measured by RT-qPCR in MIN6 cells transfected with *Smad2, Smad3* or control siRNAs.

Each dot represents an independent experiment (n=4). Data are expressed as mean ± SEM relative to control siRNA levels. Values were normalised against the endogenous control Cyclophilin (A and B) and Let-7 (C). * $p<0.05$, ** $p<0.01$, **** $p<0.0001$. P values were determined by two-way ANOVA with Sidak´s multiple comparisons test, with a single pooled variance on the log of the fold change values

4.3 Discussion

Our group unpublished data demonstrates that miR-125b is up-regulated upon specific knockout of AMPK in β-cells and its expression is induced by glucose. Interestingly, AMPK seems to play a key role in this regulation acting as a mediator of glucose regulation.

Mature miR-125b is transcribed from paralog genes located in two different loci, *MIR125B-1* and *MIR125B-2.* Remarkably, both loci have been annotated with several (2-3) TSS indicating that the precursor form of miR-125b (pri-miR-125b) could vary in a cell specific manner [528]. Indeed, it has been demonstrated that the transcription of miR-125b can occur predominantly from one or the other locus depending on the cell type [474, 529-531].

Our preliminary data indicated that the regulation of miR-125b expression mediated by AMPK could occur at the transcriptional level. Both EndoC-βH1 human pancreatic β-cells and mouse islets present a more transcriptionally active *MIR125B-2* region compared to *MIR125B-1,* thus we hypothesized that the transcriptional regulation mediated by AMPK might occur through the *MIR125B-2* locus.

Differential analysis of ATAC-seq data performed in LKB1KO islets vs control showed significantly higher chromatin accessibility in the proximity of *MIR125B-2* promoter in LKB1KO, suggesting that TFs binding to this region might be responsible for higher miR-125b expression in LKB1KO islets. Interestingly, ChIP-seq in PADC cells indicated that SMAD2/3 binds this region [534] suggesting that in β-cells the TGF-β pathway, and SMAD2/3 specifically, could regulate miR-125b transcription through *MIR125B-2* locus.

In order to confirm that the regulation of miR-125b mediated by AMPK was at the transcriptional level, first the expression of miR-125 precursor in AMPKdKO islets was measured using qPCR primers able to identify pri-miR-125b-1 and pri-miR-125b-2 independently. Surprisingly, I found that, while on one hand pri-miR-125b-1 expression was increased in AMPKdKO islets when compared to control, I failed to detect pri-miR-125b-2. It is known that different pri-miRNAs can exhibit different processing kinetics ranging from fast to slow processing [542], thus the fact that pri-miR-

125b-2 was undetected despite the ATAC-seq profile could be due to a rapid processing of miR-125b precursor from this locus.

Different signalling pathways and TFs regulate miR-125b in a cell-specific manner. In PDAC cells, for example, miR-125b transcription is induced by TGF-β through binding of SMAD2/3 to *MIR125B-1* [533]. I hypothesized that SMAD2/3 could act as a positive regulator of miR-125b transcription in β-cells and be partially responsible for the up-regulation of miR-125b expression observed in AMPKdKO and LKB1KO islets.

TGF-β superfamily signalling has been widely studied in pancreatic development and postnatal growth [543-546]. In this context, it has been shown that both SMAD2 and SMAD3 proteins play a key role in the *in vitro* trans differentiation of a duct cell line (AR42J cells) into β-cells [547]. SMAD2 and SMAD3 increased expression towards the end of gestation has also been associated with endocrine cells maturation [545]. The role of SMAD2/3 in the mature β-cells is still elusive, however it has been recently demonstrated that T2D human islets present increased levels of phosphorylated SMAD3 (pSMAD3), suggesting that TGF-β/SMAD3 signalling might have a role in β-cell apoptosis [548]. Interestingly, overexpression of SMAD3, but not SMAD2, was found to decrease GSIS in pancreatic islets [549], whereas mice with β-cell-specific expression of constitutively active SMAD3 protein presented increased apoptosis and induced loss of cell mass. β -cell dysfunction and glucose intolerance were also observed. On the contrary, β-cell-specific inactivation of SMAD3 protects from apoptosis, preserves β-cell mass and improves cell function and glucose tolerance [548].

TGF-β/SMAD pathway has also been shown to play an important role in regulating the transcription of different miRNAs in different cell types [550-554]. In the contest of β-cell, it has been shown to enhance the transcription of both miR-375 and miR-26a and play an important role in the production of β-like-cells from mesenchymal stem cells and in the modulation of insulin secretion in vitro [555-557].

We originally hypothesized that AMPK acts to prevent SMAD2/3 activity also in β-cells which would contribute to repress miR-125b expression and, possibly, TGF-β signalling. Our results confirm that

TGF-β acts as a positive regulator of miR-125b expression in human islets although, paradoxically, treatment of MIN6 cells with TGF-β activators failed to increase miR-125b expression.

Moreover, silencing of *Smad3* in MIN6 did not result in miR-125b down-regulation but instead significantly increased miR-125b expression indicating that, in these cells, *Smad3* acts as an inhibitor.

The fact that the effect of *Smad3* silencing on miR-125b expression was very small might be due to the limited strength of the siRNA knock-down achieved in our experiments (30-50%). However, in humans the heterozygous mutations in the *SMAD3* gene is sufficient to cause a genetic disease characterized by cardiac abnormalities in combination with early-onset osteoarthritis (OA) known as aneurysm-osteoarthritis syndrome [558] indicating that these small changes might still produce meaningful biological outcomes.

 Nevertheless, our results suggest that the positive effect of TGF-β in human islet miR-125b expression occurs independently of *Smad3*. Unfortunately, I was not able to obtain enough material to silence *Smad3* and measure miR-125b expression in human islets, therefore we cannot exclude that this difference is due to the use of a cell line *versus* primary islets or species-specific. TGF-β Signaling can indeed mediate its effect through SMAD-independent pathways, including Erk, SAPK/JNK, and p38 MAPK pathways [559]

It is also important to mention that, for this experiment, a pool of siRNAs was used in order to achieve *Smad2/3* silencing and therefore it cannot be excluded that the increased expression of miR-125b might be owed to potential siRNAs off-target effects and further validation would be necessary to exclude this possibility. For example, a rescue experiment control, where a recombinant SMAD3 is re-introduced into the cells, or the introduction of a silent point mutation into the cDNA encoding for *Smad3* abolishing the complementarity with the siRNA could tell us more about the specificity of our results.

Moreover, it is important to mention that monitoring both mRNA and protein levels when performing siRNAs experiments is essential to validate the approach and exclude the possibility of a slower protein turnover that in turn results in unchanged protein levels. Unfortunately, due to the

unavailability of a specific antibody able to recognize SMAD3, here I cannot confirm that the 30-50% reduction of *Smad3* mRNA levels specifically resulted in the reduction of its protein levels. Consequently, it cannot be concluded that, in this experiment and under these conditions, SMAD3 is directly responsible for changes in miR-125b expression.

Unfortunately, I also failed to confirm SMAD2/3 binding to MIR*125B-2* peak (chr16:77644425+77644490) in mouse islets and in MIN6 cells, since SMAD2/3 ChIP-qPCR failed the detection of the DNA of interest. This data could indicate that SMAD2/3 does not bind to the proposed region in *MIR125B-2* under these conditions and therefore does not activate miR-125b expression directly.

However, It is important to mention that only when both *Smad2* and *Smad3* were silenced using siRNAs and SMAD2 and SMAD3 protein levels were measured, I realized that the antibody used for the ChIP-qPCR protocol, which in other cell types has been successfully used for ChIP experiments [560, 561], in MIN6 cells was only able to recognize SMAD2. Although these two proteins often work together as heterodimers, they can also function by binding to other R-SMAD proteins independently. Therefore, there is a possibility that by immunoprecipitating SMAD2, which according to the results seemed to have no role in the regulation of miR-125b expression, I have not been able to co-precipitate SMAD3 and as a consequence the DNA bound to it.

Moreover, it is also possible that SMAD3 binding to *MIR125B-2* only occurs under specific (stimulatory) conditions. For example, the activation of TGF-β pathway could have been crucial for SMAD2/3 binding to the target-DNA region by promoting their phosphorylation and subsequent nuclear localization. It is important to mention that in the attempt to activate TGF-β pathway in MIN6 cells by treating the cells with 60ng/ml of TGF-β, no differences were found in the active phosphorylated forms of SMAD2 and SMAD3 compared to non-treated cells, indicating that the treatment failed to activate the pathway. It's also true that, for their optimal growth, MIN6 cells require high FBS concentration which is already supplemented with TGF-β amongst other growth factors, hormones and nutrients [562]. The exact concentration of this cytokine in commercial FBS is unknown, however it could have been enough to saturate our system. In this case, what we see is not a failure to activate TGF-β, but the impossibility to activate it further.

AMPK has been previously shown to inhibit SMAD2/3 activation by TGF-β in other cell types [539]. In cancer cells it inhibits TGF-β-induced SMAD2/3 phosphorylation and negatively regulates cancer cell migration whereas in myofibroblasts AMPK it has been shown to directly prevent SMAD3 from binding DNA targets [540]. To our surprise, pharmacological activation of AMPK cells resulted in a small but significant increase in *Smad3* expression in MIN6.

Moreover, activation of AMPK while *Smad3* was silenced resulted in a small but significant increase of miR-125b expression, and this difference was slightly bigger than the one observed when only *Smad3* was silenced. This could indicate that by positively regulating *Smad3*, AMPK could contribute towards the inhibition of mir-125b expression. Nevertheless, the effects observed were very small and it is therefore likely that AMPK represses miR-125b expression by SMAD2/3 independent pathways (Figure 4.10)

As mentioned earlier, overexpression of SMAD3 reduces GSIS in pancreatic islets [563], whereas its constitutive activation in β-cell increases apoptosis and induces cell mass loss. In β-cell, the role of AMPK in these two processes is still debated. However, it has been shown that its activation can negatively regulate GSIS and promote apoptosis [194, 242, 243], whereas its inhibition can protect from apoptosis [564]. Therefore, it is possible that AMPK could regulate these pathways through the positive regulation of *Smad3* and that the inhibition of SMAD3 or the regulation of its downstream effectors in β-cells might be used as a therapeutic strategy to improve insulin secretion in pre-diabetic and diabetic patients. Further research would be necessary to corroborate this hypothesis, especially considering that the effect of AMPK activation on *Smad3* we observed in our data was small.

In order to identify the TF that binds *MIR125B-2* peak*,* the reverse chromatin immunoprecipitation assay (reverse ChIP) could be used [565]. This technique uses a specific nucleic acid probe to isolate genomic DNA and the proteins associated to it. Then, using mass spectrometric analysis, it is possible to identify all the proteins bound to the locus of interest. This method allows the unbiased identification of targeted DNA locus-associated proteins and is used specially to find the regulatory proteins associated with the known DNA elements [565]. Once the TF has been identified, the role AMPK has in the regulation of this protein can be investigated. As mentioned in Chapter 3, it might

be possible that the regulation mediated by AMPK does not occur through the inhibition/stimulation of the TF factor that binds the miR-125b promoter, but it might instead occur at the level of the pri-miR-125b processing.
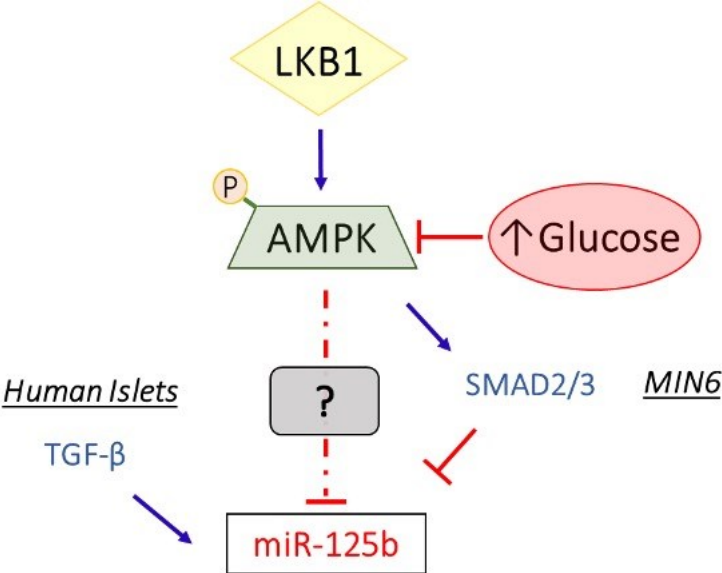
Figure 4.11: Simplified schematic representation of the results of the chapter

Chapter V: Identification of miR-125b molecular targets and function

5.1 Introduction

The ability of miRNAs to regulate multiple genes within a molecular pathway [566] makes them excellent novel targets for the treatment of diseases.

MiRNA-target interactions differ substantially between tissues and cell types [567] and, as a consequence, miRNA function is highly cell-specific. MiR-125b has been widely studied in the context of cancer [568] where it plays important roles in proliferation, apoptosis and differentiation [470]. For example, in neuroblastoma and lung fibroblast human cells miR-125b inhibits apoptosis by directly regulating p53 [569]. The regulation of p53 or its downstream proteins, such as p21, mediated by miR-125b has been suggested to occur also in colorectal cancer where higher levels of miR-125b have been associated with poor prognosis [570]. On the contrary, in osteosarcoma cells, overexpression of miR-125b suppresses proliferation and migration *in vitro* through the downregulation of the signal transducer and activator of transcription 3 (*STAT3*)[481].

Additionally, a dual role in inflammation has been recently attributed to miR-125b. In human osteoarthritis chondrocytes, miR-125b targets inflammatory genes in the TRAF6/MAPKs/NF-κB pathway and therefore negatively regulates inflammation [571]. However, in the HT29 human colon adenocarcinoma cell line, miR-125b has been suggested to induce inflammation by targeting TRAF6 and TNF alpha induced protein 3 (TNFAIP3) [572]. Interestingly, increased levels of miR-125b in vascular smooth muscle cells (VSMC) of diabetic db/db mice have been associated with increased expression of inflammatory genes. In VSMC cells, miR-125b directly targeted the histone methyltransferase *Suv39h1*, which resulted in reduced H3K9me3 levels at the promoters of monocyte chemoattractant protein-1 (*Mcp-1*) and interleukin-6 (*Il-6*) [573].

As mentioned in previous chapters (1 and 4), high levels of circulating miR-125b are associated with hyperglycaemia and miR-125b expression in islets positively correlates with the BMI of the donors (Unpublished). Nevertheless, the function of miR-125b in β-cells is not clear, and only two gene targets have been proposed in these cells: *cMaf*, a positive regulator of glucagon expression

[574] and, more recently, *Dact1*, a member of JNK signalling pathway [575]. According to our preliminary data, overexpression of miR-125b in MIN6 cells significantly reduced insulin content and impaired GSIS (Figure 1.11). Overexpression of miR-125b also affected MIN6 cell morphology, where the cells appear less rounded and more sprawl.

MiRNA-target interactions can differ substantially between organisms, tissues and cells. In order to determine the role of miR-125b in β-cells and its potential as a target for the treatment of diabetes I aimed to identify miR-125b gene targets in a high-throughput manner in a mouse insulinoma cell line.

5.2 Results

5.2.1 High-Throughput identification of miR-125b targets

5.2.1.1 miR-125b overexpression alters several genes important for β-cells function

Transcriptome profiling following miRNA overexpression is one of the most widely used approaches to identify miRNA targets since miRNAs often result in degradation of the target mRNAs. Even though this approach doesn't differentiate direct from indirect miRNA targets, it has the advantage of providing an overall picture of the effect of the microRNA at the transcriptome level and, in combination with Gene Ontology analysis, can provide important insights into the molecular roles of the miRNA.

Thus, in order to identify the gene targets and molecular function of miR-125b, mouse insulinoma MIN6 cells were transfected with control or miR-125b mimics (5 replicates each) and subjected to RNA-seq in the WTCHG facility in Oxford. Upon receipt of the Fastq files containing the sequencing reads, the first step I performed was to use FastQC [486] to assess reads number and quality (see section 2.14.3 for more details). Importantly, in addition to list the number of reads, FastQC generates a QC report that contains 12 analysis modules that also provides information on base sequence quality and content, read length, k-mer content and detect the presence of ambiguous bases and over-represented sequences (Figure 5.1 A).
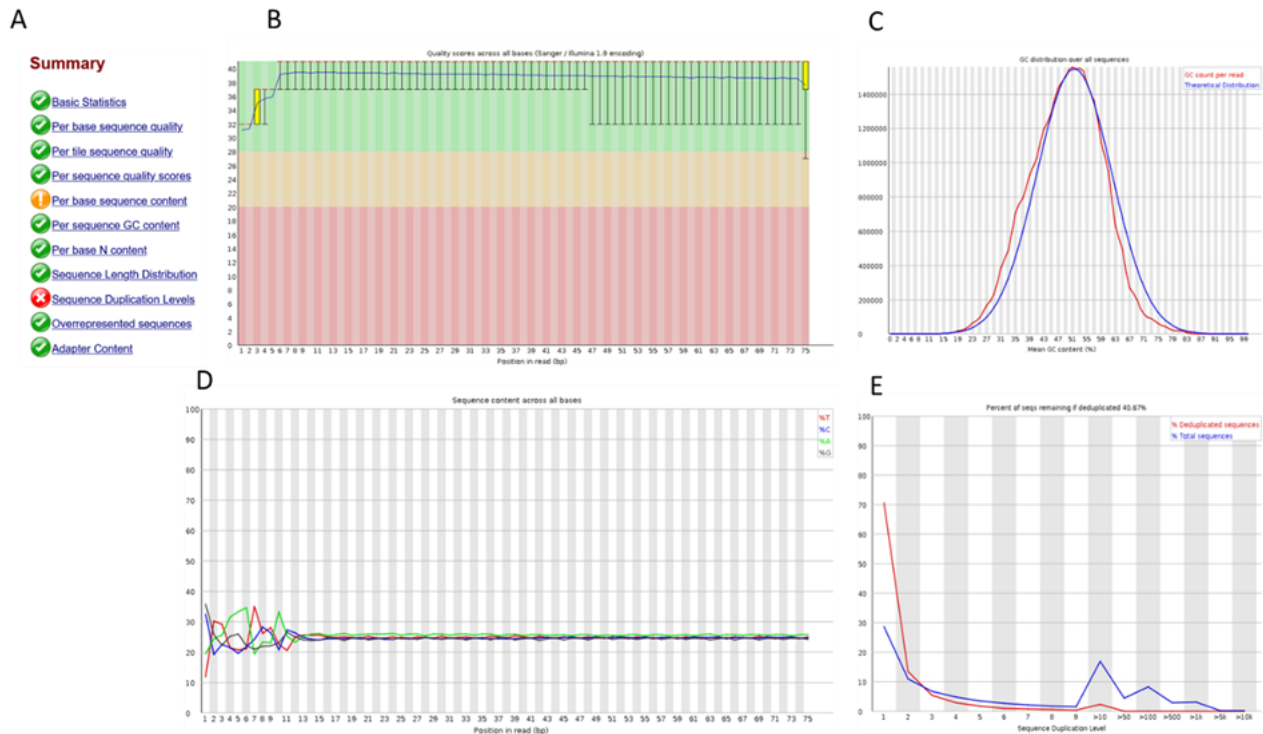
Figure 5.1: FastQC report for RNA-seq data indicating good quality across length of all sequences.
A. Summary representing the 12 analysis modules executed by FastQC on the raw sequenced reads.
B. Per Base Sequence Quality module. The Y-axis represents the quality score, while the X-axis represents the position in the read. The plot background colours identify good (Green), acceptable (Yellow), and bad (Red) quality scores.
C. Per sequence GC content. Graphic representation of GC distribution over all sequences (Red line) compared to a theoretical distribution (Blue line).
D. Per Base Sequence Content. The plot reports the percentage of bases called for each of the four nucleotides at each position across all reads in the raw sequenced reads. The X-axis plots the positions in reads. The Y-axis represents percentages of the occurrence of the bases along the reads. The module presents a warning that is normal and expected for RNA-seq data due to the nature of the library preparation protocol.
E. Sequence duplication level. The plot represents the relative number of sequences with different degrees of duplication. The X-axis represents the number of times each sequence is present in the raw sequenced reads. The Y-axis plots the percentage of reads of each given sequence.

Per Base Sequence Quality is one of the most important analysis modules (Figure 5.1 B). This plot provides an overview of the range of quality scores across all bases at each position in the FastQ file. The Y-axis represents the quality score, the X-axis represents the position in the read and the plot background is color-coded to identify good (green), acceptable (yellow), and bad (red) quality scores. A warning will be raised if the lower quartile for any base is less than 10, or if the median for any base is less than 25. All our samples showed good quality. Moreover, a good GC Content

156

score (Figure 5.1 C) and the absence of Overrepresented Sequences indicated that the libraries were free from adapter dimers or other types of contamination.

The quality module "*Per Base Sequence Content*" showed a warning in all our samples (Figure 5.1 A). As shown in the representative Figure 5.1 D there is a bias in the start positions of the reads (first 10-12 bases). This bias is quite common in RNA-seq data as a result of the sub-optimal "random" hexamer priming that occurs during RNA-seq library preparation leading to some hexamers being favoured during the priming step and eventually resulting in the biased composition over the region of the library primed by the "random" primers.

Additionally, a warning on the *Sequence Duplication Level* was also present and is indeed expected in RNA-seq (Figure 5.1 E). Normally, the over-duplication of highly expressed transcripts is tolerated in order to be able to see lowly expressed ones. That is mainly because RNA-seq counts the depth of coverage of each gene and genes with high coverage (such as INS for our dataset) are likely to be characterized by duplication.

Overall, an average of 30 million reads per sample were obtained and 90% of those were uniquely mapped using Salmon [487] to the library created combining the reference mouse GRCm38 cDNA and non-Coding RNA libraries. Salmon is a tool able to quantify the expression of transcripts starting from RNA-seq results using an expressive and realistic model of RNA-seq data that considers experimental attributes and biases commonly observed in *real* RNA-seq experiments [487].

Quantification files (quant.sf) for each sample I submitted to DESeq2 [488] for the analysis of differential gene expression. Results showed that miR-125b overexpression significantly altered (padj <0.1) the expression of many gene. The fact that *Bmf*, previously identified as miR-125b direct target, was found significantly down regulated at RNA level with a fold-change of 1.5 lead us to include all the genes that presented a fold-change of > ± 1.5 in the subsequent analysis. Importantly, of these genes, 285 were found downregulated and 305 upregulated (see Supplementary Table 2 in Appendix A).
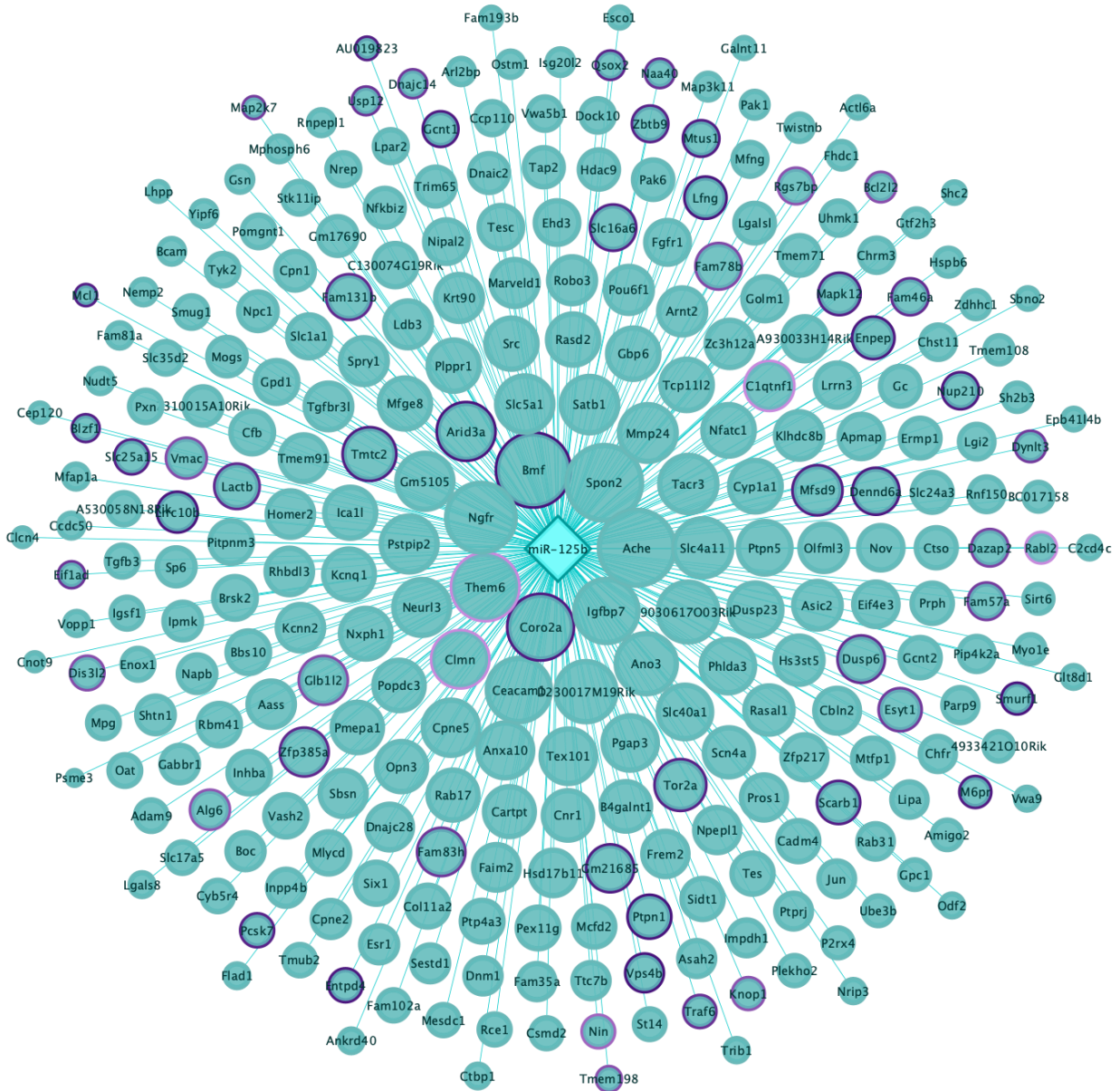
Figure 5.2: Down-regulated genes upon miR-125b overexpression in MIN6 cells
Cytoscape-generated layout of significantly down-regulated genes (fold change > - 1.5 and padj < 0.1) upon miR-125b overexpression in MIN6 cells vs. control. Node size represents the degree of the fold change (the larger, the bigger). The genes outline node colour (circles) indicates a pre-existing target prediction according to TargetScan whereas the intensity of the colour indicates the strength of the prediction score (the darker, the higher). A full list of genes can be found in Supplementary Table 2 in Appendix A

Given that miRNAs act to repress gene expression, the up-regulated genes were expected to be miR-125b indirect targets. Furthermore, since miRNAs can induce degradation of the mRNA targets, these data suggest that a considerable number of genes amongst the down-regulated ones could be potential β-cell direct target of miR-125b (Figure 5.2).

In order to have a better understanding of the biological pathways that were affected by the overexpression of miR-125b in β-cells, Gene Set Enrichment Analysis (GSEA) was carried out on the RNA-seq data results. GSEA is widely used to evaluate whether a-priori defined set of genes (gene collection) shows statistically significant, concordant expression changes between two biological conditions [576, 577]. All the genes detected in our RNA-seq were ranked according to their log2 fold-change upon miR-125b overexpression and used it as input for our GSEA analysis. GSEA allocates the ranked list an Enrichment Score (ES) for each pathway/biological process and tests for the significance against a null distribution of enrichment scores calculated for random permutated gene sets. The ES was subsequently normalised to account for the differences in the gene set sizes and the correlations between the gene set and the expression data set. This step generates a Normalized Enrichment Scores (NES) that allows to compare the results across all the gene sets obtained from the analysis. GSEA results therefore reflect the degree to which concordant over- or under-expressed genes in our gene set are overrepresented in each pathway. In other words, GSEA identifies the biological processes in which the up or down regulated genes, ranked respectively at the top or bottom of the list, are enriched.

As shown in Figure 5.3 A, GSEA revealed that genes up-regulated upon miR-125b overexpression are significantly associated with biological processes such as oxidative phosphorylation as well as ribosome and DNA replication, suggesting a role for miR-125b in these processes. On the other hand, genes ranked at the bottom of the list, thus down regulated upon miR-125b overexpression, are primarily involved in Glycosphingolipid biosynthesis, various receptor interaction pathways

and lysosomes (Figure 5.3 A). Examples of enrichment profiles for oxidative phosphorylation and Cytokine-Cytokine receptor interaction are shown in Figures 5.3 B and D.

A — Up-regulated gene sets

Gene sets	Size	NES	FDR
RIBOSOME	82	2.84	0.000
OXIDATIVE PHOSPHORYLATION	99	2.59	0.000
PARKINSONS DISEASE	101	2.40	0.000
HUNTINGTONS DISEASE	147	2.30	0.000
ALZHEIMERS DISEASE	132	2.22	0.000
PYRUVATE METABOLISM	29	2.18	0.000
PROTEASOME	40	2.15	0.000
VALINE LEUCINE AND ISOLEUCINE DEGRADATION	39	2.03	0.001
CITRATE CYCLE TCA CYCLE	26	1.99	0.001
DNA REPLICATION	35	1.87	0.006
BUTANOATE METABOLISM	25	1.82	0.009
SPLICEOSOME	119	1.81	0.009
ARACHIDONIC ACID METABOLISM	17	1.79	0.011
PYRIMIDINE METABOLISM	87	1.78	0.011
CYSTEINE AND METHIONINE METABOLISM	27	1.69	0.030
PROTEIN EXPORT	22	1.63	0.048
AMINOACYL TRNA BIOSYNTHESIS	41	1.62	0.046

C — Down-regulated gene sets

Gene set	Size	NES	FDR
GLYCOSPHINGOLIPID BIOSYNTHESIS LACTO AND NEOLACTO SERIES	16	-2.11	0.000
CYTOKINE CYTOKINE RECEPTOR INTERACTION	67	-1.89	0.025
LEISHMANIA INFECTION	34	-1.87	0.021
NEUROACTIVE LIGAND RECEPTOR INTERACTION	72	-1.79	0.046
MAPK SIGNALING PATHWAY	193	-1.68	0.102
DILATED CARDIOMYOPATHY	48	-1.67	0.089
AXON GUIDANCE	96	-1.64	0.102
CELL ADHESION MOLECULES CAMS	55	-1.63	0.091
ECM RECEPTOR INTERACTION	39	-1.61	0.100
O GLYCAN BIOSYNTHESIS	15	-1.58	0.109
N GLYCAN BIOSYNTHESIS	42	-1.52	0.164
GLYCOSAMINOGLYCAN BIOSYNTHESIS	23	-1.41	0.269
LYSOSOME	103	-1.38	0.276

Figure 5.3: Gene set enrichment analysis (GSEA) identifies miR-125b-regulated genes networks.
A and C. List of selected up-regulated and down-regulated gene sets upon miR-125b overexpression. Gene sets are ranked according to their normalized enrichment score (NES). The false discovery rate (FDR) is the estimated probability that a gene set with a given NES represents a false-positive.
B and D. Enrichment score (ES) plots for the Oxidative phosphorylation and cytokine-cytokine receptor interaction. Positive and negative ES values point to gene sets over-represented in the top up- or down-regulated genes upon miR-125b overexpression vs. control. Vertical bars refer to individual genes in a gene set and their position reflects the contribution of each gene to the ES.

5.2.1.2 Using RNA Immunoprecipitation and sequencing to identify miR-125b direct targets in MIN6 cells

As described above, miRNAs are normally found in a complex with miRISC and the sequence of the miRNA leads the complex to the target mRNAs. While the binding of miRNAs to the target mRNAs can result in their degradation, in many cases the repression occurs through the inhibition of the

mRNA translation, leaving the RNA levels of the target unchanged. Therefore, although data on the effect of miR-125b on the whole transcriptome is very informative and give us a picture of which processes might be targeted by miR-125b, this analysis not only doesn't distinguish between direct or indirect targets but also doesn't allow the detection of all those targets whose RNA levels remain unchanged and are only silenced at the translation level.

A mRNA that is actively repressed by a miRNA will be found in a complex with AGO2, one of the main protein components of miRISC. AGO2-RIP-seq has been previously used to identify all cellular mRNAs regulated by miRNAs [578-580]. Here, we designed an experimental approach that combined the overexpression of miR-125b in MIN6 cells and AGO2-RIP-seq to identify miR-125b direct targets in an unbiased manner (Figure 5.4).
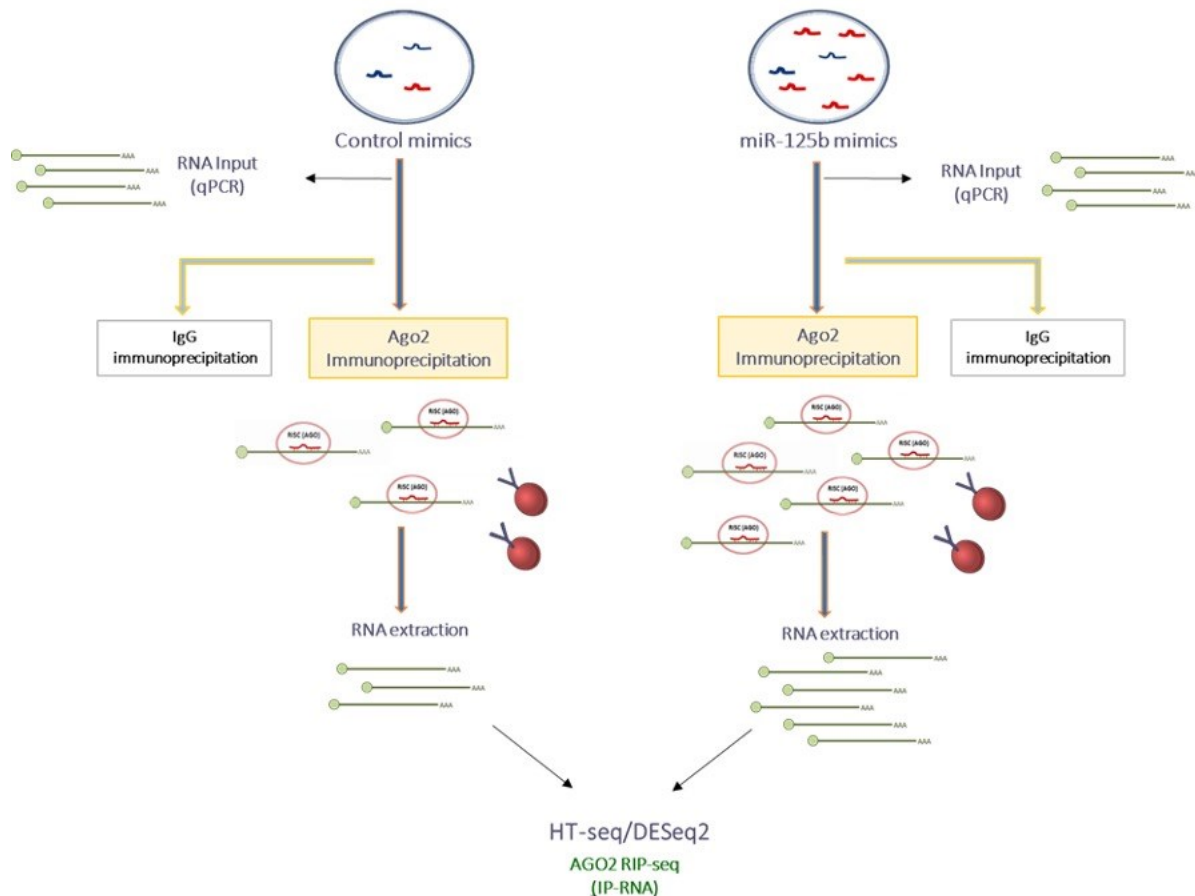
Figure 5.4: Schematic overview of the steps of the experimental analysis of RIP-seq for the identification of miR-125b direct targets.

As a result of miR-125b overexpression, increased binding of miR-125b targets to miRISC was expected. For this reason, I performed AGO2-RIP-seq in MIN6 cells transfected with miR-125b or control mimics and used DESeq2 as previously described to identify mRNAs enriched in AGO2-RIP upon miR-125b overexpression. As expected, miR-125b overexpression resulted in a significant increase in miR-125b levels (mean of 43.76 folds, P = 0.02) in these cells. Consistently, miR-125b was also increased 43 times in the AGO2-RIP (P = 0.0001), confirming that not only I was able to overexpress the miRNA but also to recover it after immunoprecipitation of miRISC. Importantly, the results obtained were specific for miR-125b since the approach didn't result in the displacement of miR-184 from the complex and its expression was unchanged both before and after immunoprecipitation conditions when miR-125b overexpressed was compared to the control samples (Figure 5.5 B).
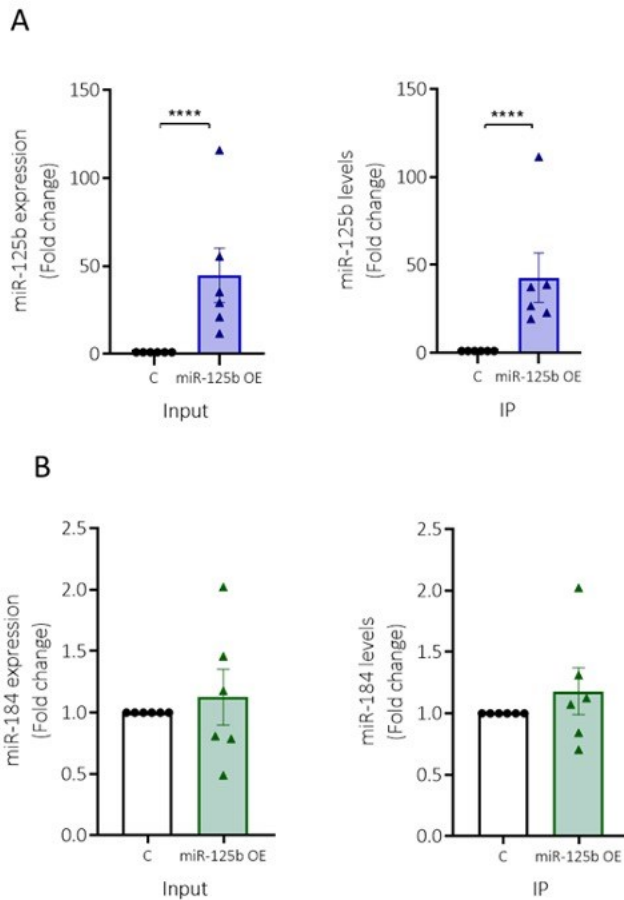
Figure 5.5: miR-125b was proportionally loaded into miRISC upon overexpression and did not cause displacement of a less expressed miRNA

A. miR-125b expression measured by RT-qPCR in MIN6 cells transfected with 5nmol of miR-125b (miR-125b-OE) or control mimic (C) for 24 hours before (Input) and after immunoprecipitation (IP).

B. miR-184 expression measured by RT-qPCR in MIN6 cells transfected with 5nmol of miR-125b (miR-125b-OE) or control mimic (C) for 24 hours before (Input) and after immunoprecipitation (IP). Each dot represents an independent experiment (n=6). Data are expressed as mean ± SEM relative to C levels. Values were normalised against the endogenous control let-7 and are represented as a normalized relative fold change to control. **** p<0.0001. P values were determined by two-tailed unpaired t-test on the log of the fold change values.

As previously reported, our earlier transcriptome analysis on cells transfected with miR-125b allowed us to identify several potential miR-125b direct targets, such as *Them6*, *Bmf* and *Hnf4g*. Moreover, *Bmf* and *Hnf4g* had already been identified by others [581, 582] as miR-125b targets in other cell types. Thus, before proceeding with the preparation of the libraries for sequencing, I used qPCR to test whether our approach was able to identify an enrichment of these mRNAs in AGO2-immunoprecipitates following miR-125b overexpression. As expected, overexpression of miR-125b resulted in a significant down regulation of *Bmf*, *Them6* and *Hnf4g* mRNAs levels (P = 0.0002, P = 0.0006 and P = 0.0024, respectively) (Figure 5.6.A) when compared to control in the cell extracts preceding immunoprecipitation (Input samples). *18s* was used as the negative control since its expression was found unchanged in the total RNA-seq data and we do not expect ribosomal RNA to be recovered in miRISC precipitates. *18s* mRNA levels, as expected, were not affected by miR-125b overexpression.

Importantly, these mRNAs were found significantly enriched in samples where AGO2 antibody was used for the immunoprecipitation (in both control or miR-125b mimic transfected cells) when compared to the IgG-immunoprecipitate samples used as negative control (Figure 5.6 B) suggesting that these mRNAs are loaded in miRISC and therefore targeted by miRNAs. Moreover, I found that the amount of *Them6* and *Hnf4g* mRNAs recovered in AGO2-IPs was significantly higher (P = 0.0056 and P = 0.0022, respectively) in cells overexpressing miR-125b, confirming that the overexpression of miR-125b enhances the binding of these targets to the miRISC (Figure 5.6 C). *Bmf* mRNA levels were also higher in the AGO2-precipitates of miR-125b overexpressing cells; However, the difference was not statistically significant (P = 0.1). As expected, *18s* was not enriched in any of the conditions. Our results suggest that these genes that in other cell lines have already been suggested to be miR-125b direct targets might be directly regulated by miR-125b also in β-cells and that our AGO2-RIP approach can successfully identify miR-125b direct targets.
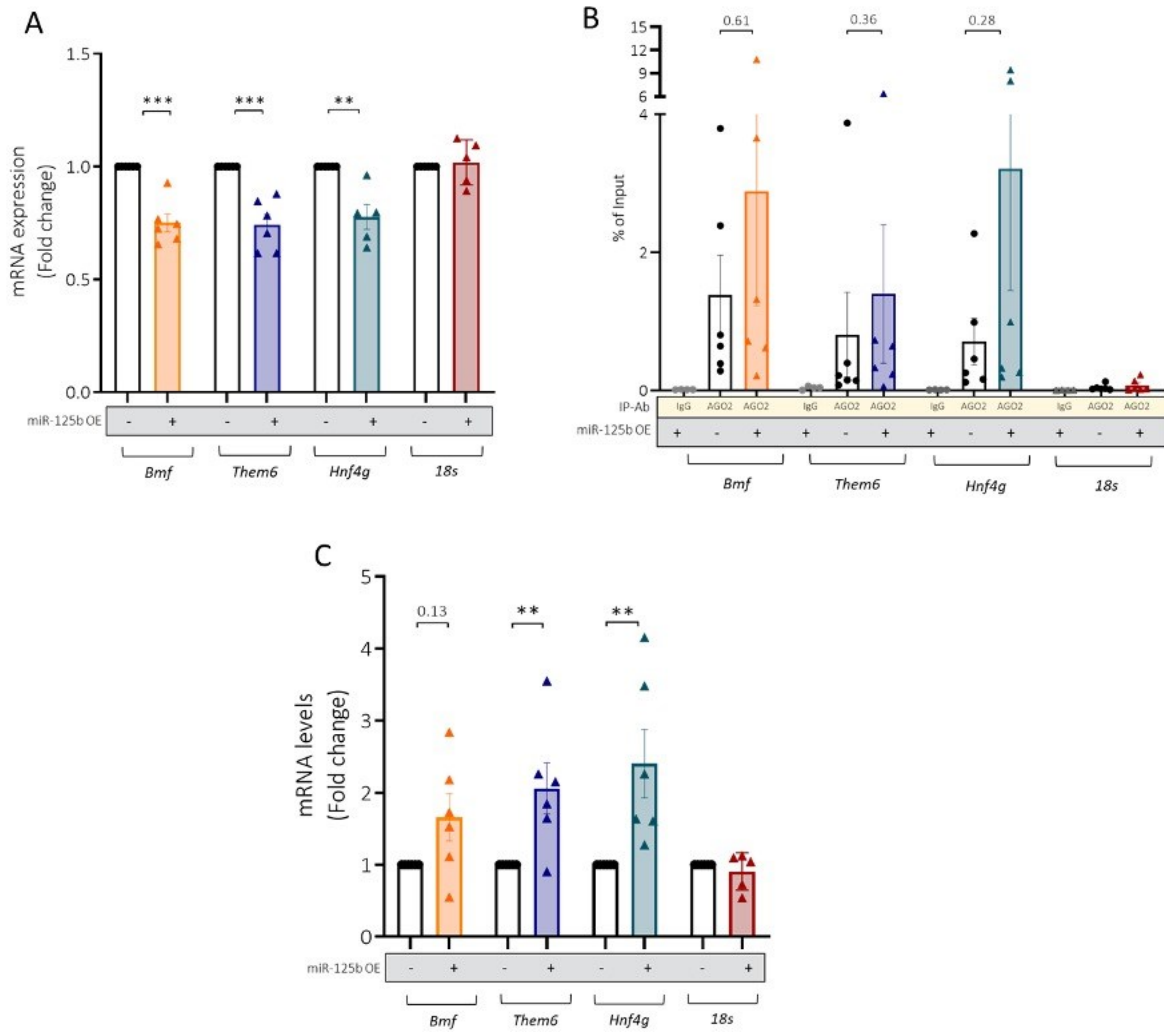
Figure 5.6: miR-125b previously identified targets were enriched after AGO2 immunoprecipitation upon miR-125b overexpression.

A. Bmf, Them6, Hnf4g and 18s mRNA expression measured by RT-qPCR in MIN6 cells transfected with 5nmol miR-125b (+) or control mimic (-) for 24 hours and whose RNA was collected before immunoprecipitation using anti-AGO2 antibody.

B. Bmf, Them6, Hnf4g and 18s mRNA expression measured by RT-qPCR in MIN6 cells transfected with 5nmol miR-125b (+) or control mimic (-) for 24 hours and whose RNA was collected after immunoprecipitation performed with anti-AGO2 antibody (AGO2) or negative control IgG.

C. Bmf, Them6, Hnf4g and 18s mRNA expression measured by RT-qPCR in MIN6 cells transfected with 5nmol miR-125b (+) or control mimic (-) for 24 hours and whose RNA was collected after immunoprecipitation using anti-AGO2 antibody. Each dot represents an independent experiment (n=6).

Data are expressed as mean ± SEM relative to control levels. Values in (A) and (C) are represented as a normalized relative fold change to control. (C) values are represented as % of input. P values for (A) and (C) were determined by two-tailed unpaired t-test on the log of the fold change values of each gene compared to its control. P values for (B) were determined by one-way ANOVA with Sidak´s multiple comparisons test, with a single pooled variance performed on the set of data of each individual gene. ** p<0.01 *** p<0.001.

Thus, the RNA obtained from AGO2-RIP was used to generate cDNA libraries that were subsequentially sequenced at the Imperial College BRC facility. The data obtained from the sequencing were analysed using a similar pipeline to the one described for total RNA-seq. The quality was assessed with FastQC and DESeq2 was used to perform a differential analysis where fold-changes between miR-125b overexpression and control samples were generated for each given gene. Amongst all the genes identified (~10,000), 109 were found significantly up-regulated (padj < 0.1) whereas 25 were found significantly down-regulated.

Interestingly, *Bmf, Them6 and Hnf4g* (fold-change = 0.94, P = 0.99; fold-change = 1.12, P = 0.99; fold-change = 0.93, P = 0.99), previously identified miR-125b direct targets by us and others [581, 582], were not amongst the significantly up-regulated genes indicating that by using only the data from the RIP-seq approach we might have missed valuable information. This led us to consider a different approach that could overcome these problems.

5.2.1.3 Identification of miR-125b direct targets in MIN6 cells by combining RIP-seq/RNA-seq data

RIP-seq protocol, which is not characterized by a cross-link step that covalently binds all the components of the miRISC complex together, mostly relies upon both a stable interaction between miRNA and mRNA targets and the miRISC ability to survive the precipitation step. This means that, if on one side it is possible to miss the identification of all those targets which binding to the miRNA is thermodynamically not strong, on the other side, since the protocol is characterized by gentle washing steps, it is reasonably probable to obtain false-positive results due to the recovery, after the immunoprecipitation step, of RNAs which expression might be indirectly upregulated upon miR-125b overexpression and unspecifically bound to miRISC.

For these reasons, we decided to combine the fold-changes obtained from both total RNA seq (T-RNA) and RIP-seq (IP-RNA) and create an IP-RNA/T-RNA ratio. We were expecting the targets of miR-125b to be enriched in IP-RNA and depending on the type of repression of the target, either reduced or unchanged in T-RNA. Therefore, using this approach, all the mRNAs with a ratio greater than 1 (~7000 genes) were considered to be miR-125b direct targets.

To further validate the approach I performed a word enrichment analysis where all the genes were sorted by their IP-RNA/T-RNA ratio and submitted to C-Words [423]. Using this unbiased method, specifically designed for the discovery of regulatory motifs, we aimed to confirm that our dataset of newly identified targets was enriched with the miR-125b seed site. As shown in Figure 5.7, the 3'UTR and CDS of the highly ranked (high RIP-RNA/T-RNA ratio) targets were significantly enriched in miR-125b seed complementary motifs, suggesting that the interaction between miR-125b and its mRNA targets occurs through the binding to both 3'UTR and CDS. However, I found that the Z-score obtained using cWORDs for the analysis performed on the 3'UTR was higher (>20) than the one obtained on CDS (>15).
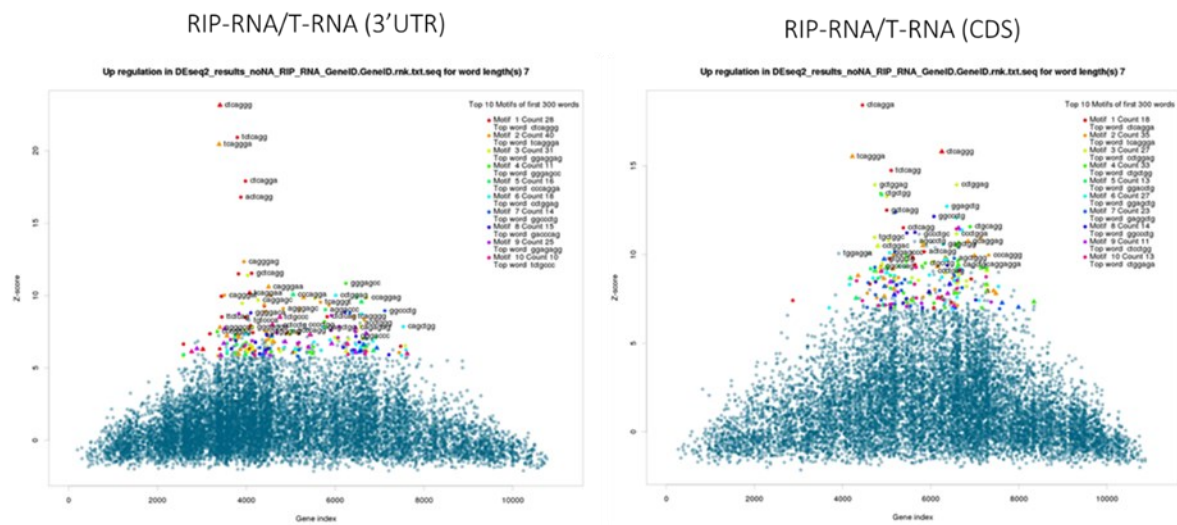
Figure 5.7: MiR-125b seed site is enriched in both 3'-untraslated region and coding region of genes identified using RIP-RNA/T-RNA.
cWords word cluster plots showing 7mer words enriched in 3'UTRs and CDSs of genes ranked by RIP-RNA/T-RNA ratio. Each dot represents a word, the Y-position reflects the maximum score of an enriched word and X-position shows the gene-index where the Z-score is maximum. Triangles annotate known seed sites of human miRNAs. Red triangles show miR-125 seed sites (TCTCAGGGA). Words are clustered by sequence similarity using the UPGMA algorithm and coloured according to what motif (or cluster) they belong to (only some amongst the top 100 words are plotted).

Figure 5.8 A reports the 180 genes that presented an IP-RNA/T-RNA ratio bigger than 1.5, value used as cut-off to narrow the analysis down and simplify the system biology visualization. Interestingly, as shown in the Venn diagram in Figure 5.8 B, out of these 180 most enriched targets only 58 were significantly down regulated at the mRNA level and only 5 of those were also significantly

enriched in the RIP-seq dataset. Additionally, the diagram shows how a considerable number of newly identified miR-125b direct targets reach statistically significant changes in only one of the two datasets, suggesting that adopting the ratio approach I was able to identify targets of miR-125b that would have otherwise been missed if the RIP-seq and RNA-seq datasets were considered independently. Moreover, according to an analysis performed using a combination of Gorilla, TargetScan and RNA hybrid, only 84 targets contained predicted binding sites for miR-125b (Supplementary Table 1 in Appendix A). Thus, our approach allowed the identification of miR-125b targets that do not contain canonical binding sites and therefore are not detected by in silico tools.
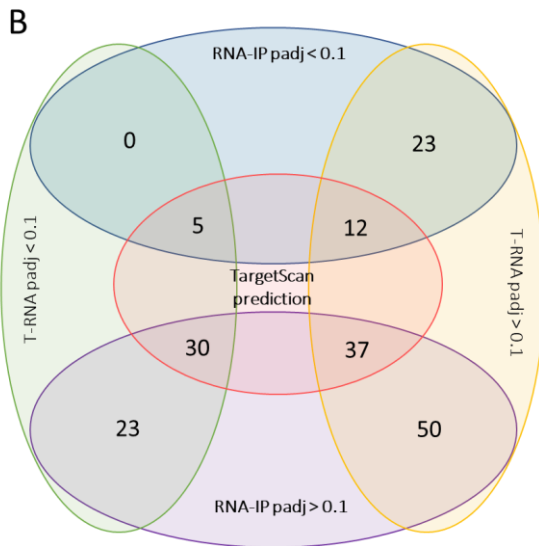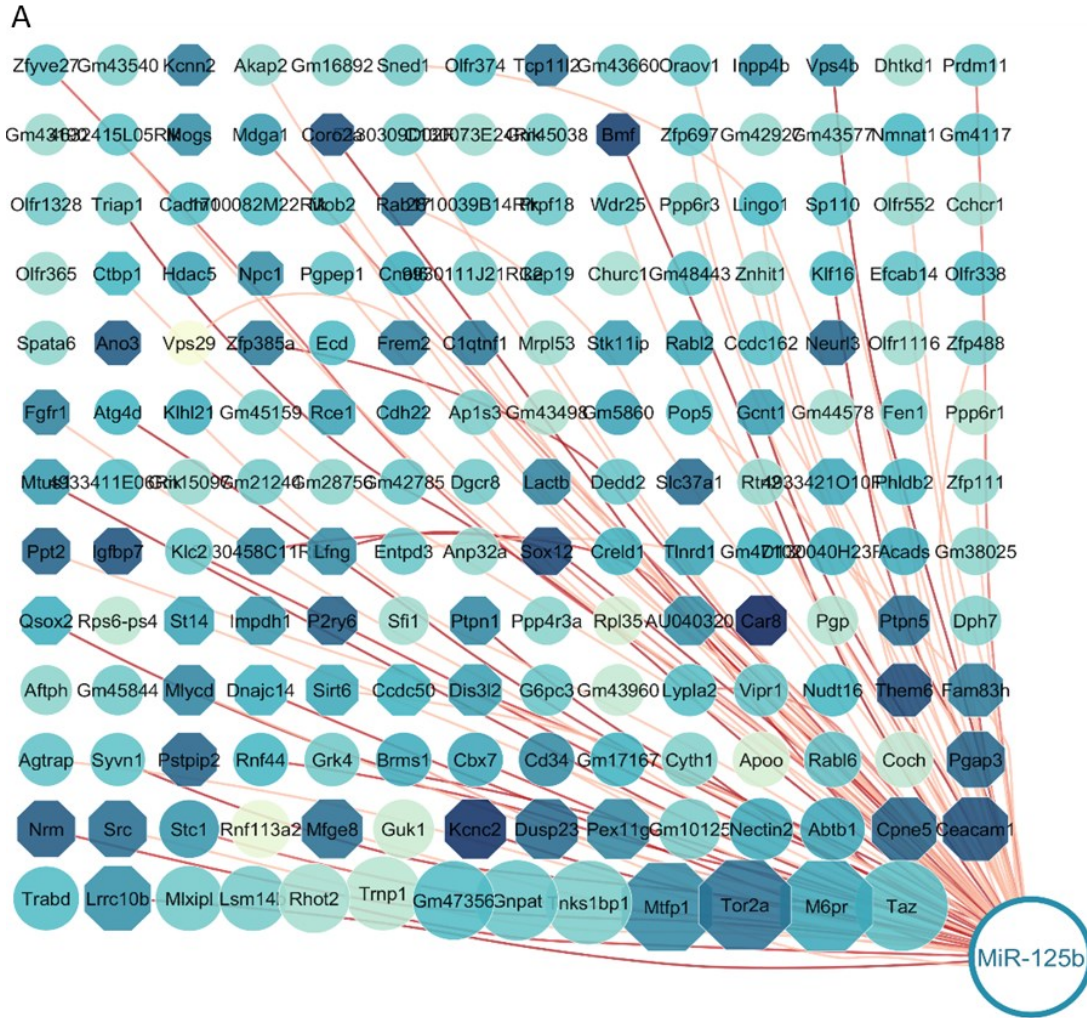
Figure 5.8: miR-125b direct target in MIN6 cells.
A.Cytoscape-generated layout of 180 genes enriched in AGO2 IP-RNA versus T-RNA following miR-125b overexpression (IP-RNA/T-RNA>1.5). Node size represents the degree of the ratio (IP-RNA/T-RNA). The genes node colour indicates T-RNA fold change (the darker, the bigger). Node shape represents the T-RNA fold change padj (hexagonal padj < 0.1; circular padj > 0.1). Red lines indicate a pre-existing target prediction according to TargetScan whereas the intensity of the colour indicates the strength of the prediction score (the darker, the higher). No lines are represented for genes having no prediction for miR-125b binding site. A full list of targets and related values can be found in Supplementary Table 1 in Appendix A.

B. Venn Diagram showing the total numbers of miR-125b direct target genes with RNA-IP padj < 0.1 (Blue), RNA-IP padj > 0.1 (Violet), T-RNA padj < 0.1 (Green), T-RNA padj > 0.1 (Yellow). In red are shown the number of genes with a pre-existing target prediction according to TargetScan.

5.2.2 Luciferase assay validation

Validation of gene-specific miRNA-target interactions is broadly used to confirm data obtained from a high-throughput experimental approach. We decided to select 7 targets among the ones with higher IP-RNA/T-RNA ratio and, using the dual luciferase assay (pmirGLO Dual-Luciferase vector) in mouse MIN6 cells, I aimed to confirm that miR-125b was directly responsible for their regulation. From the analysis performed using Gorilla, TargetScan, and RNA hybrid it was predicted that for *Taz*, *M6pr*, *Tor2a* and *Mtfp1* miR-125b binding sites were in the 3'UTR sequence, whereas for *Tnks1bp1*, *Gnpat* and *Trnp1* in the CDSs (Figure 5.9).

Gene Name	Protein	Function	Target Scan	RNA hybrid	Fold Change T-RNA	Fold Change IP-RNA	IP-RNA/ T-RNA
Taz	Tafazzin	ROS production/ Oxidative stress	3'UTR		0.9	2.2	2.5
M6pr	Mannose receptor (MPR)	Trafficking	3'UTR		0.8	2.0	2.5
Tor2a	Torsin Family 2, member A	Oxidative stress	3'UTR		0.7	1.8	2.5
Mtfp1	Mitochondrial fission process protein 1	Mitochodrial elongation/respiration	3'UTR		0.8	1.9	2.4
Tnks1bp1	Tankyrase 1 Binding Protein 1	DNA repair		CDS	0.9	2.2	2.4
Gnpat	Glyceronephosphate O-Acyltransferase	Iron metabolism		CDS	0.9	2.1	2.3
Trnp1	TMF1-regulated nuclear protein 1	DNA associated neural stem cell renewal		CDS	1.0	2.2	2.2

Figure 5.9: Top direct targets of miR-125b using IP-RNA/T-RNA approach.
The list only includes the top 7 miR-125b direct targets and details the corresponding protein names and functions. Predicted binding sites accordingly to Gorilla, TargetScan and RNAhybrid are also shown. The last three columns show the fold-change in the total RNA-seq analysis (T-RNA), RIP-seq analysis (IP-RNA) and their ratio (IP-RNA/T-RNA), respectively.

I then cloned the 3'-UTRs and/or the CDS of the mRNA targets, containing the predicted miR-125b binding sites, downstream of the *Firefly* luciferase ORF in the pmirGLO plasmid and then co-transfected the cells with the plasmid and either miR-125b or control mimics. pmirGLO also encodes for *Renilla* luciferase that was used as transfection control. With this approach I was simply testing whether the 3'UTR or the CDS of the target of interest is responsible for the regulation of protein

production. A pmirGLO vector alone with no insert was used as experimental negative control and a construct containing three fully complementary binding sites for miR-125b was used as positive control.

Using this approach, I was able to confirm that miR-125b reduced *Firefly* luciferase activity of the constructs containing the 3'UTR of *M6pr*, *Taz, Mtfp1, Tor2a* and the CDS of *Gnpat* (Figure 5.10), thus demonstrating that miR-125b silences gene expression through those sequences. As expected, the construct containing three perfectly complementary sequences to miR-125b showed a strong reduction of *Firefly* activity upon miR-125b mimic transfection whereas *Firefly* activity didn't change for the empty vector.

On the contrary, the *Firefly* luciferase activity of the plasmid containing *Trnp1* 3'UTR was not affected by miR-125b overexpression confirming that there is no binding site in this region as reported by the prediction tool. Unfortunately, I was unable to obtain a construct that contained *Trnp1* CDS downstream of the firefly luciferase ORF and therefore it cannot be excluded that miR-125b can regulate *Trnp1* expression through binding to its CDS. I also failed to validate the interaction of miR-125b with the CDS of *Tnks1bp1*, since there was no difference in the luciferase expression between the samples transfected with miR-125b mimic vs. control.
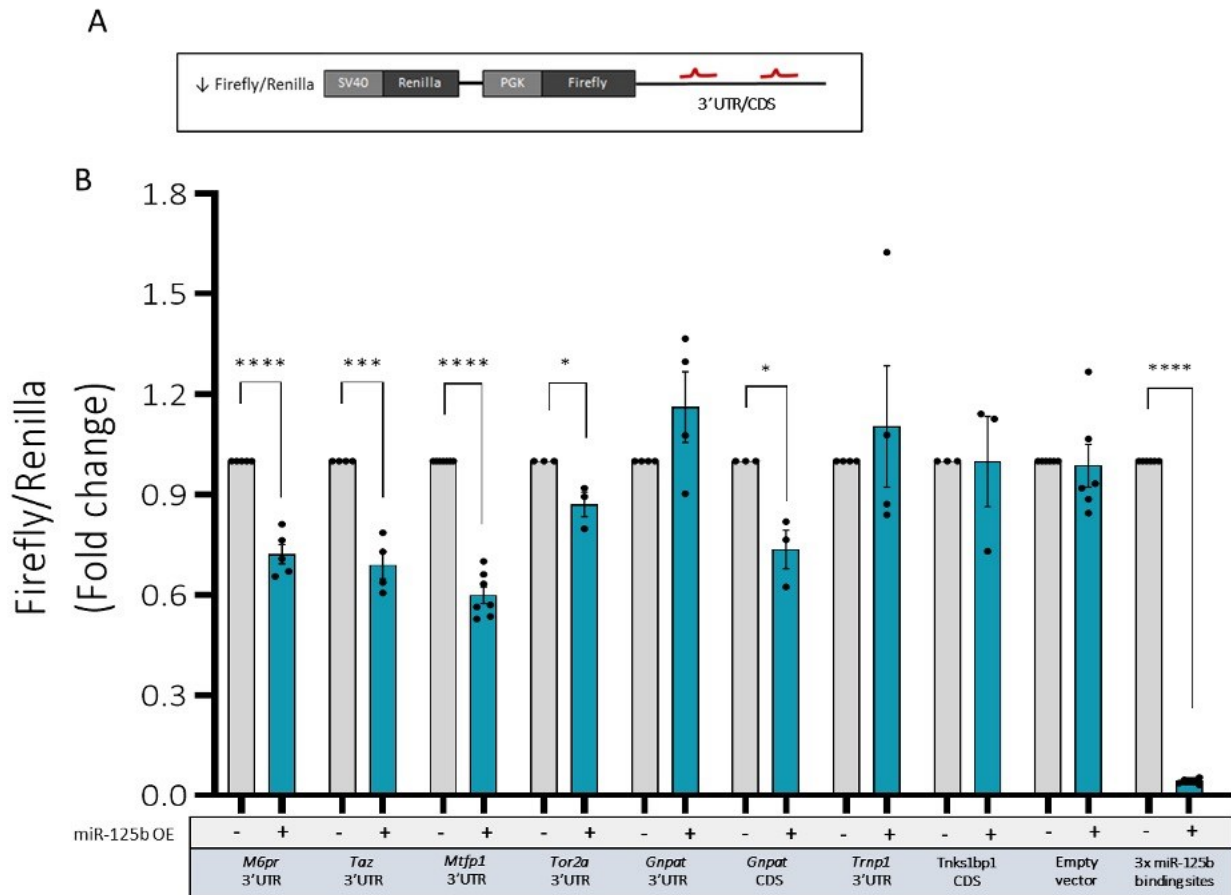
Figure 5.10: miR-125b directly affects the expression of genes with higher IP-RNA/T-RNA ratio through their 3'UTR or CDS.

A. Schematic representation of luciferase reporters containing the full-length sequence of CDSs or 3'UTRs downstream the firefly luciferase ORF. The gene encoding for Renilla luciferase is fused to a SV40 constitutive promoter and is independently expressed by the same vector.

B. Firefly luciferase activity of MIN6 cells co-transfected with luciferase reporters containing CDSs or 3'UTRs, or both, of the newly identified 6 top direct targets of miR-125b and with 5nmol of miR-125b (+) or control mimic (-). A construct containing three perfectly complementary binding sites for miR-125 (3 x miR-125b) is used as positive control whereas an empty luciferase reporter is used and negative control. Each dot represents an independent experiment (n=3-6). Data are expressed as mean ± SEM relative to control levels. Values were normalised against the levels of Renilla luciferase and are represented as a normalized relative fold change to control. * p<0.05, *** p<0.001, **** p<0.0001. P values were determined by two-tailed unpaired t-test on the log of the fold change values.

5.2.3 miR-125b role in lysosomal function

GSEA performed on both RNA-seq down-regulated genes upon miR-125b overexpression, and the genes considered direct targets in the IP-RNA/T-RNA approach (ratio >1.5) suggests that miR-125b

could have a role in the regulation of the lysosomal functions. The lysosome is a subcellular organelle that controls nutrient sensing and cellular metabolism [583]. Since nutrients tightly control the secretory machinery in β-cells, we expect the lysosomes to play a key role in β-cell function. It has been demonstrated that nutrient-deprived β-cells deliver newly synthesized secretory granules to lysosomes in the Golgi area, where degradation of their cargo (proinsulin and insulin) prevent unwanted insulin release [584], a mechanism that has also been associated with T2D [585]. Indeed, uncontrolled lysosomal degradation of insulin granules has been shown to occur in diabetic mice as well as in human and mouse pancreatic islets chronically exposed to elevated levels of glucose [585].

Considering that our preliminary data showed impaired insulin content in MIN6 cells upon miR-125b overexpression, we decided to further investigate the role of miR-125b in the control of lysosomal function. *M6pr* is one of the newly identified miR-125b top targets identified by our approach and encodes for the homonym protein M6PR (Mannose 6-Phosphate Receptor), which is involved in intracellular trafficking of lysosomal enzymes [586]. While the function of this protein has not been characterized in β-cells yet, since its dysregulation might affect lysosomal function [587] we hypothesised that, by targeting M6PR, miR-125b could impact β-cell nutrient signalling and/or insulin biosynthesis.

miR-125b directly binds to M6pr 3'UTR

As a first step, I aimed to confirm that miR-125b represses *M6pr* by directly binding to the 3'UTR of *M6Pr via* the predicted miR-125b binding site in nucleotides 851-858. In order to achieve this, I generated a pmirGLO Dual-Luciferase vector that contained a mutation in the predicted binding site and compared the luciferase activity with the one caused by the *M6pr wt* 3'UTR sequence. As shown in Figure 5.11, miR-125b lost the capacity of repressing luciferase production in the presence of mutations at the predicted binding site, confirming that miR-125b inhibit *M6pr* through direct binding to the positions 851-858 in the *M6pr* 3'UTR.

Unfortunately, I was not able to confirm these results at the protein level in MIN6 cells due to the lack of antibodies able to detect M6PR in mouse samples.

172
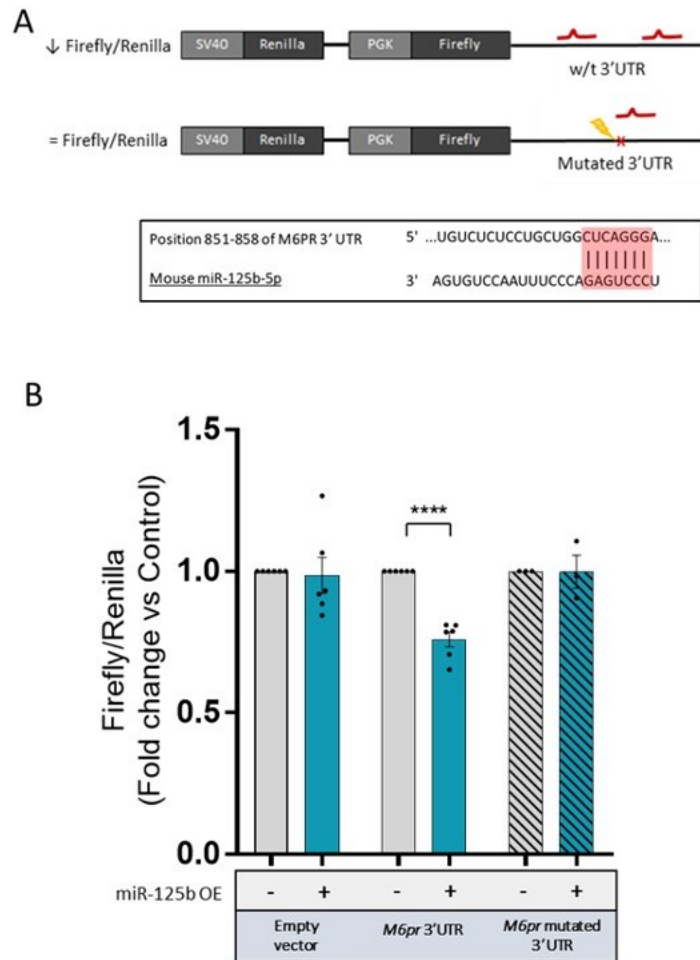
Figure 5.11: miR-125b directly targets *M6pr* by binding the 3'UTR

A. Schematic representation of luciferase reporters containing wildtype (w/t) or mutated 3'UTRs sequence downstream the firefly luciferase ORF. The gene encoding for Renilla luciferase is fused to a SV40 constitutive promoter and is independently expressed by the same vector. In red is indicated the exact position of miR-125b seed region binding to the 3'UTR of mouse *M6pr.*

B. Firefly luciferase activity of MIN6 cells co-transfected with luciferase reporters containing wild type or mutated 3'UTRs of *M6pr* and 5nmol of miR-125b (+) or control mimic (-). An empty luciferase reporter is used as negative control. Each dot represents an independent experiment (n=3-6). Data are expressed as mean ± SEM relative to control levels. Values were normalised against the levels of Renilla luciferase and are represented as a normalized relative fold change to control. **** p<0.0001. P values were determined by two-tailed unpaired t-test on the log of the fold change values.

5.2.3.1 miR-125b regulates M6PR levels in human ENDOC-βH1 cells

Protein levels of M6PR were also assessed in human β-cells (EndoC-βH1) transfected with miR-125b mimics and in a population of EndoC-βH1 cells with CRISPR-Cas9-mediated reduction of miR-

173

125b (kindly provided by Rebecca Cheung in our group). Upon miR-125b overexpression (Fold-change 0.3, P <0.0001) I saw a significant (P = 0.004) reduction of M6PR protein levels, whereas depletion (Fold-change= 412.7; P <0.0001) caused a significant (P = 0.03) increase in M6PR expression (Figure 5.12 A and B).
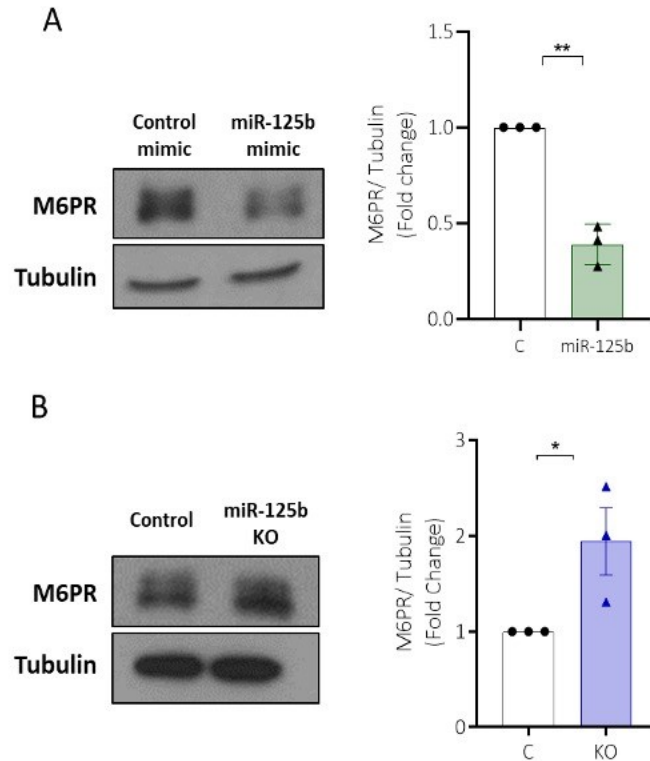
Figure 5.12: miR-125b regulates M6PR levels in ENDOC-βH1
A. Representative western blot and relative quantification of M6PR protein levels in ENDOC-βH1 transfected with 0.5nM of miR-125b or control mimic (C), for 48 hours. Tubulin is used as a loading control. n=3 independent experiments.
B. Representative western blot and relative quantification of M6PR protein levels in ENDOC-βH1 with CRISPR-Cas9-mediated reduction of miR-125b (KO) vs control (C). Tubulin is used as a loading control. Quantitative analysis of the immunoblots was determined by ImageJ. Each dot represents an independent experiment (n=3). Data are expressed as mean ± SEM relative to control levels. Values were normalised against the levels of Tubulin and are represented as a normalized relative fold change to control. * p<0.05 ** p<0.01. P values were determined by two-tailed unpaired t-test on the log of the fold change values.

5.2.3.2 Overexpression of miR-125b impairs lysosomal structure

To investigate miR-125b role in the regulation of lysosomal function, in collaboration with Dr. Alejandra Thomas we performed an Electron Microscopy (EM) analysis on MIN6 cells transfected

with miR-125b mimic or control. Results showed that miR-125b overexpression was responsible for lysosomal abnormalities, leading to an abnormally enlarged structure as visible in Figure 5.13 A.

We know that lysosomes are the primary degradative compartments of eukaryotic cells with catabolic activity [588]. Lysosomes can degrade a wide variety of intracellular material, such as proteins, nucleic acids and complex lipids, into their basic building blocks that the cell can then reuse to generate new cellular components or produce energy [589]. These organelles are also responsible for the degradation of transmembrane receptors such as Glucagon-like peptide-1 receptor (GLP-1R). In pancreatic β-cells, lysosomes have been identified as the major post-endocytic destination of this receptor [590]. Based on this, we hypothesised that, if by altering the lysosomal structure miR-125b is also responsible for a loss of lysosomal function, we would expect an impaired degradation of GLP-1R.

To test this hypothesis, I was provided with MIN6B1 cells that stably express Human GLP-1R with a small genetically encoded tag at its N-terminus (SNAP-GLP-1R). I overexpressed miR-125b for 48hr and induced the degradation of the receptor by treating the cells with GLP-1 for the following 6 hours. As shown in Figure 5.13 B , SNAP-GLP-1R degradation was induced by GLP-1 as expected but, in contrast to what was previously hypothesised, I also saw a tendency to an increased GLP-1R degradation upon miR-125b overexpression in both GLP-1 treated and not treated samples indicating that miR-125b could have a potential role in the GLP-1R degradation pathway.

It is important to mention that in this specific western blot analysis GAPDH (glyceraldehyde-3-phosphate dehydrogenase) – a protein involved in the catalysation of the sixth step of glycolysis, a process important for the breakdown of glucose – was used as the loading control. Although one of the most frequently used housekeeping genes, this protein has been found to be unstable in metabolic processes [591], making it not the most suitable candidate to use as loading control for this experiment. In fact, the selection of a different protein might be necessary for a more accurate quantification of GLP1-R.
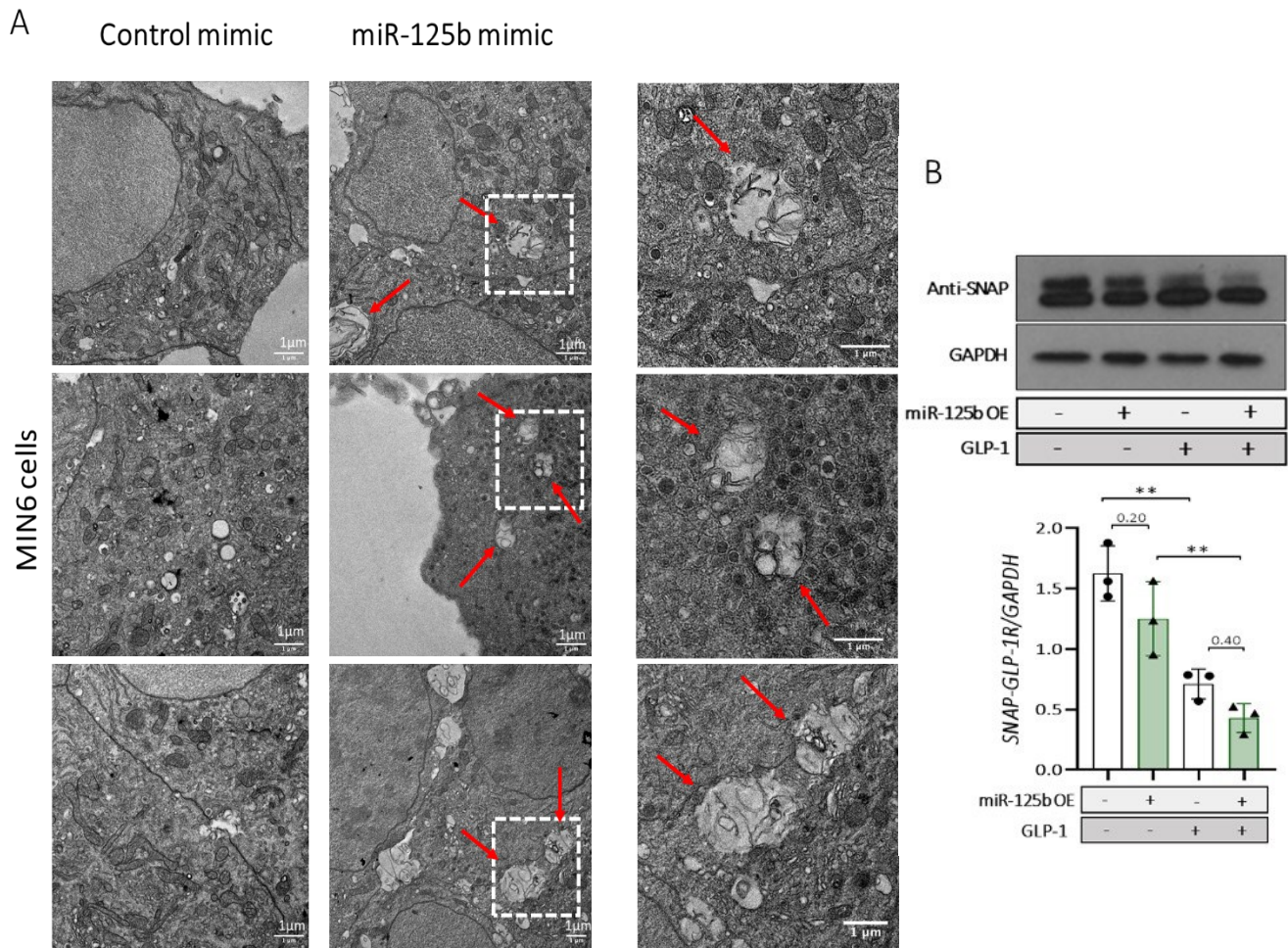
Figure 5.13: Overexpression of miR-125b alters lysosomal structure but does not affect GLP-1R degradation

A. Representative EM images showing enlarged lysosomal structures (arrows) in MIN6 cells transfected with 5nmol miR-125b or control mimic.

B. Representative western blot and relative quantification of anti-SNAP levels in MIN6B1 cells that stably express human SNAP-GLP-1R. Cells were transfected with 5nM of miR-125b (+) or control mimic (-) for 24 hours and internalization/degradation of the GLP-1R was induced by treating the cells with 100nM of GLP-1 (+) or vehicle (-) for 6 hours. GAPDH is used as a loading control.

Quantitative analysis of the immunoblots was determined by ImageJ. Each dot represents an independent experiment (n=3). Data are expressed as mean ± SEM relative to control levels. Values were normalised against the levels of Tubulin. ** $p<0.01$. P values were determined by two-way ANOVA with Sidak´s multiple comparisons test, with a single pooled variance on the normalized values.

5.3 Discussion

MiRNAs can regulate multiple genes within a molecular pathway [566] and this makes them excellent novel targets for the treatment of disease, however, potential off-target effects and safety are the main challenges for developing miRNA-based therapeutics [445]. Since miRNA-target interactions can differ substantially between different organisms, tissue or even cell-types [567], one of the best way to understand the peculiar function fulfilled by a certain miRNA in a specific cell type, developmental stage or even in response to external stimuli is to identify miRNAs direct targets [446]. Only in this way miRNA-based therapeutics inducible toxic side effects can be prevented [445].

The importance of identifying miRNAs direct targets in order to unravel their molecular mechanism of action in specific cell types/tissues led to the development of numerous methods in the past few years (More in section 1.5.9).

For miR-125b direct target identification we decided to combine the data obtained from AGO2-miRNA-mRNA complex immunoprecipitation followed by high throughput sequencing with a normal transcriptome analysis, both performed upon miR-125b overexpression in MIN6 cells. Using the two datasets I generated a ratio (IP-RNA/T-RNA) and I was expecting the direct targets to have this ratio >1. Whit this method I was able to identify 180 direct targets of miR-125b in MIN6 cells and validated 4 of the one with the bigger ratio.

Both cWORDs analysis and luciferase assay validations confirmed that our approach was successful. A significant enrichment of miR-125b seed complementary motifs was observed in the 3'UTRs and CDSs of the highly ranked (high RIP-RNA/T-RNA ratio) genes using cWORDs and some of these intersections were successfully confirmed using luciferase assay. Nevertheless, our study is not without limitations.

As described before, both total-RNA-seq and RIP-seq were performed in cells in which miR-125b was overexpressed between 20 and 100 times and these non-physiological levels of exogenous miRNA has been shown to generate high false positive rates [592]. A quick way to overcome this

problem could be the use of a control in which miR-125b is silenced using oligonucleotides. However, if on one hand the overexpression is problematic for the rate of false positive finders generally due to a non-physiological stoichiometry of the overexpressed miRNA that results in artificial repression of non-target genes [593], the inhibition of specific miRNAs using oligonucleotides can be challenging if the endogenous expression of the miRNA of interest is high [594]. Interestingly, when inhibition of the endogenous levels of miR-15 and miR-16 was performed in parallel to over-expression experiments, the up-regulation of the mRNA expression induce by the inhibition was limited compared to the downregulation obtained by the overexpression [595], suggesting that although the latter gives more artefacts, it still allows an easier detection of the targets.

Overexpression of a specific miRNA can also saturate miRISC and displace other endogenous miR-NAs [596]. This can certainly be a problem when only total RNA-seq is used to identify a miRNA-specific molecular function, however, since RIP and RNA-seq approaches were combined, I strongly think that the method was able to detect the most important molecular functions miR-125b in mouse insulinoma cell line. Indeed, when I measured miR-184 and compared its expression between cells transfected with miR-125b or control mimic, I couldn't detect any differences either before or after immunoprecipitation of the miRISC complex. This indicates that endogenous miR-184, although being less abundant of many other β-cell specific miRNAs including miR-125b, was not displaced by our approach.

Another problem could derive from the fact that RIP-seq approach does not require a crosslink step, thus the mRNA targets at the time of the immunoprecipitation were not bound covalently to AGO2 leading to a missed detection of the ones characterized by a thermodynamically weak binding. However, I might have been able to overcome this problem for genes that are regulated by miR-125b through RNA degradation since the method considered direct targets all the gens with the ratio IP-RNA/T-RNA >1.

An example can be *Bmf*. As mentioned before this gene has been identified as miR-125b direct target in other cells type [581, 597]. In our RNA-seq dataset *Bmf* is significantly downregulated by miR-125b overexpression (Fold-change 0.63 and P = 0.00018), however, in the RIP-seq analysis

the fold-change resulted unchanged (0.95). These data were confirmed also in the analysis performed on selected genes selected before preparing the samples for HT-seq. *Bmf* mRNA levels in the input samples were significantly reduced upon miR-125b overexpression, as well as for *Them6* and *Hnf4g*, however in the IPed samples, where a significant increase of *Them6* and *Hnf4g* was clear, for *Bmf* there was only a tendency, indicating that probably some of *Bmf* mRNA was lost during the manipulation of the samples.

At the end of the analysis, because the results from both RNA-seq and RIP-seq experiments were combined, regardless the fact that *Bmf* was not enriched in the RIP-seq dataset, it was one of miR-125b direct target in β-cells with an IP-RNA/T-RNA of 1.52.

The use of quantitative proteomic strategies has also emerged as a key technique for experimental identification of miRNA targets [598]. Mass spectrometry-based proteomic approaches allow direct determination of proteins whose levels are altered because of translational suppression and, unlike the transcriptome analysis, is also able to identifies targets which suppression doesn't occur as a result of RNA degradation [320, 599]. Moreover, mRNA levels do not necessarily correlate with the levels of protein expression therefore determining only the mRNA expression levels can be not sufficient [600, 601]. Certainly, the combination of proteomic and transcriptomic can be a valid approach to overcome these problematics and gain a better understanding of miRNA role in specific cell types or tissues, however the results obtained could still include indirect targets. Indeed, the addition of mass spectrometry-based proteomic analysis to our experimental approach could have been able to reduce the limitations described above.

In point of fact, I have also been working on the optimization of a method based on the crosslinking of miRNAs and their targets, immunoprecipitation of AGO2-miRNA-mRNA complexes and high throughput sequencing that is able to identify miRNA-mRNA physiological interactions at single nucleotide resolution. For more details see Chapter VI: .

Nevertheless, amongst the 180 direct targets identified TAZ, M6PR, MTFP1, TOR2A and GNPAT were validated using luciferase assay.

TOR2A (Torsin Family 2, member A) protein is a member of the AAA+ ATPase superfamily and reside in the endoplasmic reticulum and perinuclear space [602] and its role in β-cells has not been characterized yet. the precise molecular function(s) of Torsins remain poorly understood also in other cell types, however, studies in animal models or cell-based systems suggest that these proteins are mainly involved molecular chaperone/protein quality control, vesicle/protein trafficking, cellular architecture, and NE vesiculation/trafficking [603].

GNPAT (Glyceronephosphate O-Acyltransferase), like TOR2A, has not been characterized in β-cells, however, its role in other cells has been related to pathways that control metabolism and glycerophospholipid biosynthesis. Enzymatic defects in GNPAT has been shown to disrupt the biosynthesis of plasmalogens [604], a subclass of ether phospholipids that are commonly found in cell membranes in the nervous, immune and cardiovascular systems [605], and result in the accumulation of the fatty alcohols. Plasmalogens can act as a natural antioxidant [606] and lipidomic profiling of multiple populations and clinical cohorts has identified decreased levels of plasmalogens to be associated with obesity [607] pre-diabetes and type 2 diabetes [607].

TAZ, also called Tafazzin, is an enzyme associated with the rare inherited x-linked disorder Barth Syndrome that causes dilated cardiomyopathy, neutropenia and skeletal myopathy resulting in death in male infancy or early childhood from septicaemia and/or cardiac decompensation. The role of this protein in β-cells hasn't been identified yet. Mice containing a doxycycline inducible shRNA against TAZ were subject to a wide spectrum of mitochondria abnormalities in skeletal and cardiac muscles [608-610], whereas knockdown of TAZ in yeast [611], mice [612] and in human iPSCs-cardiomyocytes [613] was associated with increased production of ROS.

MTFP1 (Mitochondrial fission process protein 1), also called MTP18, is an integral protein of the mitochondrial inner membrane that like TAZ, has an important role in the regulation of mitochondrial function [614].

It has been demonstrated that loss of MTFP1 results in a hyperfused mitochondrial reticulum, kidney fibroblast-like cell line whereas its overexpression stimulates fragmentation in mouse and human [614-616]. Recently, this process has been associated to a nutrient-sensing mechanism where the nutrient/energy/redox sensor mammalian target of rapamycin complex 1 (mTORC1)

has been found to stimulate translation of MTFP1 to control mitochondrial fission and apoptosis [615]. MTFP1 was already described as miR-125b target in human monocyte [617]. In these cells, by silencing MTFP1 it was shown to promote the elongation of the mitochondrial network and lead to apoptosis [617]. In the same cell type, miR-125b has been also shown to attenuate the mitochondrial respiration through the silencing of the BH3-only proapoptotic protein [617]. More-over, in the thermogenic beige adipocytes, miR-125b-5p has been shown to negatively regulate mitochondrial biogenesis and play an important role in the repression of beige adipocyte function through the modulation of oxygen consumption and mitochondrial gene expression [618].

AMPK, direct regulator of miR-125b expression in β-cells, also controls different aspects of mito-chondrial biology and homeostasis in other cell types [619]. In skeletal muscle It controls mito-chondrial number through stimulation/inhibition of mitochondrial biogenesis. AMPK activation in fact was found to induce mitochondrial biogenesis through activation of nuclear respiratory fac-tor-1 (NRF-1) [620] whereas inhibition showed a decrease in mitochondrial content [621]. AMPK was also found to promotes mitochondrial fission by phosphorylating mitochondrial fission factor (MFF) [622] and mitophagy thought activation of the kinase ULK1 upon energy stress [623].

By simultaneously regulating mitochondrial biogenesis, fission and mitophagy, AMPK could also maintain mitochondrial health in pancreatic β-cells, and this could be achieved, partially, by sup-pressing miR-125b.Future experiments will need to investigate whether an excessive mitochon-drial fission/fusion turnover caused by the over- or under-expression of miR-125b could induce the formation of defective mitochondria that could then directly impact β-cells functions and mimic T2D pathogenesis.

This might be particularly important because, although only *Taz* and *Mtfp1* have been specifically described here, dozens of other genes that control oxidative phosphorylation and other mitochon-dria functions were identified as direct and indirect targets of miR-125b in β-cells.

Considering that mitochondrial respiration is a major source of reactive oxidative species (ROS) and oxidative stress [624], it is therefore not surprising that defects in mitochondrial function are associated with the development of diabetes [624] and that miR-125b could have a role in the

development or regulation of these processes. Indeed, further investigations are required to elucidate the underling mechanism involved in this regulation.

GSEA of genes down-regulated upon miR-125b overexpression revealed that also lysosomal function together with the secretory pathway, Golgi apparatus and transport were amongst the pathways affected. Not much is known about this regulation in β-cells or other cell types, but it is possible that alterations in the lysosomal function might modulate insulin degradation, whereas changes to the secretory pathway would most likely impact insulin maturation and packaging into mature secretory granules derived from the trans-Golgi network and might contribute to a reduced insulin content.

According to our data, M6PR, a receptor involved in the intracellular trafficking of lysosomal enzymes, is one of the top targets of miR-125b in β-cells. I found that M6PR regulation by miR-125b occurs in both mouse and human β-cells and, at least in mouse, take place through the binding of miR-125b to the position 851-858 of *M6pr* 3'UTR. The role of this receptor has not been investigated in pancreatic β-cells yet, however the global M6PR knock-out mouse model is characterized by elevated levels of phosphorylated lysosomal enzymes in the bloodstream and accumulates lysosomal material in the lysosomes [625]. Interestingly, a similar phenotype has been observed also in Mucolipidosis types II and III, a group of inherited metabolic diseases where abnormal amounts of carbohydrates and fatty materials (lipids) accumulate in cells [626]. The primary defect in this disease is a partial or complete deficiency in the enzyme N-acetylglucosamine-1-phosphotransferase involved in the synthesis of the Mannose 6-Phosphate (M6P) recognition marker on lysosomal enzymes [626].

Thus, we hypothesised that by inhibiting *M6pr*, miR-125b might induce a defect in lysosomal hydrolases, that is responsible for an accumulation of lysosomal structures with non-digested organelles and proteins and could indirectly result in altered Golgi-trafficking and protein processing and altered insulin production.

According to our data miR-125b overexpression in MIN6 cells was responsible for lysosomal abnormally enlarged structures. Lysosomal size has been found altered in several human diseases [627], however, nothing is known about how this factor can influence lysosomal function in β-cell.

In the data presented in this chapter enlarged lysosomes didn't affect lysosomal mediated degradation of GLP-1R. These results indicate that perhaps the altered lysosomal structure is not involved in this specific regulation.

On the contrary, we noticed that, upon miR-125b overexpression, there was an unexpected tendency towards the reduction of SNAP-GLP-1R. This is an exogenous protein expressed by the cells, therefore, miR-125b cannot regulate its expression by targeting the 3'UTR. The regulation could occur through the CDS however, the fact that GLP-1R RNA levels were unchanged upon miR-125b overexpression in the total RNA-seq dataset and it was not enriched in the RIP-seq dataset, indicates that according to our data, GLP-1R is not a target of miR-125b in β-cells and suggests that perhaps miR-125b could regulate this important receptor indirectly. Unfortunately, I was not able to detect the endogenous levels of GLP-1R in these cells due to the lack of the antibody therefore, further investigations are required to elucidate the underling mechanism involved in this regulation.

In β-cells, AMPK has not been directly linked to the regulation of lysosomal function, however, it is known that lysosomes control nutrient sensing and cellular metabolism [583] and that AMPK can promote autophagy during starvation to promote cell survival. In fact, by activating autophagy, the cells recycle nutrients from organelles, proteins, and different macromolecules that are delivered to the lysosomes for degradation [588]. It has been shown that AMPK α1 knockout HEK293T cells fail to induce activation of the autolysosomes mainly due to interference with the fusion of autophagosomes with lysosomes, indicating that AMPK is required for efficient autophagosome maturation and lysosomal fusion [628]. It has also been demonstrated that AMPK activity decreases in T2D [516] and, not surprisingly, loss of autophagy in the β-cell results in an increased incidence of diabetes [629]. Whether elevated miR-125b levels during hyperglycaemia/loss of AMPK activity can contribute to this process remains to be studied, however, it is indeed possible that AMPK might be necessary for the regulation of autophagy and by targeting miR-125b targets involved in the regulation of such process or miR-125b itself we might be able to improve T2D pathogenesis.

Chapter VI: eiCLIP

6.1 Introduction

Approaches that use first crosslinking and then AGO2 immunoprecipitation followed by sequencing are powerful tools that allow the identification of miRNAs direct targets. In the past few years, not only they have been widely and successfully used for this purpose but have also been implemented to overcome all the limitations present in the previous versions.

iCLIP (Individual nucleotide resolution UV cross-linking and immunoprecipitation) can determine the interactions between RBPs and RNA targets at nucleotide resolution [630]. This is achieved by covalent crosslink between RBPs and RNAs obtained using UV-light and purification of the complexes under stringent conditions. More importantly, the UV crosslink causes reverse transcription stalling at the crosslinking site that helps to identify the precise location of the RBP through high throughput sequencing of the cDNA [630] (Figure 6.1).

As shown in Figure 6.1, the single nucleotide resolution of the iCLIP protocol is obtained through the circularization of the cDNA fragments, however it has been often observed that this step is inefficient [631]. Moreover, it has been reported that when iCLIP was performed by the ENCODE consortium at large scale, the generation of valuable libraries was very hard to achieve for many RBPs and even harder for the ones that lacked canonical RNA binding domains [631].

For this reasons improvement of CLIP protocols is necessary to achieve better results in terms of both technical and biological reproducibility, especially in conditions where the amount of available sample is limited or the abundance of RBP of interest is particularly low.

This chapter focuses on the adaptation of enhanced iCLIP (eiCLIP) of AGO2 in human EndoC-βH1 to identify the precise location of miRISC in human β-cells. This technique is a variation of the iCLIP technique developed by our collaborator Dr. Sibley (University of Edinburgh) that doesn't rely on

cDNA circularization to identify the exact position of the crosslinking. By computationally deter-mining the sequences matching β-cell miRNA seed reagions in the proximity of eiCLIP-identified miRISC positions, we aim to define all (or at least, most) the direct target of the β-cell miRNAome.
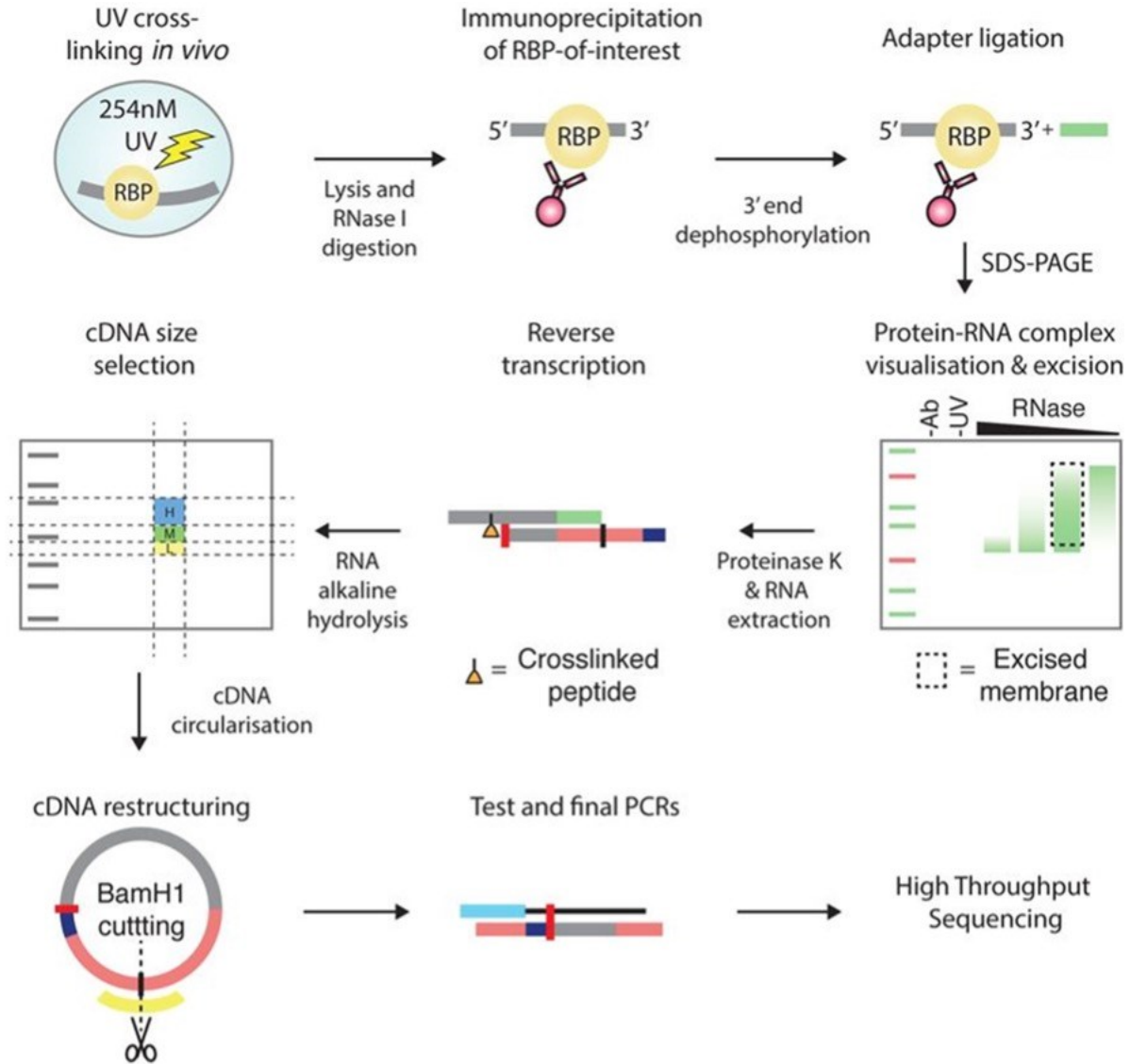
Figure 6.1: Schematic representation of iCLIP protocol
Modified from Sibley et al. [9]

6.2 Results

6.2.1 AGO2-eiCLIP optimization in Hela cells

AGO2-eiCLIP protocol optimization was performed under the guidance of Dr. Sibley. This is a complex, multi-step protocol (Figure 6.2) that is difficult to execute, therefore, we decided to initially
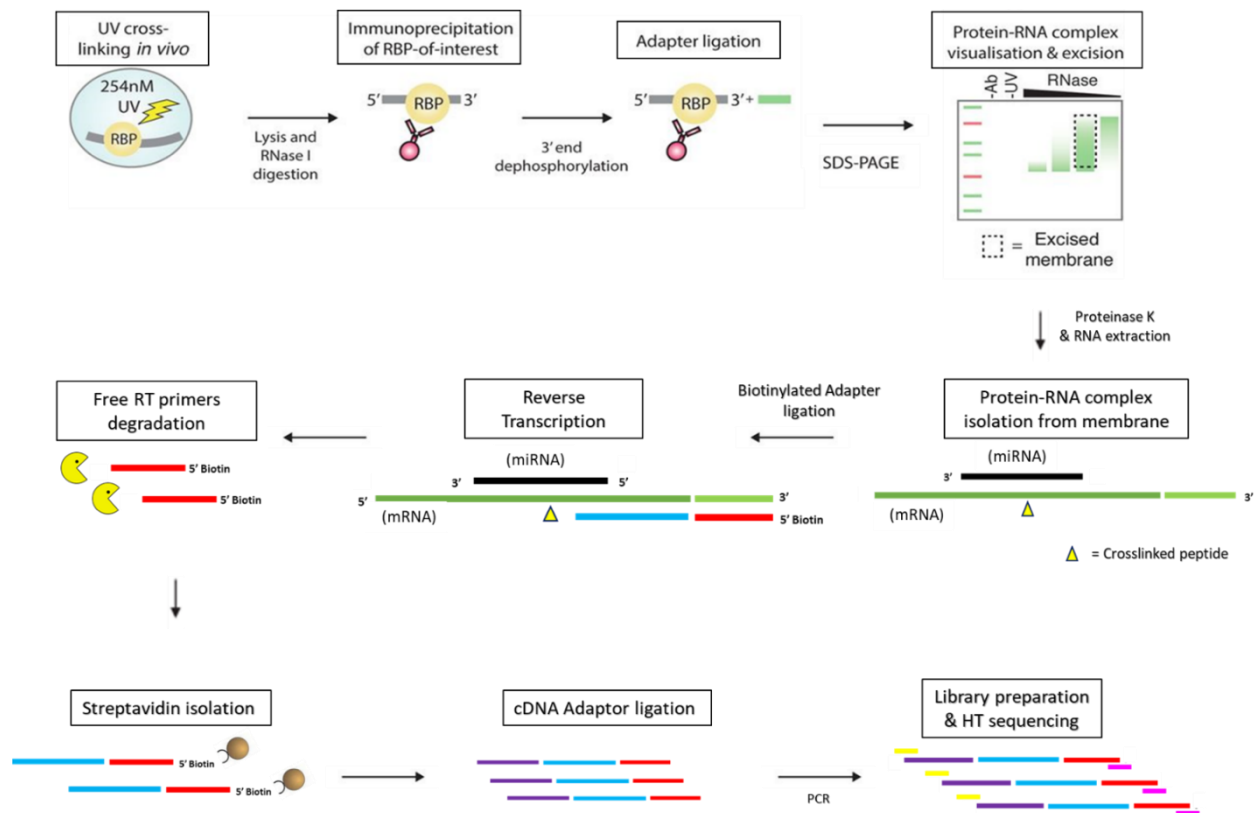
Figure 6.2: Schematic representation of eiCLIP.
Cells are irradiated with UV-C light on ice, leading to formation of a covalent bond between protein and RNA. This is followed by partial RNase digestion and an immunoprecipitation with protein-specific antibodies. For the library preparation and visualization, the RNA is dephosphorylated, a 3' end fluorescent adapter is ligated. The complexes are separated by SDS–PAGE and isolated from a nitrocellulose membrane according to the expected size. The protein is then digested by proteinase K, and reverse transcription (RT) is performed through the ligation of a biotinylated adaptor. cDNA synthesis is truncated at the remaining polypeptide. cDNA molecules are isolated using streptavidin coated beads and free adaptors are removed by enzymatic digestion. cDNA libraries are amplified by PCR using specific index that anneal to both 3' and 5' end before the HT sequencing.

focus ourselves on the optimization of this protocol on human HeLa cells, which have a higher growth rate compared to the EndoC-βH1 cells, using an antibody against the highly expressed RNA binding protein heterogeneous nuclear ribonucleoproteins C1/C2 (hnRNPC) that was then used as positive control in subsequent experiments.

6.2.1.1 UV-Crosslinking

The first step of the protocol requires the covalent binding of the protein-RNA complexes using UV-crosslinking. For this step I simply followed the conditions optimized in Huppertz et al. [630]:cells grown on 15cm adherent plate (~15.0 x 10^6 cells) were irradiated with 150 mJ/cm^2 at 254 nm using a Stratalinker UV crosslinker.

6.2.1.2 Selection of optimal RNAse I treatment

After UV-crosslink, cells were collected, pelleted and washed for subsequent lysis with the aim of exposing the protein-RNA complexes. The RNAs, still bound covalently to the proteins, are then partially digested using a controlled amount of RNAse I enzyme. The size of the RNAs obtained from this step determines the size of the cDNA library produced for the final sequencing analysis, which in turn influences the quality of the data produced by the sequencing. It is therefore important to carefully optimize it.

Multiple factors need to be considered when improving this step:

1) RBP-RNA interactions tend to be quite distinctive; they can be cell-specific and depend on the extract and is thus important to optimize this step for each family of RBPs in a cell-specific manner.

2) The size of the final DNA sequence after all the adapters are ligated in order to obtain the final cDNA library starting from the RNA fragments needs to be within the optimal range for the sequencing step.

3) In the 80% of the cases [632], the cDNA synthesis performed by the reverse transcriptase stops at the position where the short polypeptide has been covalently bound to the RNAs. This means that the cDNA molecules produced by this step are half the size of their corresponding RNAs [630].

187

Taking into account the above, I determined that RNA size between 50 and 300 nucleotides was optimal to obtain cDNA libraries of 100-250 nucleotides. Different RNAse I dilutions, ranging from 1:10 to 1:2000, were tested and samples were subsequently immunoprecipitated using a specific antibody against hnRNPC. Dephosphorylation of the 3' end of the RNAs fragments and subsequent ligation of a fluorescent adaptor enabled hnRNPC-RNA complexes visualization on nitrocellulose membrane after SDS-PAGE.

Figure 6.3A (upper panel) displays the image taken using the fluorescent imager Odyssey LI-COR CLx, which can detect infrared fluorescence. From the Figure it is possible to identify both hnRNPC (~50kDa) and its dimers (~100 kDa) bound to the florescent RNA molecules that are detected by the imager as a red smear. Crucially, no signal was observed in the two negative controls in which no antibody (No Ab) was used for the immunoprecipitation or no UV crosslink (No UV) was performed on the cells. It is also clear that the smear that represents the RNA molecules is inversely proportional to the concentration of RNAse I used for the digestion. Given that the average molecular weight of 70nt of RNA has been estimated to be ~20 kDa and that the fluorescent adaptor is ~45 nt long, the ideal position of RNA-protein complexes that will generate RNA fragment of 55 nt is ~35 kDa above the expected molecular weight of the protein of interest.

As shown in Figure 6.3B, RNA fragments of ~50, ~ 100 and ~300 nt were detected when using 1:250, 1:500, 1:1000 and 1:2000 RNAse I dilutions, respectively, whereas no RNA fragments were observed when samples were treated with a 1:10 dilution. This indicates that the fragments bound to hnRNPC that emitted infrared fluorescence signal after SDS-PAGE (Figure 6.2A , 3rd lane) were smaller than 25 nt, and therefore not detectable on the UREA gel.

At this point, RNA molecules need to be released from both the nitrocellulose membrane and the hnRNPC proteins. The first was achieved using a cutting mask that was produced by drawing a box around the protein-RNA signal that started ~15Kda above the molecular weight of hnRNPC. For this test the RNAs bound to hnRNPC dimers was not considered. Using a mask, I cut the corresponding nitrocellulose membrane pieces. On the other side, the isolation of the RNA fragment from hnRNPC was obtained using proteinase K. The RNA was then purified using phenol chloroform and the size of the fragments assessed by running the sample on a urea gel (Figure 6.3A,

lower pannel). RNAse I 1:1000 dilution was identified as the optimal dilution to obtain RNA fragments between 50 and 300 nt.
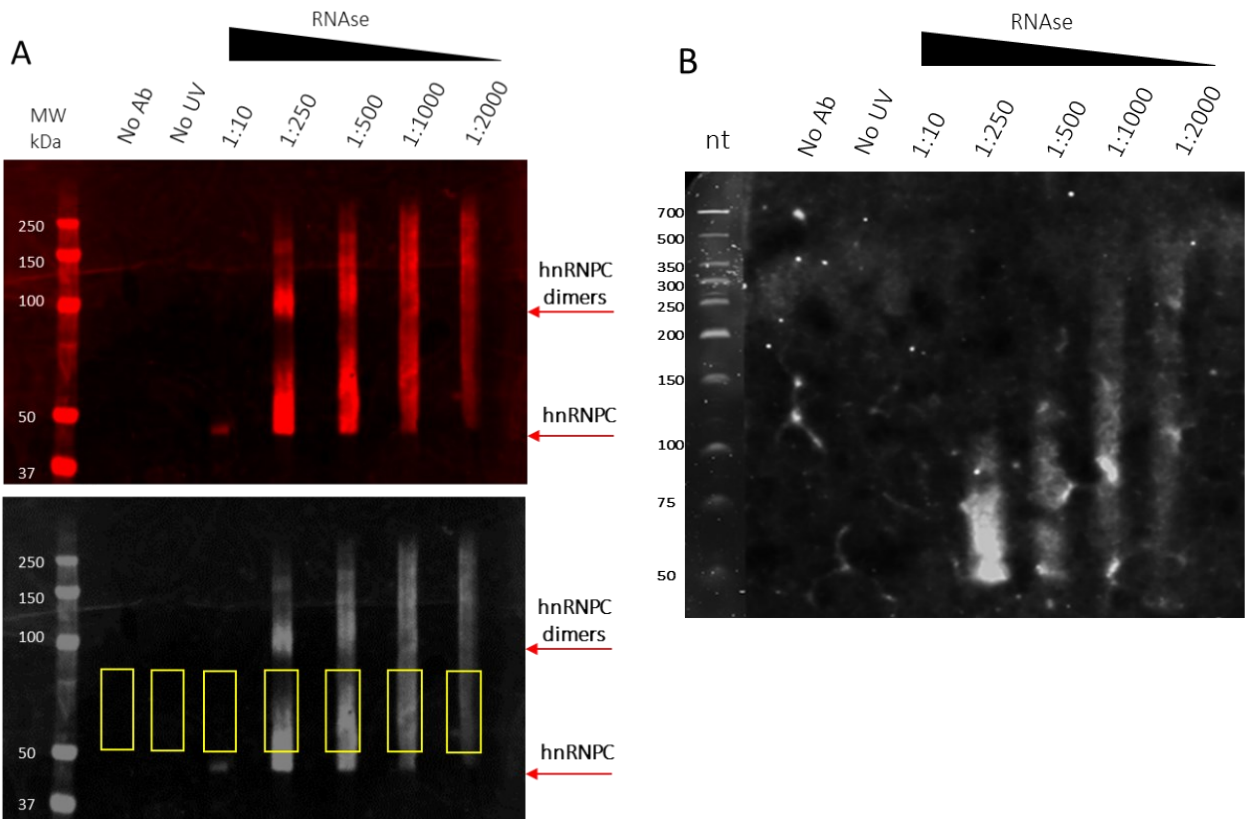
Figure 6.3: hnRNPC eiCLIP protein-RNA complex visualisation and RNase digestion analysis in Hela cells
A. Upper panel shows fluorescent adapter labelled RNA-protein complexes visualised following SDS-PAGE using a Li-cor imager in Hela cells. Samples include no antibody (No Ab) and UV-crosslinking (no UV) negative controls, and a gradient of RNAse I concentration/dilutions from which suitable digestions patterns can be determined. hnRNP C antibody was used for the immunoprecipitated and the arrows indicate its monomers (~50 kDa) and dimers (~100 kDa). Lower panel shows grayscale version with solid yellow markers that indicate regions excised from the membrane using a cutting mask.
B. RNA size distributions from the different regions and different RNase concentrations cut from Figure 6.3A. Extracted RNA was denatured and run for 40 minutes at 180V on a 6% TBE-UREA gel.

6.2.1.3 AGO2-eiCLIP in Hela cells

The same conditions were applied to Hela cells samples when the immunoprecipitation was performed using an anti-AGO2 antibody. Considering how miRNAs repress their targets, after crosslinking, our protein-RNA complexes were expected to be mainly formed by AGO2 bound to both

the miRNA and its target RNA. For this test I only used two dilutions of RNAse I (1:10 and 1:1000) and included as positive control a sample in which the immunoprecipitation was performed using hnRNPC.

As showed in Figure 6.4 A (upper panel), RNAse treatment generated a diffused signal of protein–RNA complexes starting from AGO2 molecular weight (100 kDa). Consistently with previous results, low dilution of RNAse I enzyme resulted in an almost complete degradation of the RNA, whereas dilution 1:1000 resulted in RNA fragments of the desired size. As expected, both No UV crosslink (no UV) and no antibody (No Ab) negative controls resulted in no visible RNA. The positive control, in which hnRNPC was used for the immunoprecipitation, showed a much higher efficiency compared to the samples immunoprecipitated using AGO2.

In order to further optimize the protocol, an additional control called size matched input (SMI) was added. This consists of RNA extracted from the protein-RNA complexes present in the sample before the immunoprecipitation and is used to normalize the sequencing results for nonspecific background signal coming from a part of the membrane that has an identical size range of our RNAs of interest. The addition of this control improves the signal-to-noise ratio of the technique. Pieces of the membrane containing the fragmented RNA were then excised using the method described in the previous section (Figure 6.4 A, lower panel).

Since our collaborator's lab has been working on further refinement and optimization of the protocol, from this step onward the protocol will differ from the one previously published for iCLIP. In fact, this new protocol skips both the size selection of the cDNA products using gel electrophoresis and the circularization of the cDNA that is not necessary to obtain single nucleotide resolution. Instead, after RNA extraction, the cDNA synthesis is performed using biotinylated primers that perfectly anneal to the fluorescent adaptors on the 3' end of the RNA molecules. Then, the newly formed DNA molecules are isolated using streptavidin and the free adaptors are removed from the sample through a RecJ enzyme digestion. At this point, another adaptor is ligated to the 5' end of the cDNA molecules and then specific indexes, able to anneal to both 3' and 5' end, are used to amplify the cDNA libraries by PCR (Figure 6.4 B). Notably, each index is characterized by a unique barcode that helps to distinguish different samples when pooled together for the sequencing step.

Importantly, in order to obtain a single nucleotide resolution eiCLIP relies on the fact that the cDNA synthesis started by the biotinylated primer terminates at the position where the RBP was covalently bound to the RNA target during the crosslink step.

Figure 6.4 B reports the cDNA libraries amplified by qPCR. As expected, no libraries were obtained from the negative controls No Ab (1A and 1B) and No UV (2) and from the sample treated with high amount of RNAse I (3) due to the absence of RNA in the samples. What we indeed obtained as expected was detectable cDNA libraries from the samples treated with optimal RNAse I concentrations and precipitated with AGO2 and hnRNPC antibodies (4 and 5, respectively), as well as the SMI. The libraries produced from the AGO2 immunoprecipitated sample (4) and its matching SMI control (6A) were characterized by bigger DNA fragments (150-400nt) when compared to the hnRNPC immunoprecipitated sample (5) and its correspondent SMI (6B) where the fragments varied from 150 to 250 nt. In order to obtain the best compromise between the over-amplification of secondary products, which normally appear above the expected library size, and a too weak amplification of the band of interest that results in low number of reads detected by the sequencer [633], the optimal number of PCR cycles was carefully identified using Figure 6.4 B as a guide. For these samples, 21 cycles were used.

For the sequencing analysis that used a NextSeq500 platform, we combined the cDNA library produced for AGO2 in this experiment with another one produced in a second independent experiment and pooled both with several other libraries generated by our collaborator. Sequencing results indicated that ~4.5 million reads were obtained with an average length of 41 nt. Of those, ~2 million reads (45.7%) were uniquely mapped to the Homo sapiens (human) genome assembly GRCh38 (hg38).
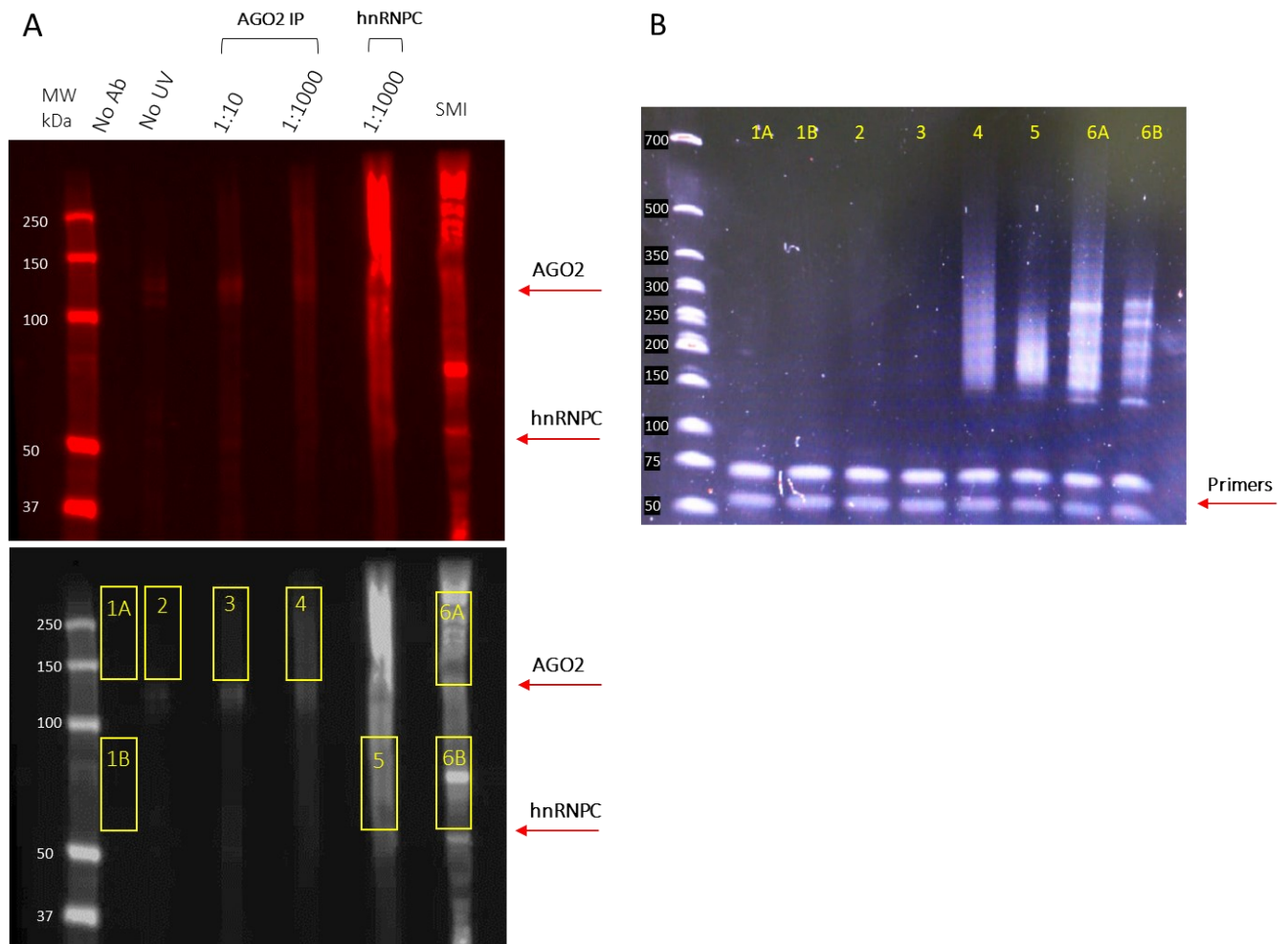
Figure 6.4: AGO2-eiCLIP protein-RNA complex visualisation and cDNA library PCRs in Hela cells
A. Upper panel shows fluorescent adapter labelled RNA-protein complexes visualised following SDS-PAGE using Li-cor imager in Hela cells. Samples include no antibody (Non Ab) and no UV negative controls, 1:10 and 1:1000 RNase I conditions for samples immunoprecipitated with AGO2 antibody (AGO2 IP), 1:1000 RNase I conditions for positive control immunoprecipitated using hnRNPC antibody and size matched input (SMI) sample. The arrows indicate AGO2 (~100 kDa) and hnRNPC (~50 kDa). Lower panel shows grayscale version with solid yellow markers that indicate regions excised from the membrane using a cutting mask. B. cDNA libraries prepared using RNA extracted from the regions cut from Figure 6.4 A. 22 cycles of PCR were performed to amplify these libraries.

6.2.1.4 AGO2 crosslinking occurs between 5-10 nucleotides from miRNA's seed region

Our collaborator Nejc Haberman (MRC-LMS, Prof. Lenhard's lab) computationally interrogated the 60 nucleotide sequences around the AGO2 crosslinking sites to identify sequences complementary to the 7-nucleotide seeds of the most abundantly expressed miRNAs in HeLa [490]. The sequence enrichment of the 7-mer reverse complements were plotted as an heatmap (Figure 6.5) where the strongest crosslink position, or rather the ones that had the maximum number of reads (maximum coverage), were used as the reference point. Maximum coverage was observed at a position 5-6 nucleotides upstream of the position complementary to most HeLa miRNA-seeds, suggesting that the technique was successful in identifying miRNA binding site locations.
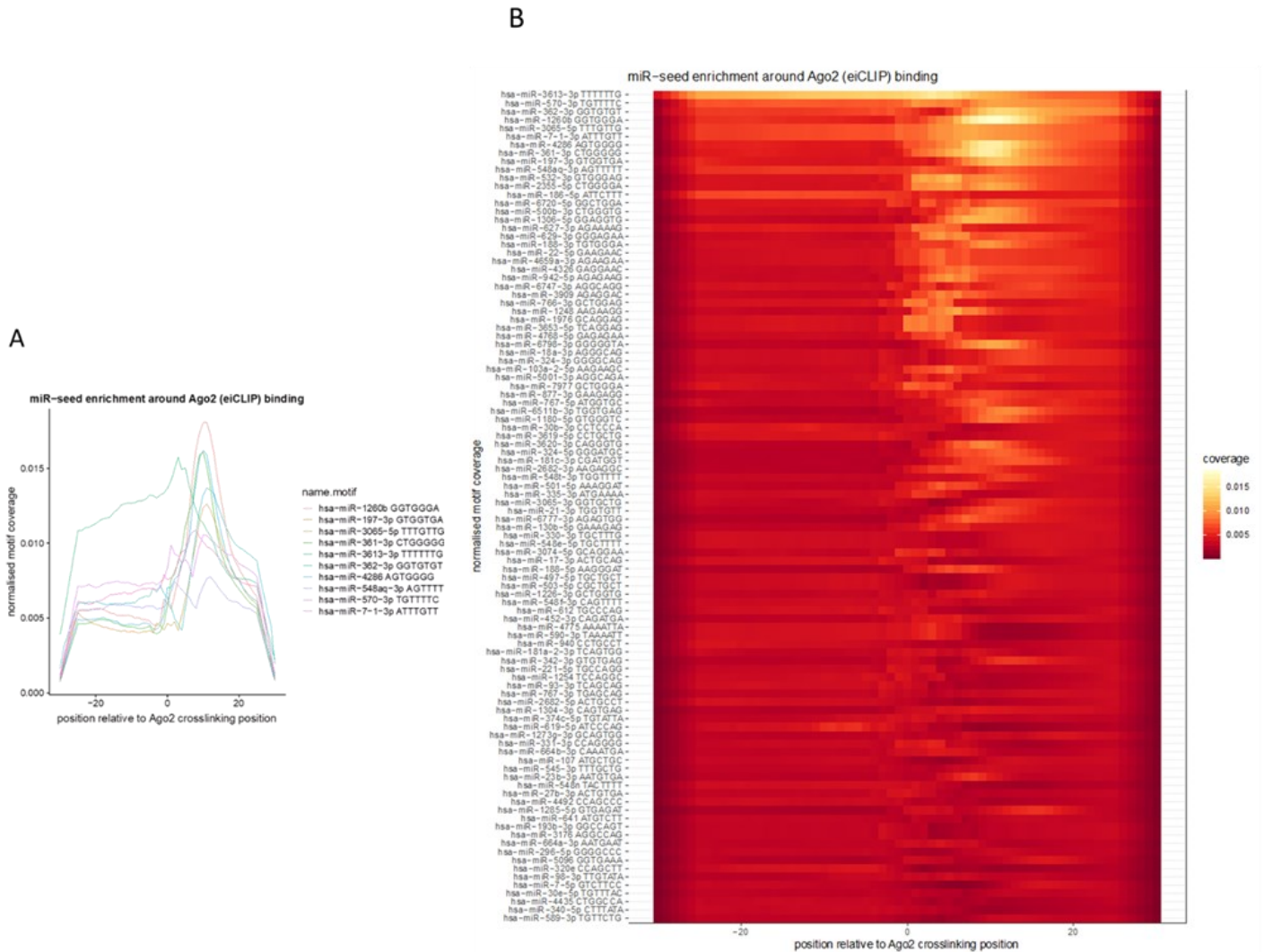
Figure 6.5: Crosslink between AGO2 and miRNAs occurs at 5-6 nucleotides from seed region

A. Kmer graph showing the miR-seed enrichment around AGO2 crosslinking for the 11 most abundant miR-NAs in Hela cells. The Y axis indicates the position (± 30nt) relative to the AGO2 crosslink site. The X axis represents the number of reads (Coverage) identified for each motif (defined as 6 nucleotides complementary to the miRNA seed)

B. Heatmap showing the miR-seed enrichment around AGO2 crosslinking for the most abundant miRNAs in Hela cells. The X axis indicates the position (± 30nt) relative to the AGO2 crosslink site. The Y axis shows the name of the miRNAs and seed motif. The colours indicate the relative number of reads (coverage) identified as complementary to miRNA seed regions, obtained for each position relative to AGO2 crosslinking (the lighter, the higher).

Bed and BedGraph files containing AGO2-crosslink positions were uploaded for the visualization of the coordinates of the AGO2 crosslink and the coverage depth (peaks), respectively, on UCSC Genome Browser. Figure 6.6 shows the results obtained for the 3'UTR of CDKN1B, a gene that encodes for the cell cycle regulator cyclin dependent kinase inhibitor 1B (p27[Kip1]).

The major function of p27[Kip1] is to arrest cell cycle and it is therefore generally considered a tumour suppressor [634]. The physiological role of p27 [Kip1] is determined by different post-translational modifications [635], for example phosphorylation regulates p27 [Kip1] capacity to bind to cyclin-CDK complexes and promote their inhibition [636]. However, p27 [Kip1] activity can also be regulated by changes in its subcellular localization [637], mRNA translation and/or ubiquitin-mediated prote-olysis [638]. Not surprisingly, in cancer, p27 [Kip1] is often inactivated via impaired synthesis, accel-erated degradation, or mis-localization [635].

MiRNAs can also contribute to the cell-cycle dependent regulation of p27[Kip1] expression [639]. Different miRNAs are predicted to bind CDKN1B 3'UTR (Figure 6.6 A) and some of these predic-tions have been experimentally validated in different tumour cells lines [640], including human cervical cancer Hela cells [641, 642]. Our AGO2-eiCLIP data confirmed that different AGO2 cross-link sites were found in the CDKN1B 3'UTR (Figure 6.6 B). As an example, Figure 6.5 B shows the coordinates of AGO2 crosslink site in proximity of the miR-221 binding site, one of the first miRNA to be identified as an endogenous regulator of p27[Kip1] protein expression through the inhibition of CDKN1B 3'UTR [642].
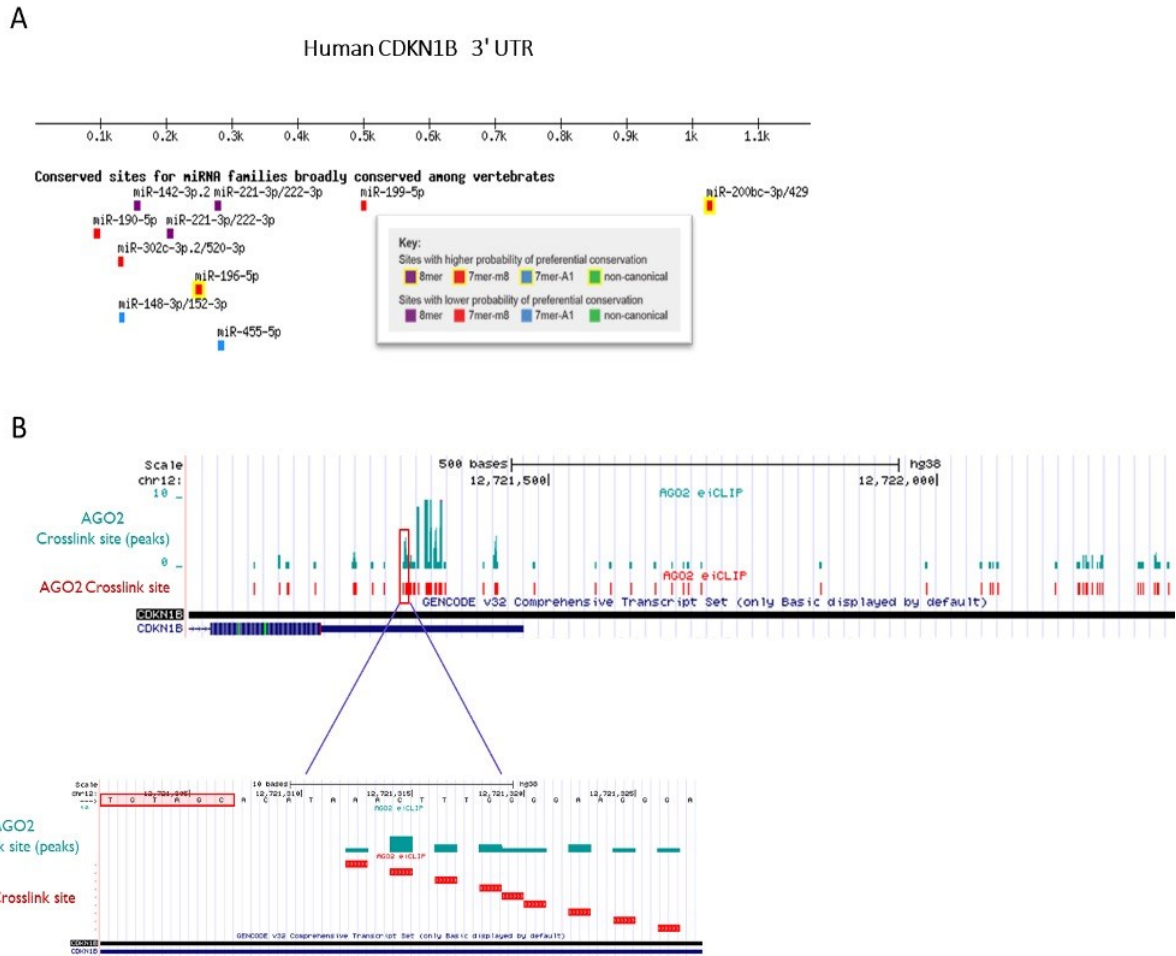
Figure 6.6: Human CDKN1B locus presents different binding sites for miRNAs
A. Predicted binding site of miRNAs in the 3'UTR of *CDKN1B* based on TargetScan 7.2 database.
B. Genomic region correspondent to the 3'UTR of CDKN1B. In red are reported the coordinates of AGO2 crosslink sites detected with eiCLIP. The green peaks show the coverage depth of sequencing for each AGO2 crosslink sites.
C. Zoom-in of the CDKN1B 3'UTR region that encodes for the binding site of miR-221. The solid red rectangle indicates the region that transcribes for the 7nt binding site sequence recognized by miR-221.

6.2.2 AGO2-eiCLIP in EndoC-βH1

Finally, I used the optimized eiCLIP protocol for AGO2 immunoprecipitation in EndoC-βH1 cells. No UV and No Ab conditions were used as negative controls, whereas AGO2 immunoprecipitation in Hela cells (Hela AGO2 IP) and hnRNPC in EndoC-βH1 (EndoC-βH1 hnRNPC IP) were used as positive controls.

As shown in Figure 6.7A (upper panel), for both Hela and EndoC-βH1 cells low dilution of RNAse I enzyme (1:1000) generated a smear of protein–RNA complexes starting from AGO2 molecular weight (100 kDa), whereas the signal was almost absent when the cells were treated with lower dilutions of the enzyme (1:10). Infrared fluorescence was also absent in No UV and No Ab negative controls, as expected. The fragments of membrane with the RNA of interest were excised (Figure 6.7, lower panel), RNA was extracted from them as described earlier and cDNA libraries were generated. As shown in Figure 6.7B, with an amplification of 25 cycles I was able to produce AGO2-eiCLIP libraries from EndoC-βH1 of a size that is comparable to those obtained in Hela cells.
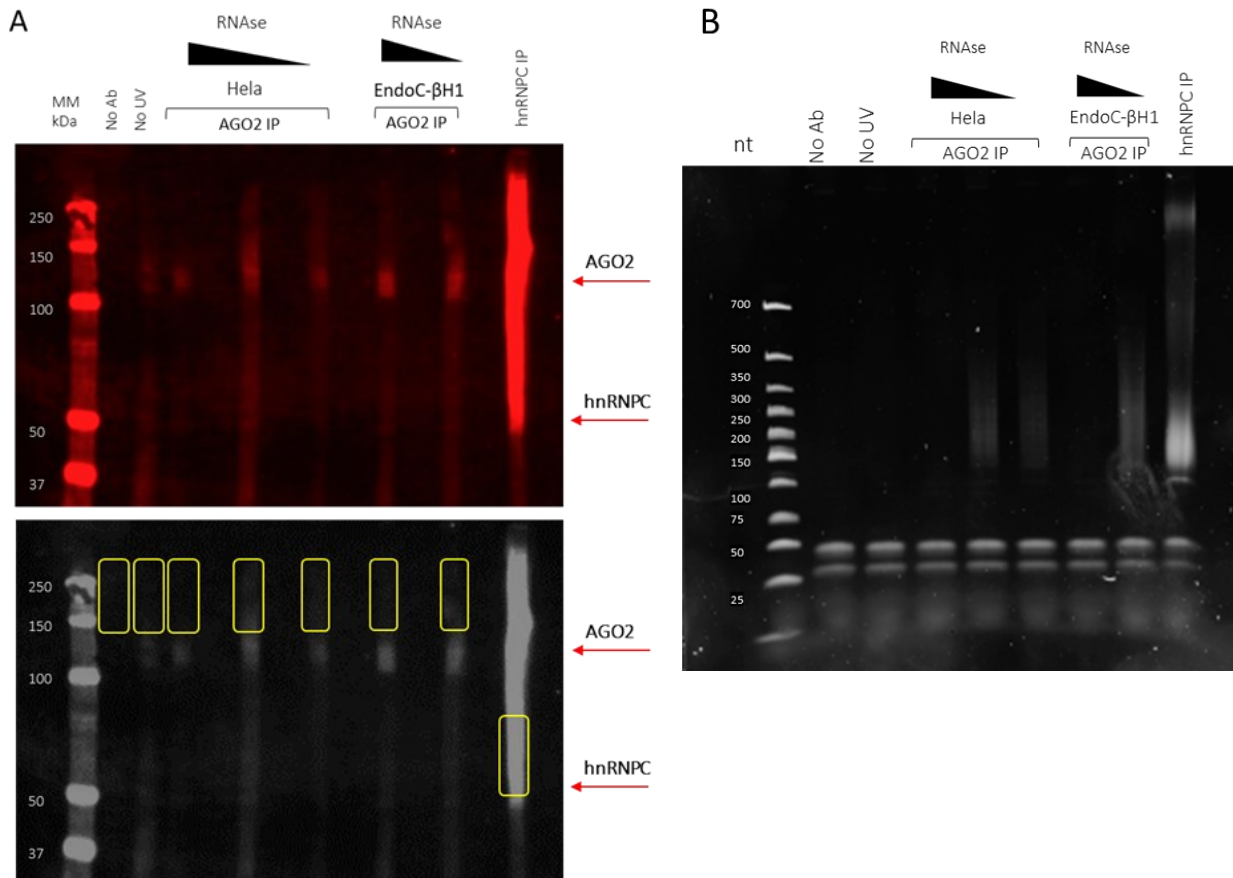
Figure 6.7: AGO2-eiCLIP protein-RNA complex visualisation and cDNA library PCRs in EndoC-βH1 cells

A. Upper panel shows fluorescent adapter labelled RNA-protein complexes visualised following SDS-PAGE using Li-cor imager in Hela and EndoC-βH1. Samples include no antibody (Non Ab) and no UV negative controls, 1:10 and 1:1000 RNase I conditions for samples immunoprecipitated with AGO2 antibody (AGO2 IP) in Hela cells, 1:10 and 1:1000 RNase I conditions for samples immunoprecipitated using AGO2 antibody and positive control immunoprecipitated using hnRNPC antibody. The arrows indicate AGO2 (~100 kDa) and hnRNPC (~50 kDa). Lower panel shows grayscale version with solid yellow markers that indicate regions excised from the membrane using a cutting mask.

B. cDNA libraries prepared using RNA extracted from the regions cut from figure 6.7 A. 25 PCR cycles were performed to amplify the libraries.

6.3 Discussion

Different techniques that use the immunoprecipation of AGO proteins to directly capture miR-ISC-bound mRNAs have become available in the past few years (see section 1.6.9).

RIP-based approaches described in Chapter V: are one of the first methods used to identify miR-NAs direct targets. Although straightforward and useful, RIP-based approaches are characterized by a significant false positive discovery rate due to the immunoprecipitation of RNAs that can non-specifically interact with proteins and that are not eliminated because of the low stringency puri-fication protocols. Conversely, loose interactions can be lost during the immunoprecipitation and important targets remain undetected due to the lack of a covalent binding of Protein-RNA com-plexes. Moreover, these techniques are typically performed following overexpression/inhibition of the miRNA of interest, another approach that can lead to false results due to a displacement of other miRNA-mRNA interactions that occur in the cells under normal conditions.

For these reasons, stabilization of RNA-protein interaction through covalent binding has become increasingly popular. Crosslinking and Immunoprecipitation methods (CLIP) use UV to induce pro-tein–RNA crosslinks preceding immunoprecipitation of miRISC. Many different CLIP assays have been developed in the past few years (see section 1.6.9 , however, even with the more sophisti-cated ones, data reproducibility is still an issue that can be strongly affected by technical differ-ences between various protocols [643] and by the potential false positives resulted from back-ground binding [644].

iCLIP is one of the most recently developed CLIP-based protocols and can be used to determine the interactions between RBPs and RNA targets at single-nucleotide resolution [630]. As previously mentioned, this result is achieved by the stalling of the protein at the crosslink site during the reverse transcription step and subsequent circularization of the cDNA fragments. However, it has often been observed that this step is inefficient [631]. Moreover, the generation of good quality libraries has also been found very hard to achieve for many RBPs and even harder for the ones that lacked canonical RNA binding domains [631].

Dr Chris Sibley's group has been working on the optimization of the iCLIP protocol and introduced a new version named eiCLIP, with the aim to achieve better results in terms of both technical and biological reproducibility, especially in conditions where the amount of available sample is limited or the abundance of the RBP of interest is particularly low. Under his guidance, I worked in parallel to try and optimize this more sophisticated CLIP technique in order to identify all miRNA-mRNA interactions in human β-cells using AGO2 immunoprecipitation.

Due to the complex multi-step protocol that characterizes this technique, we decided to start the optimization using high-proliferative human HeLa cells. The results in this chapter show that we were able to successfully produce, and sequence, libraries obtained from eiCLIP using an anti-AGO2 antibody for the immunoprecipitation despite the fact that the RNA recovered was very limited compared to when an antibody against the abundant RBP hnRNPC was used. This is because miRNA molecules are positioned between AGO2 and the mRNA targets and this makes the crosslink less efficient. MiRNAs can also interfere with the reverse transcriptase, making the cDNA generation less efficient.

The data analysis carried out by our collaborator Dr Haberman identified that, in Hela cells, AGO2 crosslink occurs consistently at 5-10 nucleotides from the miRNA seed region. When we looked at the *CDKN1B* locus, a gene that in Hela cells has been shown to be highly regulated by miRNAs, our AGO2-eiCLIP data confirmed that different AGO2 crosslink sites were found within the CDKN1B 3'UTR. Interestingly, also CDKN1B 5'UTR and CDS are characterized by this crosslink sites indicating that the regulation by miRNAs might also occur through these regions.

Nevertheless, it is also possible that the regulation of 5'UTR and CDS might be miRNA independent. It has been shown that AGO2 is also present in the nucleus [645] where it regulates gene expression not only at post-transcriptional level but also by regulating chromatin modifications and alternative splicing [307, 646]. Ameyar et al. [307], for example, described a mechanism by which AGO2 might act through the methylation of histone H3 lysine 9 (H3K9) and the recruitment of the spliceosome complex to regulate gene expression. Nevertheless, the details of this mechanism are still largely unknown, and more investigations are needed.

Argonaute proteins can also regulate transcription by inducing gene silencing [647, 648]. This mechanism is well known in Schizosaccharomyces pombe [649] where AGO proteins are only involved in the epigenetic regulation and not in the post-transcriptional gene silencing machinery [650]. In the past few years, it has been shown that AGO2 has retained this so-called transcriptional gene silencing (TGS) role in mammalian cells and it is necessary for the inhibition of gene transcription mediated by 21-base duplex RNAs called small interfering antigen RNA (agRNA) that are normally complementary to gene promoters and can inhibit gene expression by silencing transcription rather than translation [294].

Recently, AGO2 has also been found to interact with the ATP-dependent chromatin remodelling complex SWI/SN around TSSs of target DNAs while loaded with a novel class of short RNA called swiRNA [651]. The authors of this study suggested a role for AGO2 in the establishment of nucleosome occupancy on target TSSs, implying that this protein could repress and/or activate transcription [651]. Epigenetic changes play important roles in the normal physiological function of pancreatic β-cells indicating that their alteration could contribute to the development of diabetes [652].

It is important to mention that, although AGO2 is the most studied Argonaute [294] and is the only AGO to function as an endonuclease in mammals [295, 296], in humans all four AGO proteins (AGO1–4) have been associated with almost indistinguishable sets of miRNAs [297-299]. Therefore, by only analysing the RNAs bound to AGO2, the data described in this chapter lack of information about mRNA targets that are normally silenced through other members of the AGO family.

Moreover, we have also not included the SMI meaning that the signal obtained from the sequencing was not normalized for sequences that have been unspecifically enriched during the immunoprecipitation, library preparation or by sequencing bias. It is therefore possible that some of the crosslink sites identified by the bioinformatic analysis might not specific.

Importantly, the AGO2-eiCLIP protocol optimized in Hela cells was successfully applied also in EndoC-βH1 obtaining cDNA libraries with a comparable size to the one produced in Hela cells. Our future aim is to produce more technical replicates that can be therefore sequenced and include

the SMI sample that will be used to normalize the sequencing results as a function of the quantified background signal.

We are also considering improving the protocol by introducing a further step named crosslinking ligation and sequencing of hybrids (CLASH)[653] where the ends of RNAs associated with the protein of interest are ligated together. This method has already been used to determine miRNA-mRNA specific interactions [653]. By using the CLASH protocol, Helwak et al.[653] were able to identify ∼15,000 unique miRNA-mRNA interaction sites. Although this might appear a conspicuous number, the authors concluded that it was not enough to obtain a reliable comparison of alterations in miRNA targets under changing physiological conditions. They in fact estimated that the number of chimeric reads obtained from the sequencing data was quite variable and lower than 2% of all sequencing reads, mainly because of one of the biggest limitations of the protocol which is the low efficiency of RNA-RNA ligation. Further optimizations are thus required in other to adopt this new approach.

Chapter VII: Final Discussion

Diabetes is a complex chronic disease characterized by increased level of glucose in the blood [89]. The most common type of diabetes is T2D, representing almost 90% of all the cases. High blood sugar (hyperglycaemia) affects people who have diabetes and is the result of a combination of insulin resistance in skeletal muscle, liver and adipose tissue and an insufficient secretion of insulin from pancreatic β-cells [144]. During the early stages of diabetes, β-cells are still able to compensate from peripheral insulin resistance releasing more insulin. Nevertheless, in frank diabetes, β-cells lose this compensatory capacity, leading to chronic hyperglycaemia [145]. Diabetic complications are an important cause of disability, reduced quality of life and death and have heavy impact on health care costs [89].

AMP-activated protein kinase (AMPK) is a protein that acts as a nutrient and energy sensor, it is involved in the maintenance of cellular energy homeostasis and its activation is the suggested mechanism of action of some anti-diabetic drugs [654]. At systemic level AMPK activation promotes glucose uptake in muscle [654] and shuts down glucose production by the liver during feeding. In β-cells, AMPK plays a role in the control of insulin secretion [192, 245, 247], as proven by the fact that specific deletion of both AMPK catalytic subunits α1 and α2 (AMPKdKO) in mice caused altered glucose-stimulated insulin secretion [247]. Nevertheless, the effect of long-term activation and the mechanisms of action of this enzyme in the β-cells is still contested.

MiRNAs, small non-coding RNAs that silence gene expression post-transcriptionally, have also been found altered in T2D [397]. MiRNAs expression can change in response to fluctuations in glucose levels and this highlights the importance of these little molecules in the rapid physiological response to nutrient-induced insulin secretion [389]. Molecular mechanisms by which miRNAs are altered in T2D or hyperglycaemia are not known but recent data from our group show that AMPK plays a key role in the glucose-mediated regulation of miR-184 [458] and miR-125b (unpublished).

Many studies have identified the important functions of miRNAs and have shown that the rescue or inhibition of aberrant miRNA gene expression patterns could be a valid therapeutic approach

for different diseases (Table 1.3). However, paradoxically, transcriptional regulation of miRNAs, which is a critical step for the regulation of their expression, remains poorly understood [655]. This is due to previous limitations in methodology available to study primary miRNA transcripts combined with the fact that annotation of miRNA promoters is incomplete [655]. Interestingly, data from our group indicate that AMPK might regulate miR-184 [458] and miR-125b (Unpublished) at the transcriptional level. Importantly, miR-184 has already been shown to play an important role in the regulation of insulin secretion and β-cell proliferation during compensatory expansion in pregnancy and obesity [397, 400]. On the contrary, very little is known about miR-125b function in β-cell.

Therefore, with this PhD project I aimed to widen the understanding of the mechanisms underlying glucose and AMPK-mediated regulation of miR-184 and miR-125b and to elucidate miR-125b function in pancreatic β-cells.

I found that in MIN6 cells CTCF binds a region ~25 kb upstream of MIR184, however the regulation of this TF binding is not affected by glucose (low sugar vs. normal diet) or the absence of AMPK (LKB1KO) indicating that neither glucose nor AMPK regulate miR-184 transcription through CTCF. I also found that both glucose and AMPK positively regulate CTCF protein expression in MIN6 cells. Since glucose, in turn, also represses AMPK activity, it is possible that this regulation is a mechanism used by the cells to maintain CTCF activity balance.

In order to identify MIR184 TSS, we have taken advantage of several published datasets of ChIP-seq histone modifications in mouse islets and combined them with SLIC-CAGE data obtained by our collaborators. Using this approach, I identified a region located ~121 Kb upstream of MIR184 (chr9:89802260-89802328) that presents modifications associated with actively transcribed promoters (SLIC-CAGE/H3K4me3) and is positioned within chromatin regions significantly less accessible in LKB1KO islets. Using this approach, I was also able to establish that the region 78 kb upstream of MIR184, which was proposed by others as a potential mir-184 promoter [495], is an enhancer. More studies are needed to confirm this region as a miR-184 promoter. However, the reduced accessibility in LKB1KO islets could indicate that, as hypothesised, AMPK might induce

miR-184 expression at transcriptional level through the regulation of a TF binding. By understanding this regulation, we could also learn more about the mechanisms that characterize the activation of AMPK in β-cell. As mentioned before, its activation is a suggested mechanism of action of some anti-diabetic drugs [654], however the effect of long-term activation and the mechanisms of action of this enzyme is still contested in β-cell. MiR-184 expression in human pancreatic islets is considerably higher if compared to other tissues such as liver and skeletal muscle [467], indicating that some of the effect of this AMPK-mediated regulation might be β-cell specific. It has been shown that its expression is downregulated in prediabetic and diabetic mice [397] and its silencing is triggered by insulin resistance to increase proliferation and accommodate for the elevated demand for insulin [400]. The activation of miR-184 transcription mediated by AMPK might have the opposite role. Indeed, in MIN6 cells, miR-184 has been found to inhibit insulin secretion by targeting Slc25a22, a glutamate transporter that plays a role in the control of GSIS [468] whereas, in AMPKdKO islets, Slc25a22 expression was found significantly increased [247]. The link between AMPK, miR-184 and Slc25a22 in the control of insulin secretion has not been investigated yet, however it is possible that, by inducing miR-184 expression, AMPK might inhibit insulin secretion. This axis could also explain the increased secretion of insulin observed in isolated AMPKdKO islets [247].

Unfortunately, I was not able to clearly decipher the mechanism involved in the AMPK mediated regulation of miR-125b expression. Indeed, I found that TGF-β positively regulates this miRNA in human islets, however, in MIN6 cells, one of its downstream pathway effectors, SMAD3, had the opposite role. In MIN6 cells, SMAD3 was positively regulated by AMPK, whereas its inhibition resulted in a small increase of miR-125b expression. These data suggest that AMPK-mediated regulation of SMAD3 might contribute towards the inhibition of mir-125b expression. Since the effects observed were very small, we also concluded that most likely AMPK represses miR-125b expression through a SMAD2/3 independent pathway, which remains to be identified.

Interestingly, it has been shown that SMAD3 overexpression is responsible for a reduced GSIS in pancreatic islets [563], whereas its constitutive activation increases apoptosis and induces β-cell mass loss. Although the role of AMPK in these two processes is still controversial, the fact that

when activated it negatively regulates GSIS [242, 243] and promotes apoptosis [656] suggests that AMPK could regulate these pathways through the positive regulation of SMAD3. Further research would be necessary to corroborate this hypothesis, especially since the effect of AMPK activation on Smad3 was small.

High levels of circulating miR-125b have been associated with hyperglycaemia (HbA1c) in prediabetic [469], T2D [469] and T1D subjects [424] and its expression in islets correlates with the BMI of the donors. These data, together with our results showing that islet miR-125b is regulated by glucose via AMPK, suggest miR-125b as a contributor to the development of diabetes. Nevertheless, miR-125b function in β-cell has not been fully characterized yet. According to our group preliminary data, miR-125b overexpression in MIN6 cells significantly reduced insulin content, impaired GSIS and affected MIN6 cell morphology. At the start of this project, only two gene targets had been proposed in β-cells: cMaf, a positive regulator of glucagon expression [574] and, more recently, Dact1, a member of JNK signalling pathway [575].

As mentioned before, the ability of miRNAs to regulate multiple genes within a molecular pathway [566] makes them excellent novel targets for the treatment of disease. However, miRNA-target interactions can differ substantially between organisms, tissues and cells. Therefore, we aimed to identify miR-125b gene targets in a high-throughput manner, in order to determine the role and mechanism of action of miR-125b in β-cells and its potential as a target for the treatment of diabetes.

I used a combined approach in which miR-125b overexpression in MIN6 cells (vs. control) was followed by both total mRNA sequencing (RNA-seq: T-RNA) and AGO2 RNA immunoprecipitation and sequencing (AGO2 RIP-seq: IP-RNA). I generated a ratio between the fold-change obtained by the differential analysis carried out on both sequencing data (IP-RNA/T-RNA) and considered all the genes with a ratio greater than 1 to be miR-125b direct targets. Importantly, the top targets identified with this approach revealed that miR-125b in β-cells might have an important role in the regulation of lysosomal and mitochondrial functions.

Indeed, both lysosomal and mitochondrial function are altered in diabetes. Pancreatic β-cells store insulin within secretory granules that undergo exocytosis upon detection of increased glucose levels in the blood. In order to control this process and maintain an adequate insulin secretion, β-cells activate several adaptive cellular mechanisms, including autophagy [657]. Through autophagy, cells get rid of all the damaged or old granules by delivering them to acidic lysosomes for intracellular degradation. Using the same mechanism, β-cells ensures that an optimal intracellular turnover of secretory granules is maintained maintained [658] to preserve cell homeostasis [659]. However, this process must be balanced as excessive degradation of insulin granules can impair β-cell function and cause diabetes [660].

Stress-induced nascent granule degradation has been associated with β-cell failure in T2D [585]. It has also been suggested that a small portion of degradation-derived products could be released from the cells, and this could partially explain why proinsulin, the major cargo in nascent granules, and other partially degraded granule proteins, including products of insulin degradation, are released in higher amounts in diabetic patients [661, 662]. For example, deletion of Atp6ap2, an essential component of the vascular ATPase required for lysosomal degradative functions, was found responsible for a dramatic accumulation of large, multigranular vacuoles in the cytoplasm, which resulted in reduced insulin content and compromised glucose homeostasis in β-cells [660]. These data suggest that by targeting M6PR, a protein involved in intracellular trafficking of lysosomal enzymes [586], miR-125b might affect lysosomal function [587] and be responsible for a similar phenotype observed upon the deletion of Atp6ap2. However, further studies are needed to confirm this effect. These should take into account the fact that not only the overall number of insulin granules could be affected by the deletion of M6PR but also the ratio between mature and immature ones. Altered ratio between these two types of insulin granules in the β-cells overexpressing miR-125b could indicate that this miRNA impacts β-cells insulin processing and/or biosynthesis.

The importance of understanding miR-125b mediated regulation of M6PR could be also extended to other cell types where dysfunctional lysosomes cause lysosomal storage disorders (LSDs). The scientific community has identified more than 40 types of LSDs, and this number keeps growing.

All LSDs are progressive and can affect different body organs or systems including central nervous system, heart, liver, and lungs [663]. Clear evidence indicating dysfunctionality of M6PR in LSDs is still lacking but, due to its role, alerting the expression of this protein could improve the enzyme replacement therapy strategies and favour the delivery of specific therapeutic compounds to lysosomes.

Growing evidence indicates that, additionally to lysosomal dysfunction, also impaired mitochondrial functions are central contributors to β-cell failure in T2D [664]. It was established several decades ago that inhibition of the respiratory chain was responsible for impaired GSIS [665], whereas it has more recently been demonstrated that diabetes and hyperglycaemia are responsible for an impaired glucose-dependent increase of mitochondrial metabolism [666]. Under those conditions, high levels of glucose are not able to induce NADH or increase the rate of ATP synthesis, confirming that dysfunctional mitochondria might be responsible for a reduced GSIS [666]. Metabolic stress has also been found to induce the production of ROS by β-cell mitochondria [667], and this in turn contributes to mitochondria damage and decreased GSIS [668].

It is still unclear if mitochondrial function targeting can be used as a therapeutic approach for the treatment of T2D, however it might be a viable treatment strategy to improve β-cell function. According to the unpublished data produced by other members of our group, overexpression of miR-125b induces elongation of mitochondria structures in MIN6 mouse insulinoma cells, whereas its inhibition in human β-cells produces the opposite effect with mitochondria appearing more rounded. These data indicate that miR-125b could be a key regulator of mitochondrial fission/fusion dynamics in β-cells.

Fusion–fission balance is crucial for cellular fitness in response to environmental stress and extracellular stimuli [669]. Therefore, it is reasonable to think that alterations of this balance could lead to oxidative stress, mitochondrial dysfunction, and metabolic alterations. Indeed, increased fragmentation is a common feature of mitochondrial morphology in T2D and it has already been linked to overproduction of ROS induced by high glucose levels [670]. Therefore, it is possible that, by targeting MTFP1, a protein involved in the control of mitochondrial fission, and other proteins

important for various mitochondrial functions, miR-125b might at least partially induce the mitochondria dysfunctions observed in T2D. Obviously, further investigations are needed to confirm this.

It is important to mention that MIN6 cells, the model I used for miR-125b direct target identification, were originally established from an insulinoma developed in an IT6 transgenic C57BL/6 mouse line that expresses the SV40 T antigen under the control of the human insulin promoter [671]. These cells, in addition to altered gene expression and GCK/HK activity, present also a left-shifted secretory response [672] indicating that the glucose sensing is altered and suggesting that they cannot be considered as true pancreatic β-cells.

For these reasons it was important for our group to also generate a mice model with β-cell specific, inducible, miR-125b overexpression (unpublished). Preliminary data indicate that these mice are characterized by severe hyperglycaemia and impaired GSIS in vivo. In vitro, the islets extracted from these mice present less insulin content and secretion and the calcium response to glucose is impaired. Indeed, data obtained by the identification of miR-125b direct targets described in this thesis helped our group to precisely identify which type of investigations were required to understand the mechanisms underlying the phenotype observed in vivo.

Interestingly, Wei et al. have very recently identified miR-125b as an important regulator of high-fat diet induced fat accumulation and insulin resistance in mice [673]. The authors generated a whole body miR-125b-2 knockout mouse using CRISPR/CAS9 technology and found that the mice presented increased liver weight and increased adipocyte volume upon high-fat diet. Moreover, in these mice miR-125b-2KO also altered HFD-induced changes in glucose tolerance and insulin resistance. Thus, the authors concluded that overexpression of miR-125b might be a novel potential target for the regulation of fat accumulation, and this strategy might be used to treat obesity and diabetes [673]. Our results, on the contrary, demonstrated that overexpression of miR-125b in MIN6 cells repress genes important for both mitochondrial and lysosomal function and led to lysosomal hypertrophy. These data, together with results observer in vivo where β-cell specific miR-125b overexpression impairs glucose tolerance and insulin secretion, strongly suggest that uncontrolled miR-125b overexpression will lead to failure of β-cell function and impairment of

insulin secretion. Therefore, we can speculate that miR-125b is targeted in a cell-specific manner and that the systemic delivery of a miR-125b mimic in order to treat diabetes could not have the results expected by Wei et al. but, instead, might be deleterious for β-cell function. This further highlights the importance of precise miRNA target identification to accurately consider all the regulatory networks involved before proceeding with the development of new therapeutic strategy.

For these reasons, my PhD project also aimed to optimize AGO2 eiCLIP, a sophisticated CLIP technique that allows the identification of miRNA direct target in an unbiased manner. More than 300 miRNAs are expressed in β-cells but very little is known about the specific role of most of them and their contribution to T2D pathogenesis. The selection of relevant miRNAs for in-depth studies is not always straightforward and can there be very time-consuming.

Under the guidance of our collaborator Dr. Sibley, we successfully optimized the AGO2 eiCLIP protocol initially in the fast-growing Hela cells and subsequently in human ENDOC-βH1 β-cells. The computational analysis performed on the sequencing results obtained in Hela cells confirmed that the technique was able to identify AGO2 crosslinking sites and precisely map them within 5-10 nucleotides from the miRNAs seed region. This information will be most useful to computationally determine the miRNAs responsible for the targeting of miRISC to the EndoC-βH1 transcriptome in our future experiments. Additionally, the AGO2 eiCLIP cDNA libraries generated in the ENDOCβH1 cells during this PhD project will be sequenced and results analysed to precisely map the entire miRNA targetome in β-cells.

Furthermore, when results obtained in wild-type β-cells will be compared to those obtained in models that mimic diabetes we will be able to provide a comprehensive understanding of the role of miRNAs and their implication in T2D. Possibly, these results will also aid in the identification of new targets for the treatment of this complex and multifactorial disease.

Remarkably, it is only in the recent years that the role of RBPs dysregulation came to be appreciated in the context of pancreatic endocrine cells function and only a small number of them have so far been studied in β-cells [674]. For example, it has only very recently been demonstrated that treatment of EndoC-βH1 with cell stressors such as high and low levels of glucose or cytokines can

induce dysregulation of many RBPs [675], thus indicating that their dysregulation can contribute to the development of diabetes. In this context, our newly optimized eiCLIP protocol could prove useful to identify the roles that different RPBs have in the development of diabetes. This rapidly emerging field will doubtless provide a new perspective on the therapeutic treatments for the extremely complex and multifaceted disease that is diabetes.

In summary, the results shown in this thesis mainly suggest that miR-125b is an important regulator of the β-cell lysosomal system and mitochondria dynamics and that it might play a critical role during hyperglycaemia. These results, together with the ones obtained *in vivo* by our group, strongly indicate that non-targeted administration of miR-125b mimics will lead to defective β-cell function and worsen diabetes. Future studies focussed on cell-specific targeting of miR-125b will be essential to confirm the potential therapeutic benefit of targeting this miRNA for the treatment of diabetes and other metabolic disease.

Bibliography

1.   Röder, P.V., et al., *Pancreatic regulation of glucose homeostasis*, in *Exp Mol Med*. 2016. p. e219-.
2.   Alsahli, M., M.Z. Shrayyef, and J.E. Gerich, *Normal Glucose Homeostasis*, in *Principles of Diabetes Mellitus*, L. Poretsky, Editor. 2017, Springer International Publishing: Cham. p. 23-42.
3.   Wahren, J., P. Felig, and L. Hagenfeldt, *Physical exercise and fuel homeostasis in diabetes mellitus.* Diabetologia, 2020. **14**(4): p. 213-222.
4.   Consoli, A., et al., *Determination of Krebs cycle metabolic carbon exchange in vivo and its use to estimate the individual contributions of gluconeogenesis and glycogenolysis to overall glucose output in man.* J Clin Invest, 1987. **80**(5): p. 1303-10.
5.   Freychet, L., et al., *Effect of intranasal glucagon on blood glucose levels in healthy subjects and hypoglycaemic patients with insulin-dependent diabetes.* Lancet, 1988. **1**(8599): p. 1364-6.
6.   McTernan, P.G., et al., *Insulin and rosiglitazone regulation of lipolysis and lipogenesis in human adipose tissue in vitro.* Diabetes, 2002. **51**(5): p. 1493-8.
7.   Rossetti, L. and A. Giaccari, *Relative contribution of glycogen synthesis and glycolysis to insulin-mediated glucose uptake. A dose-response euglycemic clamp study in normal and diabetic rats.* J Clin Invest, 1990. **85**(6): p. 1785-92.
8.   Walton, P.E. and T.D. Etherton, *Stimulation of lipogenesis by insulin in swine adipose tissue: antagonism by porcine growth hormone.* J Anim Sci, 1986. **62**(6): p. 1584-95.
9.   Pozefsky, T., et al., *Amino acid balance across tissues of the forearm in postabsorptive man. Effects of insulin at two dose levels.* J Clin Invest, 1969. **48**(12): p. 2273-82.
10.   Gittes, G.K., *Developmental biology of the pancreas: a comprehensive review.* Dev Biol, 2009. **326**(1): p. 4-35.
11.   Slack, J.M., *Developmental biology of the pancreas.* 1995.
12.   Williams, J.A., *Regulation of acinar cell function in the pancreas.* Curr Opin Gastroenterol, 2010. **26**(5): p. 478-83.
13.   Weir, G.C. and S. Bonner-Weir, *Islets of Langerhans: the puzzle of intraislet interactions and their relevance to diabetes.* J Clin Invest, 1990. **85**(4): p. 983-7.
14.   Brissova, M., et al., *Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy.* J Histochem Cytochem, 2005. **53**(9): p. 1087-97.
15.   Alanentalo, T., et al., *Tomographic molecular imaging and 3D quantification within adult mouse organs.* Nature Methods, 2006. **4**(1): p. 31-33.
16.   Pillay, K. and P. Govender, *Amylin Uncovered: A Review on the Polypeptide Responsible for Type II Diabetes.* Biomed Res Int, 2013. **2013**.
17.   Godoy-Matos, A.F., *The role of glucagon on type 2 diabetes at a glance.* Diabetology & Metabolic Syndrome, 2014. **6**(1): p. 1-5.
18.   Tricia M. Tan, S.R.B., *Chapter 176 - Pancreatic Polypeptide*. Handbook of Biologically Active Peptides (Second Edition). 2013.
19.   Sakata, N., G. Yoshimatsu, and S. Kodama, *Development and Characteristics of Pancreatic Epsilon Cells*, in *Int J Mol Sci*. 2019.
20.   Dolenšek, J., M. Rupnik, and A. Stožer, *Structural Similarities and Differences Between the Human and the Mouse Pancreas.* Islets, 2015. **7**(1).
21.   Saito, K., N. Yaginuma, and T. Takahashi, *Differential volumetry of A, B and D cells in the pancreatic islets of diabetic and nondiabetic subjects.* Tohoku J Exp Med, 1979. **129**(3): p. 273-83.
22.   Henderson, J.R. and M.C. Moss, *A morphometric study of the endocrine and exocrine capillaries of the pancreas.* Q J Exp Physiol, 1985. **70**(3): p. 347-56.

23. Hutton, J.C., *Insulin secretory granule biogenesis and the proinsulin-processing endopeptidases.* Diabetologia, 1994. **37 Suppl 2**: p. S48-56.

24. Olofsson, C.S., et al., *Fast insulin secretion reflects exocytosis of docked granules in mouse pancreatic B-cells.* Pflugers Arch, 2002. **444**(1-2): p. 43-51.

25. Daniel, S., et al., *Identification of the docked granule pool responsible for the first phase of glucose-stimulated insulin secretion.* 1999.

26. Bratanova-Tochkova, T.K., et al., *Triggering and augmentation mechanisms, granule pools, and biphasic insulin secretion.* Diabetes, 2002. **51 Suppl 1**: p. S83-90.

27. BB, B., R. CJ, and G. JS, *The dynamic plasticity of insulin production in β-cells.* Molecular metabolism, 2017. **6**(9).

28. Mueckler, M. and B. Thorens, *The SLC2 (GLUT) family of membrane transporters.* Mol Aspects Med, 2013. **34**(2-3): p. 121-38.

29. De Vos, A., et al., *Human and rat beta cells differ in glucose transporter but not in glucokinase gene expression.* J Clin Invest, 1995. **96**(5): p. 2489-95.

30. Coppieters, K.T., et al., *Persistent glucose transporter expression on pancreatic beta cells from longstanding type 1 diabetic individuals.* Diabetes Metab Res Rev, 2011. **27**(8): p. 746-54.

31. Newgard, C.B. and J.D. McGarry, *Metabolic coupling factors in pancreatic beta-cell signal transduction.* Annu Rev Biochem, 1995. **64**: p. 689-719.

32. Guillam, M.T., et al., *Early diabetes and abnormal postnatal pancreatic islet development in mice lacking Glut-2.* Nat Genet, 1997. **17**(3): p. 327-30.

33. Thorens, B., et al., *Transgenic reexpression of GLUT1 or GLUT2 in pancreatic beta cells rescues GLUT2-null mice from early death and restores normal glucose-stimulated insulin secretion.* J Biol Chem, 2000. **275**(31): p. 23751-8.

34. Thorens, B., *GLUT2, glucose sensing and glucose homeostasis.* Diabetologia, 2015. **58**(2): p. 221-32.

35. Matschinsky, F.M., *Banting Lecture 1995. A lesson in metabolic regulation inspired by the glucokinase glucose sensor paradigm.* Diabetes, 1996. **45**(2): p. 223-41.

36. Matschinsky, F.M., *Regulation of pancreatic beta-cell glucokinase: from basics to therapeutics.* Diabetes, 2002. **51 Suppl 3**: p. S394-404.

37. Iynedjian, P.B., *Mammalian glucokinase and its gene.* Biochem J, 1993. **293 (Pt 1)**: p. 1-13.

38. Vionnet, N., et al., *Nonsense mutation in the glucokinase gene causes early-onset non-insulin-dependent diabetes mellitus.* Nature, 1992. **356**(6371): p. 721-722.

39. Froguel, P., et al., *Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus.* Nature, 1992. **356**(6365): p. 162-4.

40. Terauchi, Y., et al., *Pancreatic β-Cell-specific Targeted Disruption of Glucokinase Gene.* 1995.

41. Maechler, P., S. Carobbio, and B. Rubi, *In beta-cells, mitochondria integrate and generate metabolic signals controlling insulin secretion.* The international journal of biochemistry & cell biology, 2006. **38**(5-6): p. 696-709.

42. Inagaki, N., et al., *Reconstitution of IKATP: An Inward Rectifier Subunit Plus the Sulfonylurea Receptor.* Science (New York, N.Y.), 1995. **270**(5239).

43. Markworth, E., S. C, and S. M, *ATP4- Mediates Closure of Pancreatic Beta-Cell ATP-sensitive Potassium Channels by Interaction With 1 of 4 Identical Sites.* Diabetes, 2000. **49**(9).

44. Tarasov, A., D. J, and A. F, *Metabolic Regulation of the Pancreatic Beta-Cell ATP-sensitive K+ Channel: A Pas De Deux.* Diabetes, 2004. **53 Suppl 3**.

45. Safayhi, H., et al., *L-type Calcium Channels in Insulin-Secreting Cells: Biochemical Characterization and Phosphorylation in RINm5F Cells.* Molecular endocrinology (Baltimore, Md.), 1997. **11**(5).

46.   Ohara-Imaizumi, M., et al., *Site of Docking and Fusion of Insulin Secretory Granules in Live MIN6 Beta Cells Analyzed by TAT-conjugated Anti-Syntaxin 1 Antibody and Total Internal Reflection Fluorescence Microscopy.* The Journal of biological chemistry, 2004. **279**(9).

47.   Sher, E., et al., *Voltage-operated Calcium Channel Heterogeneity in Pancreatic Beta Cells: Physiopathological Implications.* Journal of bioenergetics and biomembranes, 2003. **35**(6).

48.   Ashcroft, F., et al., *Stimulus-secretion Coupling in Pancreatic Beta Cells.* Journal of cellular biochemistry, 1994. **55 Suppl**.

49.   Steiner, D.F. and P.E. Oyer, *THE BIOSYNTHESIS OF INSULIN AND A PROBABLE PRECURSOR OF INSULIN BY A HUMAN ISLET CELL ADENOMA.* Proc Natl Acad Sci U S A, 1967. **57**(2): p. 473-80.

50.   Lomedico, P.T., et al., *The structure of rat preproinsulin genes.* Ann N Y Acad Sci, 1980. **343**: p. 425-32.

51.   Ohneda, K., H. Ee, and M. German, *Regulation of insulin gene transcription.* Semin Cell Dev Biol, 2000. **11**(4): p. 227-33.

52.   Melloul, D., S. Marshak, and E. Cerasi, *Regulation of insulin gene transcription.* Diabetologia, 2002. **45**(3): p. 309-26.

53.   German, M., et al., *The insulin gene promoter: A simplified nomenclature.* Diabetes, 1995. **44**(8): p. 1002-1004.

54.   Andrali, S.S., et al., *Glucose regulation of insulin gene expression in pancreatic beta-cells.* Biochem J, 2008. **415**(1): p. 1-10.

55.   Goodison, S., S. Kenna, and S.J. Ashcroft, *Control of insulin gene expression by glucose.* Biochem J, 1992. **285**(Pt 2): p. 563-8.

56.   Poitout, V., et al., *Regulation of the Insulin Gene by Glucose and Fatty Acids1.* J Nutr, 2006. **136**(4): p. 873-6.

57.   Sharma, A., et al., *The NeuroD1/BETA2 Sequences Essential for Insulin Gene Transcription Colocalize With Those Necessary for Neurogenesis and p300/CREB Binding Protein Binding.* Molecular and cellular biology, 1999. **19**(1).

58.   Qiu, Y., et al., *Insulin Gene Transcription Is Mediated by Interactions between the p300 Coactivator and PDX-1, BETA2, and E47*, in *Mol Cell Biol.* 2002. p. 412-20.

59.   Mosley, A. and O. S, *Glucose Regulates Insulin Gene Transcription by Hyperacetylation of Histone h4.* The Journal of biological chemistry, 2003. **278**(22).

60.   Eskridge, E. and S. D, *Cell-free processing and segregation of insulin precursors.* The Journal of biological chemistry, 1983. **258**(19).

61.   Goodge, K.A. and J.C. Hutton, *Translational regulation of proinsulin biosynthesis and proinsulin conversion in the pancreatic beta-cell.* Semin Cell Dev Biol, 2000. **11**(4): p. 235-42.

62.   Eliasson, L., et al., *Novel Aspects of the Molecular Mechanisms Controlling Insulin Secretion.* The Journal of physiology, 2008. **586**(14).

63.   Heaslip, A.T., et al., *Cytoskeletal Dependence of Insulin Granule Movement Dynamics in INS-1 Beta-Cells in Response to Glucose*, in *PLoS One.* 2014.

64.   Dawicki-McKenna, J., G. YE, and O. EM, *Sites of Glucose transporter-4 Vesicle Fusion With the Plasma Membrane Correlate Spatially With Microtubules.* PloS one, 2012. **7**(8).

65.   Fimia, G. and S.-C. P, *Cyclic AMP signalling.* Journal of cell science, 2001. **114**(Pt 11).

66.   Tengholm, A., *Cyclic AMP dynamics in the pancreatic β-cell.* Upsala journal of medical sciences, 2012. **117**(4).

67.   Seino, S. and S. T, *PKA-dependent and PKA-independent Pathways for cAMP-regulated Exocytosis.* Physiological reviews, 2005. **85**(4).

68.   Hatakeyama, H., et al., *Two cAMP-dependent pathways differentially regulate exocytosis of large dense-core and small vesicles in mouse β-cells.* J Physiol, 2007. **582**(Pt 3): p. 1087-98.

69. Landa, L., et al., *Interplay of Ca2+ and cAMP Signaling in the Insulin-Secreting MIN6 Beta-Cell Line.* The Journal of biological chemistry, 2005. **280**(35).

70. Dyachok, O., et al., *Oscillations of Cyclic AMP in Hormone-Stimulated Insulin-Secreting Beta-Cells.* Nature, 2006. **439**(7074).

71. Shibasaki, T., et al., *Essential Role of Epac2/Rap1 Signaling in Regulation of Insulin Granule Dynamics by cAMP.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(49).

72. Xiong, Q., et al., *Key proteins involved in insulin vesicle exocytosis and secretion.* Biomedical reports, 2017. **6**(2).

73. Bruns, D. and J. R, *Molecular Determinants of Exocytosis.* Pflugers Archiv : European journal of physiology, 2002. **443**(3).

74. Thurmond, D.C., *Regulation of Insulin Action and Insulin Secretion by SNARE-Mediated Vesicle Exocytosis.* 2013.

75. Barg, S., et al., *A Subset of 50 Secretory Granules in Close Contact With l-Type Ca2+ Channels Accounts for First-Phase Insulin Secretion in Mouse β-Cells.* 2002.

76. Ashcroft, F., H. DE, and A. SJ, *Glucose induces closure of single potassium channels in isolated rat pancreatic beta-cells.* Nature, 1984. **312**(5993).

77. Miki, T., et al., *Defective insulin secretion and enhanced insulin action in KATP channel-deficient mice.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(18).

78. Dunne, M., et al., *Hyperinsulinism in infancy: from basic science to clinical disease.* Physiological reviews, 2004. **84**(1).

79. Sato, Y., et al., *Glucose Regulation of Insulin Secretion Independent of the Opening or Closure of Adenosine Triphosphate-Sensitive K+ Channels in β Cells.* Endocrinology, 2020. **140**(5): p. 2252-2257.

80. Straub, S., et al., *Glucose Activates Both K(ATP) Channel-Dependent and K(ATP) Channel-Independent Signaling Pathways in Human Islets.* Diabetes, 1998. **47**(5).

81. Gembal, M., G. P, and H. JC, *Evidence That Glucose Can Control Insulin Release Independently From Its Action on ATP-sensitive K+ Channels in Mouse B Cells.* The Journal of clinical investigation, 1992. **89**(4).

82. Prentki, M., C. BE, and M. SRM, *Lipid-associated metabolic signalling networks in pancreatic beta cell function.* Diabetologia, 2020. **63**(1).

83. Nolan, C., et al., *Fatty acid signaling in the beta-cell and insulin secretion.* Diabetes, 2006. **55 Suppl 2**.

84. Ferdaoussi, M., et al., *G protein-coupled receptor (GPR)40-dependent potentiation of insulin secretion in mouse islets is mediated by protein kinase D1.* Diabetologia, 2012. **55**(10).

85. Yokoi, N., et al., *β-Cell glutamate signaling: Its role in incretin-induced insulin secretion.* Journal of diabetes investigation, 2016. **7 Suppl 1**(Suppl 1).

86. Kim, W. and E. JM, *The role of incretins in glucose homeostasis and diabetes treatment.* Pharmacological reviews, 2008. **60**(4).

87. Jones, B., et al., *Control of insulin secretion by GLP-1.* Peptides, 2018. **100**.

88. Drucker, D., *Incretin action in the pancreas: potential promise, possible perils, and pathological pitfalls.* Diabetes, 2013. **62**(10).

89. Association, A.D., *2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2019.* Diabetes care, 2019. **42**(Suppl 1).

90. Organization, W.H., *Global report on diabete.* WHO, 2017.

91. Seuring, T., O. Archangelidi, and M. Suhrcke, *The Economic Costs of Type 2 Diabetes: A Global Systematic Review.* Pharmacoeconomics, 2015. **33**(8): p. 811-31.

92. Urakami, T., *Maturity-onset diabetes of the young (MODY): current perspectives on diagnosis and treatment.* Diabetes, metabolic syndrome and obesity : targets and therapy, 2019. **12**.

93. Achenbach, P., et al., *Natural history of type 1 diabetes.* Diabetes, 2005. **54 Suppl 2**: p. S25-31.

94. Federation, I.D. *International Diabetes Federation*. 2020; Available from: https://www.idf.org/.

95. Rewers, M. and L. J, *Environmental risk factors for type 1 diabetes.* Lancet (London, England), 2016. **387**(10035).

96. Rui, J., et al., *beta Cells that Resist Immunological Attack Develop during Progression of Autoimmune Diabetes in NOD Mice.* Cell Metab, 2017. **25**(3): p. 727-738.

97. Pulgaron, E. and D. AM, *Obesity and type 2 diabetes in children: epidemiology and treatment.* Current diabetes reports, 2014. **14**(8).

98. Roumen, C., B. EE, and C. E, *Lifestyle Intervention for Prevention of Diabetes: Determinants of Success for Future Implementation.* Nutrition reviews, 2009. **67**(3).

99. Balk, E., et al., *Combined Diet and Physical Activity Promotion Programs to Prevent Type 2 Diabetes Among Persons at Increased Risk: A Systematic Review for the Community Preventive Services Task Force.* Annals of internal medicine, 2015. **163**(6).

100. Hu, F.B., *Globalization of diabetes: the role of diet, lifestyle, and genes*, in *Diabetes Care*. 2011: United States. p. 1249-57.

101. Barnett, A.H., et al., *Diabetes in identical twins. A study of 200 pairs.* Diabetologia, 1981. **20**(2): p. 87-93.

102. Sladek, R., et al., *A genome-wide association study identifies novel risk loci for type 2 diabetes.* Nature, 2007. **445**(7130).

103. Mahajan, A., et al., *Fine-mapping of an expanded set of type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps.* Nat Genet, 2018. **50**(11): p. 1505-13.

104. DeFronzo, R.A., et al., *Type 2 diabetes mellitus.* Nature Reviews Disease Primers, 2015. **1**(1): p. 1-22.

105. Peifer, M. and P. P, *Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus.* Science (New York, N.Y.), 2000. **287**(5458).

106. Voight, B., et al., *Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis.* Nature genetics, 2010. **42**(7).

107. Le Bacquer, O., et al., *TCF7L2 rs7903146 impairs islet function and morphology in non-diabetic individuals.* Diabetologia, 2012. **55**(10).

108. GA, R., *Dorothy Hodgkin Lecture 2014. Understanding genes identified by genome-wide association studies for type 2 diabetes.* Diabetic medicine : a journal of the British Diabetic Association, 2014. **31**(12).

109. Rutter, G., *Dorothy Hodgkin Lecture 2014. Understanding genes identified by genome-wide association studies for type 2 diabetes.* Diabetic medicine : a journal of the British Diabetic Association, 2014. **31**(12).

110. DeFronzo, R.A., *From the Triumvirate to the Ominous Octet: A New Paradigm for the Treatment of Type 2 Diabetes Mellitus*, in *Diabetes*. 2009. p. 773-95.

111. Abdul-Ghani, M.A., D. Tripathy, and R.A. DeFronzo, *Contributions of beta-cell dysfunction and insulin resistance to the pathogenesis of impaired glucose tolerance and impaired fasting glucose.* Diabetes Care, 2006. **29**(5): p. 1130-9.

112. Zheng, Y., L. SH, and H. FB, *Global aetiology and epidemiology of type 2 diabetes mellitus and its complications.* Nature reviews. Endocrinology, 2018. **14**(2).

113. Knowler, W., et al., *Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin.* The New England journal of medicine, 2002. **346**(6).

114. Pan, X., et al., *Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study.* Diabetes care, 1997. **20**(4).

115. Di Cianni, G., et al., *Intermediate metabolism in normal pregnancy and in gestational diabetes.* Diabetes Metab Res Rev, 2003. **19**(4): p. 259-70.

116. Catalano, P.M., et al., *Longitudinal changes in insulin release and insulin resistance in nonobese pregnant women.* Am J Obstet Gynecol, 1991. **165**(6 Pt 1): p. 1667-72.

117. Parsons, J.A., T.C. Brelje, and R.L. Sorenson, *Adaptation of islets of Langerhans to pregnancy: increased islet cell proliferation and insulin secretion correlates with the onset of placental lactogen secretion.* Endocrinology, 1992. **130**(3): p. 1459-66.

118. Barbour, L.A., et al., *Cellular mechanisms for insulin resistance in normal pregnancy and gestational diabetes.* Diabetes Care, 2007. **30 Suppl 2**: p. S112-9.

119. Kaaja, R. and T. Ronnemaa, *Gestational diabetes: pathogenesis and consequences to mother and offspring.* Rev Diabet Stud, 2008. **5**(4): p. 194-202.

120. Plows, J.F., et al., *The Pathophysiology of Gestational Diabetes Mellitus.* Int J Mol Sci, 2018. **19**(11).

121. Peters, R.K., et al., *Long-term diabetogenic effect of single pregnancy in women with previous gestational diabetes mellitus.* Lancet, 1996. **347**(8996): p. 227-30.

122. Shostrom, D.C.V., et al., *History of Gestational Diabetes Mellitus in Relation to Cardiovascular Disease and Cardiovascular Risk Factors in US Women.* Front Endocrinol (Lausanne), 2017. **8**: p. 144.

123. Vohr, B.R. and C.M. Boney, *Gestational diabetes: the forerunner for the development of maternal and childhood obesity and metabolic syndrome?* J Matern Fetal Neonatal Med, 2008. **21**(3): p. 149-57.

124. Rahier, J., et al., *Pancreatic Beta-Cell Mass in European Subjects With Type 2 Diabetes.* Diabetes, obesity & metabolism, 2008. **10 Suppl 4**.

125. Butler, A., et al., *Adaptive Changes in Pancreatic Beta Cell Fractional Area and Beta Cell Turnover in Human Pregnancy.* Diabetologia, 2010. **53**(10).

126. Van Assche, F., A. L, and D.P. F, *A Morphological Study of the Endocrine Pancreas in Human Pregnancy.* British journal of obstetrics and gynaecology, 1978. **85**(11).

127. Maclean, N. and O. RF, *Quantitative Estimation of the Pancreatic Islet Tissue in Diabetic Subjects.* Diabetes, 1955. **4**(5).

128. Butler, A., et al., *Beta-cell Deficit and Increased Beta-Cell Apoptosis in Humans With Type 2 Diabetes.* Diabetes, 2003. **52**(1).

129. Klöppel, G., et al., *Islet Pathology and the Pathogenesis of Type 1 and Type 2 Diabetes Mellitus Revisited.* Survey and synthesis of pathology research, 1985. **4**(2).

130. Bensellam, M., L. DR, and J. JC, *The molecular mechanisms of pancreatic β-cell glucotoxicity: recent findings and future research directions.* Molecular and cellular endocrinology, 2012. **364**(1-2).

131. Oh, Y.S., et al., *Fatty Acid-Induced Lipotoxicity in Pancreatic Beta-Cells During Development of Type 2 Diabetes.* Front Endocrinol (Lausanne), 2018. **9**.

132. Fonseca, S., L. KL, and U. F, *Endoplasmic Reticulum Stress Signaling in Pancreatic Beta-Cells.* Antioxidants & redox signaling, 2007. **9**(12).

133. Laybutt, D., et al., *Endoplasmic Reticulum Stress Contributes to Beta Cell Apoptosis in Type 2 Diabetes.* Diabetologia, 2007. **50**(4).

134. Abedini, A. and A.M. Schmidt, *Mechanisms of Islet Amyloidosis Toxicity in Type 2 Diabetes.* FEBS Lett, 2013. **587**(8): p. 1119-27.

135. Huang, C., et al., *High Expression Rates of Human Islet Amyloid Polypeptide Induce Endoplasmic Reticulum Stress Mediated Beta-Cell Apoptosis, a Characteristic of Humans With Type 2 but Not Type 1 Diabetes.* Diabetes, 2007. **56**(8).

136. Davidson, M., *Pathogenesis of Impaired Glucose Tolerance and Type II Diabetes Mellitus--Current Status.* The Western journal of medicine, 1985. **142**(2).

137. Rosengren, A., et al., *Reduced Insulin Exocytosis in Human Pancreatic β-Cells With Gene Variants Linked to Type 2 Diabetes.* Diabetes, 2012. **61**(7).

138. Deng, S., et al., *Structural and Functional Abnormalities in the Islets Isolated From Type 2 Diabetic Subjects.* Diabetes, 2004. **53**(3).

139. Del Guerra, S., et al., *Functional and Molecular Defects of Pancreatic Islets in Human Type 2 Diabetes.* Diabetes, 2005. **54**(3).

140. Doliba, N., et al., *Glucokinase Activation Repairs Defective Bioenergetics of Islets of Langerhans Isolated From Type 2 Diabetics.* American journal of physiology. Endocrinology and metabolism, 2012. **302**(1).

141. Ostenson, C., et al., *Impaired Gene and Protein Expression of Exocytotic Soluble N-ethylmaleimide Attachment Protein Receptor Complex Proteins in Pancreatic Islets of Type 2 Diabetic Patients.* Diabetes, 2006. **55**(2).

142. Tsunoda, K., et al., *Single Nucleotide Polymorphism (D68D, T to C) in the Syntaxin 1A Gene Correlates to Age at Onset and Insulin Requirement in Type II Diabetic Patients.* Diabetologia, 2001. **44**(11).

143. Ostenson, C. and E. S, *Islet Gene Expression and Function in Type 2 Diabetes; Studies in the Goto-Kakizaki Rat and Humans.* Diabetes, obesity & metabolism, 2007. **9 Suppl 2**.

144. Escribano, O., et al., *β-Cell Hyperplasia Induced by Hepatic Insulin Resistance.* 2009.

145. Cerf, M.E., *Beta Cell Dysfunction and Insulin Resistance.* Front Endocrinol (Lausanne), 2013. **4**.

146. Campos, C., *Chronic hyperglycemia and glucose toxicity: pathology and clinical sequelae.* Postgrad Med, 2012. **124**(6): p. 90-7.

147. Brownlee, M., *The pathobiology of diabetic complications: a unifying mechanism*, in *Diabetes*. 2005: United States. p. 1615-25.

148. Verkhratsky, A. and F. P, *Mitochondrial malfunction and Ca2+ dyshomeostasis drive neuronal pathology in diabetes.* Cell Calcium, 2020. **44**(1): p. 112-122.

149. Brownlee, M., *A radical explanation for glucose-induced beta cell dysfunction.* J Clin Invest, 2003. **112**(12): p. 1788-90.

150. Brereton, M.F., et al., *Hyperglycaemia induces metabolic dysfunction and glycogen accumulation in pancreatic β-cells.* Nature Communications, 2016. **7**(1): p. 1-15.

151. C, T., et al., *Pancreatic β cell dedifferentiation as a mechanism of diabetic β cell failure.* Cell, 2012. **150**(6).

152. Gao, T., et al., *Pdx1 maintains β-cell identity and function by repressing an α-cell program.* Cell Metab, 2014. **19**(2): p. 259-71.

153. Taylor, B.L., F.F. Liu, and M. Sander, *Nkx6.1 is essential for maintaining the functional state of pancreatic beta cells.* Cell Rep, 2013. **4**(6): p. 1262-75.

154. Swisa, A., et al., *PAX6 maintains β cell identity by repressing genes of alternative islet cell types*, in *J Clin Invest*. p. 230-43.

155. Yin, Q., et al., *Raptor determines β-cell identity and plasticity independent of hyperglycemia in mice.* Nature communications, 2020. **11**(1).

156. Pullen, T., et al., *Identification of genes selectively disallowed in the pancreatic islet.* Islets, 2010. **2**(2).

157. Pullen, T. and R. GA, *When less is more: the forbidden fruits of gene repression in the adult β-cell.* Diabetes, obesity & metabolism, 2013. **15**(6).

158. Sekine, N., et al., *Low lactate dehydrogenase and high mitochondrial glycerol phosphate dehydrogenase in pancreatic beta-cells. Potential role in nutrient sensing.* The Journal of biological chemistry, 1994. **269**(7).

159. Talchai, C., et al., *Pancreatic β cell dedifferentiation as a mechanism of diabetic β cell failure.* Cell, 2012. **150**(6).

160. Wang, C., et al., *5-amino-imidazole carboxamide riboside acutely potentiates glucose-stimulated insulin secretion from mouse pancreatic islets by KATP channel-dependent and -independent pathways.* Biochemical and biophysical research communications, 2005. **330**(4).

161. Li, X., et al., *Structural basis of AMPK regulation by adenine nucleotides and glycogen.* Cell research, 2015. **25**(1).

162. Cinti, F., et al., *Evidence of β-Cell Dedifferentiation in Human Type 2 Diabetes.* The Journal of clinical endocrinology and metabolism, 2016. **101**(3).

163. Spijker, H., et al., *Loss of β-Cell Identity Occurs in Type 2 Diabetes and Is Associated With Islet Amyloid Deposits.* Diabetes, 2015. **64**(8).

164. Brereton, M., et al., *Reversible changes in pancreatic islet structure and function produced by elevated blood glucose.* Nature communications, 2014. **5**.

165. Jonas, J., et al., *Chronic Hyperglycemia Triggers Loss of Pancreatic Beta Cell Differentiation in an Animal Model of Diabetes.* The Journal of biological chemistry, 1999. **274**(20).

166. Butler, A., et al., *β-Cell Deficit in Obese Type 2 Diabetes, a Minor Role of β-Cell Dedifferentiation and Degranulation.* The Journal of clinical endocrinology and metabolism, 2016. **101**(2).

167. Abdul-Ghani, M., et al., *Initial Combination Therapy With Metformin, Pioglitazone and Exenatide Is More Effective Than Sequential Add-On Therapy in Subjects With New-Onset Diabetes. Results From the Efficacy and Durability of Initial Combination Therapy for Type 2 Diabetes (EDICT): A Randomized Trial.* Diabetes, obesity & metabolism, 2015. **17**(3).

168. Harrison, L., et al., *β-Cell Function Preservation After 3.5 Years of Intensive Diabetes Therapy.* Diabetes care, 2012. **35**(7).

169. DeFronzo, R., et al., *Combination of Empagliflozin and Linagliptin as Second-Line Therapy in Subjects With Type 2 Diabetes Inadequately Controlled on Metformin.* Diabetes care, 2015. **38**(3).

170. Agarwal, A.A., P.R. Jadhav, and Y.A. Deshmukh, *Prescribing pattern and efficacy of anti-diabetic drugs in maintaining optimal glycemic levels in diabetic patients*, in *J Basic Clin Pharm*. 2014. p. 79-83.

171. Davies, M., et al., *Management of Hyperglycemia in Type 2 Diabetes, 2018. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD).* Diabetes care, 2018. **41**(12).

172. Hundal, R., et al., *Mechanism by Which Metformin Reduces Glucose Production in Type 2 Diabetes.* Diabetes, 2000. **49**(12).

173. Hawley, S., et al., *The Antidiabetic Drug Metformin Activates the AMP-activated Protein Kinase Cascade via an Adenine Nucleotide-Independent Mechanism.* Diabetes, 2002. **51**(8).

174. Zhou, G., et al., *Role of AMP-activated Protein Kinase in Mechanism of Metformin Action.* The Journal of clinical investigation, 2001. **108**(8).

175. Madiraju, A., et al., *Metformin Suppresses Gluconeogenesis by Inhibiting Mitochondrial Glycerophosphate Dehydrogenase.* Nature, 2014. **510**(7506).

176. Hardie, D.G., *AMPK − sensing energy while talking to other signaling pathways.* Cell Metab, 2014. **20**(6): p. 939-52.

177. Stephenne, X., et al., *Metformin Activates AMP-activated Protein Kinase in Primary Human Hepatocytes by Decreasing Cellular Energy Status.* Diabetologia, 2011. **54**(12).

178. Foretz, M., et al., *Metformin Inhibits Hepatic Gluconeogenesis in Mice Independently of the LKB1/AMPK Pathway via a Decrease in Hepatic Energy State.* The Journal of clinical investigation, 2010. **120**(7).

179. Proks, P., et al., *Sulfonylurea Stimulation of Insulin Secretion.* 2002.

180. Roumie, C., et al., *Comparative Effectiveness of Sulfonylurea and Metformin Monotherapy on Cardiovascular Events in Type 2 Diabetes Mellitus: A Cohort Study.* Annals of internal medicine, 2012. **157**(9).

181. Cervera, A., et al., *Mechanism of Action of Exenatide to Reduce Postprandial Hyperglycemia in Type 2 Diabetes.* American journal of physiology. Endocrinology and metabolism, 2008. **294**(5).

182. Singh, A., *Dipeptidyl peptidase-4 Inhibitors: Novel Mechanism of Actions.* Indian journal of endocrinology and metabolism, 2014. **18**(6).

183. Cox, M.E., et al., *Dipeptidyl peptidase-4 inhibitors in the management of type 2 diabetes: safety, tolerability, and efficacy*, in *Drug Healthc Patient Saf*. 2010. p. 7-19.

184. Garcia, D. and R.J. Shaw, *AMPK: mechanisms of cellular energy sensing and restoration of metabolic balance.* Mol Cell, 2017. **66**(6): p. 789-800.

185. Zhang, B., Z. G, and L. C, *AMPK: an emerging drug target for diabetes and the metabolic syndrome.* Cell metabolism, 2009. **9**(5).

186. Salt, I., et al., *AMP-activated protein kinase: greater AMP dependence, and preferential nuclear localization, of complexes containing the alpha2 isoform.* Biochem J, 1998. **334**(Pt 1): p. 177-87.

187. Da Silva Xavier, G., et al., *Role for AMP-activated protein kinase in glucose-stimulated insulin secretion and preproinsulin gene expression.* The Biochemical journal, 2003. **371**(Pt 3).

188. Hardie, D., *AMPK: a key regulator of energy balance in the single cell and the whole organism.* International journal of obesity (2005), 2008. **32 Suppl 4**.

189. D, G. and S. RJ, *AMPK: Mechanisms of Cellular Energy Sensing and Restoration of Metabolic Balance.* Molecular cell, 2017. **66**(6).

190. Oakhill, J., et al., *β-Subunit myristoylation is the gatekeeper for initiating metabolic stress sensing by AMP-activated protein kinase (AMPK).* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(45).

191. Warden, S., et al., *Post-translational modifications of the beta-1 subunit of AMP-activated protein kinase affect enzyme activity and cellular localization.* The Biochemical journal, 2001. **354**(Pt 2).

192. Sun, G., et al., *Ablation of AMP-activated protein kinase alpha1 and alpha2 from mouse pancreatic beta cells and RIP2.Cre neurons suppresses insulin release in vivo.* Diabetologia, 2010. **53**(5).

193. Fu, A., E. CE, and S. RA, *Role of AMPK in Pancreatic Beta Cell Function.* Molecular and cellular endocrinology, 2013. **366**(2).

194. Leclerc, I. and R. GA, *AMP-activated protein kinase: a new beta-cell glucose sensor?: Regulation by amino acids and calcium ions.* Diabetes, 2004. **53 Suppl 3**.

195. Xue, B. and B.B. Kahn, *AMPK integrates nutrient and hormonal signals to regulate food intake and energy balance through effects in the hypothalamus and peripheral tissues.* J Physiol, 2006. **574**(Pt 1): p. 73-83.

196. Rourke, J., H. Q, and S. RA, *AMPK and Friends: Central Regulators of β Cell Biology.* Trends in endocrinology and metabolism: TEM, 2018. **29**(2).

197. Gowans, G., et al., *AMP is a true physiological regulator of AMP-activated protein kinase by both allosteric activation and enhancing net phosphorylation.* Cell metabolism, 2013. **18**(4).

198. Oakhill, J., et al., *AMPK is a direct adenylate charge-regulated protein kinase.* Science (New York, N.Y.), 2011. **332**(6036).

199. Woods, A., et al., *Ca2+/calmodulin-dependent protein kinase kinase-beta acts upstream of AMP-activated protein kinase in mammalian cells.* Cell metabolism, 2005. **2**(1).

200. Momcilovic, M., H. SP, and C. M, *Mammalian TAK1 activates Snf1 protein kinase in yeast and phosphorylates AMP-activated protein kinase in vitro.* The Journal of biological chemistry, 2006. **281**(35).

201. Shackelford, D. and S. RJ, *The LKB1-AMPK pathway: metabolism and growth control in tumour suppression.* Nature reviews. Cancer, 2009. **9**(8).

202. Hardie, D., *Keeping the home fires burning: AMP-activated protein kinase.* Journal of the Royal Society, Interface, 2018. **15**(138).

203. Stahmann, N., et al., *Activation of AMP-activated Protein Kinase by Vascular Endothelial Growth Factor Mediates Endothelial Angiogenesis Independently of Nitric-oxide Synthase*,* in *J Biol Chem.* 2010. p. 10638-52.

204. Yang, Y., et al., *Hunger states switch a flip-flop memory circuit via a synaptic AMPK-dependent positive feedback loop.* Cell, 2011. **146**(6).

205. Andersson, U., et al., *AMP-activated protein kinase plays a role in the control of food intake.* The Journal of biological chemistry, 2004. **279**(13).

206. Hawley, S., et al., *Calmodulin-dependent protein kinase kinase-beta is an alternative upstream kinase for AMP-activated protein kinase.* Cell metabolism, 2005. **2**(1).

207. Marcelo, K.L., A.R. Means, and B. York, *The Ca2+/Calmodulin/CaMKK2 Axis: Nature's Metabolic CaMshaft.* Trends Endocrinol Metab, 2016. **27**(10): p. 706-18.

208. Chen, S., et al., *Complementary regulation of TBC1D1 and AS160 by growth factors, insulin and AMPK activators.* The Biochemical journal, 2008. **409**(2).

209. Pehmøller, C., et al., *Genetic disruption of AMPK signaling abolishes both contraction- and insulin-stimulated TBC1D1 phosphorylation and 14-3-3 binding in mouse skeletal muscle.* American journal of physiology. Endocrinology and metabolism, 2009. **297**(3).

210. Barnes, K., et al., *Activation of GLUT1 by metabolic and osmotic stress: potential involvement of AMP-activated protein kinase (AMPK).* Journal of cell science, 2002. **115**(Pt 11).

211. Marsin, A., et al., *The stimulation of glycolysis by hypoxia in activated monocytes is mediated by AMP-activated protein kinase and inducible 6-phosphofructo-2-kinase.* The Journal of biological chemistry, 2002. **277**(34).

212. Saltiel, A. and K. CR, *Insulin signalling and the regulation of glucose and lipid metabolism.* Nature, 2001. **414**(6865).

213. Andreelli, F., et al., *Liver adenosine monophosphate-activated kinase-alpha2 catalytic subunit is a key target for the control of hepatic glucose production by adiponectin and leptin but not insulin.* Endocrinology, 2006. **147**(5).

214. Habets, D., et al., *Crucial role for LKB1 to AMPKalpha2 axis in the regulation of CD36-mediated long-chain fatty acid uptake into cardiomyocytes.* Biochimica et biophysica acta, 2009. **1791**(3).

215. Winder, W., et al., *Activation of AMP-activated protein kinase increases mitochondrial enzymes in skeletal muscle.* Journal of applied physiology (Bethesda, Md. : 1985), 2000. **88**(6).

216. Merrill, G., et al., *AICA riboside increases AMP-activated protein kinase, fatty acid oxidation, and glucose uptake in rat muscle.* The American journal of physiology, 1997. **273**(6).

217. Kroemer, G., M. G, and L. B, *Autophagy and the integrated stress response.* Molecular cell, 2010. **40**(2).

218. Kim, J., et al., *AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1.* Nature cell biology, 2011. **13**(2).

219. Jung, C., et al., *mTOR regulation of autophagy.* FEBS letters, 2010. **584**(7).

220. Sudarsanam, S. and J. DE, *Functional consequences of mTOR inhibition.* Current opinion in drug discovery & development, 2010. **13**(1).

221. Minokoshi, Y., et al., *AMP-kinase regulates food intake by responding to hormonal and nutrient signals in the hypothalamus.* Nature, 2004. **428**(6982).

222. Li, Y., et al., *AMPK phosphorylates and inhibits SREBP activity to attenuate hepatic steatosis and atherosclerosis in diet-induced insulin-resistant mice.* Cell metabolism, 2011. **13**(4).

223. Kahn, B., et al., *AMP-activated protein kinase: ancient energy gauge provides clues to modern understanding of metabolism.* Cell metabolism, 2005. **1**(1).

224. Wafa, A., et al., *Stimulation of AMP-activated protein kinase (AMPK) is associated with enhancement of Glut1-mediated glucose transport.* Archives of biochemistry and biophysics, 2000. **380**(2).

225. Kurth-Kraczek, E., et al., *5' AMP-activated protein kinase activation causes GLUT4 translocation in skeletal muscle.* Diabetes, 1999. **48**(8).

226. Inoki, K., et al., *TSC2 integrates Wnt and energy signals via a coordinated phosphorylation by AMPK and GSK3 to regulate cell growth.* Cell, 2006. **126**(5).

227. Zhang, S. and K. KH, *Glucose activation of acetyl-CoA carboxylase in association with insulin secretion in a pancreatic beta-cell line.* The Journal of endocrinology, 1995. **147**(1).

228. Ronnebaum, S., et al., *Chronic suppression of acetyl-CoA carboxylase 1 in beta-cells impairs insulin secretion via inhibition of glucose rather than lipid metabolism.* The Journal of biological chemistry, 2008. **283**(21).

229. Diraison, F., et al., *Over-expression of sterol-regulatory-element-binding protein-1c (SREBP1c) in rat pancreatic islets induces lipogenesis and decreases glucose-stimulated insulin release: modulation by 5-aminoimidazole-4-carboxamide ribonucleoside (AICAR).* The Biochemical journal, 2004. **378**(Pt 3).

230. Pende, M., et al., *Hypoinsulinaemia, glucose intolerance and diminished beta-cell size in S6K1-deficient mice.* Nature, 2000. **408**(6815).

231. Eto, K., et al., *Genetic manipulations of fatty acid metabolism in beta-cells are associated with dysregulated insulin secretion.* Diabetes, 2002. **51 Suppl 3**.

232. Shimano, H., *Sterol regulatory element-binding proteins (SREBPs): transcriptional regulators of lipid synthetic genes.* Progress in lipid research, 2001. **40**(6).

233. Horton, J., G. JL, and B. MS, *SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver.* The Journal of clinical investigation, 2002. **109**(9).

234. Rutter, G., D.S.X. G, and L. I, *Roles of 5'-AMP-activated protein kinase (AMPK) in mammalian glucose homoeostasis.* The Biochemical journal, 2003. **375**(Pt 1).

235. Emens, L., L. DW, and M. LG, *Hepatocyte nuclear factor 1 alpha is expressed in a hamster insulinoma line and transactivates the rat insulin I gene.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(16).

236. Bartoov-Shifman, R., et al., *Activation of the insulin gene promoter through a direct effect of hepatocyte nuclear factor 4 alpha.* The Journal of biological chemistry, 2002. **277**(29).

237. Düfer, M., et al., *Activation of the AMP-activated protein kinase enhances glucose-stimulated insulin secretion in mouse β-cells.* Islets, 2010. **2**(3).

238. Akkan, A. and M. WJ, *Insulinotropic action of AICA riboside. I. Insulin release by isolated islets and the perfused pancreas.* Diabetes research (Edinburgh, Scotland), 1994. **25**(1).

239. Wang, X., et al., *Troglitazone acutely activates AMP-activated protein kinase and inhibits insulin secretion from beta cells.* Life sciences, 2007. **81**(2).

240. Lamontagne, J., et al., *Pioglitazone Acutely Reduces Insulin Secretion and Causes Metabolic Deceleration of the Pancreatic β-Cell at Submaximal Glucose Concentrations*, in *Endocrinology*. 2009. p. 3465-74.

241. Zhou, L., et al., *Berberine acutely inhibits insulin secretion from beta-cells through 3',5'-cyclic adenosine 5'-monophosphate signaling pathway.* Endocrinology, 2008. **149**(9).

242. Targonsky, E., et al., *alpha-lipoic acid regulates AMP-activated protein kinase and inhibits insulin secretion from beta cells.* Diabetologia, 2006. **49**(7).

243. Cai, Y., et al., *Increased oxygen radical formation and mitochondrial dysfunction mediate beta cell apoptosis under conditions of AMP-activated protein kinase stimulation.* Free radical biology & medicine, 2007. **42**(1).

244. Leclerc, I., et al., *Metformin, but not leptin, regulates AMP-activated protein kinase in pancreatic islets: impact on glucose-stimulated insulin secretion.* American journal of physiology. Endocrinology and metabolism, 2004. **286**(6).

245. Beall, C., et al., *Loss of AMP-activated protein kinase α2 subunit in mouse β-cells impairs glucose-stimulated insulin secretion and inhibits their sensitivity to hypoglycaemia*, in *Biochem J*. 2010. p. 323-33.

246. Wicksteed, B., et al., *Conditional gene targeting in mouse pancreatic ß-Cells: analysis of ectopic Cre transgene expression in the brain.* Diabetes, 2010. **59**(12).

247. Kone, M., et al., *LKB1 and AMPK Differentially Regulate Pancreatic β-Cell Identity.* FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 2014. **28**(11).

248. Fu, A., et al., *Loss of Lkb1 in adult beta cells increases beta cell mass and enhances glucose tolerance in mice.* Cell metabolism, 2009. **10**(4).

249. Granot, Z., et al., *LKB1 Regulates Pancreatic β Cell Size, Polarity, and Function.* Cell Metab, 2009. **10**(4): p. 296-308.

250. Salt, I.P., et al., *AMP-activated protein kinase is activated by low glucose in cell lines derived from pancreatic beta cells, and may regulate insulin release.* Biochem J, 1998. **335**(Pt 3): p. 533-9.

251. da Silva Xavier, G., et al., *Role for AMP-activated protein kinase in glucose-stimulated insulin secretion and preproinsulin gene expression*, in *Biochem J*. 2003. p. 761-74.

252. Viollet, B., et al., *The AMP-activated protein kinase alpha2 catalytic subunit controls whole-body insulin sensitivity.* The Journal of clinical investigation, 2003. **111**(1).

253. Jorgensen et al, S., et al., *Knockout of the alpha2 but not alpha1 5'-AMP-activated protein kinase isoform abolishes 5-aminoimidazole-4-carboxamide-1-beta-4-ribofuranosidebut not contraction-induced glucose uptake in skeletal muscle.* The Journal of biological chemistry, 2004. **279**(2).

254. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. **75**(5): p. 843-54.

255. Wightman, B., I. Ha, and G. Ruvkun, *Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans.* Cell, 1993. **75**(5): p. 855-62.

256. Reinhart, B.J., et al., *The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans.* Nature, 2000. **403**(6772): p. 901-6.

257. Calin, G.A., et al., *Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.* 2002.

258. Pasquinelli, A.E., et al., *Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA.* Nature, 2000. **408**(6808): p. 86-9.

259. Griffiths-Jones, S., et al., *miRBase: tools for microRNA genomics.* Nucleic Acids Research, 2020. **36**(suppl_1).

260. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs.* Genome Res, 2009. **19**(1): p. 92-105.

261. de Rie, D., et al., *An integrated expression atlas of miRNAs and their promoters in human and mouse.* Nat Biotechnol, 2017. **35**(9): p. 872-878.

262. Kim, Y.K. and V.N. Kim, *Processing of intronic microRNAs.* Embo j, 2007. **26**(3): p. 775-83.

263. Borchert, G.M., W. Lanier, and B.L. Davidson, *RNA polymerase III transcribes human microRNAs.* Nat Struct Mol Biol, 2006. **13**(12): p. 1097-101.

264. Lau, N.C., et al., *An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans.* Science, 2001. **294**(5543): p. 858-62.

265. Ha, M. and V.N. Kim, *Regulation of microRNA biogenesis.* Nat Rev Mol Cell Biol, 2014. **15**(8): p. 509-24.

266. Servín-González, L.S., A.J. Granados-López, and J.A. López, *Families of microRNAs Expressed in Clusters Regulate Cell Signaling in Cervical Cancer.* International Journal of Molecular Sciences, 2015. **16**(6): p. 12773-12790.

267. Catalanotto, C., C. Cogoni, and G. Zardo, *MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions.* Int J Mol Sci, 2016. **17**(10).

268. Han, J., et al., *The Drosha-DGCR8 complex in primary microRNA processing.* Genes Dev, 2004. **18**(24): p. 3016-27.

269. Ameres, S.L. and P.D. Zamore, *Diversifying microRNA sequence and function.* Nature Reviews Molecular Cell Biology, 2013. **14**(8): p. 475-488.

270. Lee, Y., et al., *The nuclear RNase III Drosha initiates microRNA processing.* Nature, 2003. **425**(6956).

271. Gregory, R., et al., *The Microprocessor complex mediates the genesis of microRNAs.* Nature, 2004. **432**(7014).

272. Landthaler, M., Y. A, and T. T, *The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis.* Current biology : CB, 2004. **14**(23).

273. Denli, A.M., et al., *Processing of primary microRNAs by the Microprocessor complex.* Nature, 2004. **432**(7014): p. 231-5.

274. Shiohama, A., et al., *Molecular cloning and expression analysis of a novel gene DGCR8 located in the DiGeorge syndrome chromosomal region.* Biochem Biophys Res Commun, 2003. **304**(1): p. 184-90.

275. Finnegan, E.F. and A.E. Pasquinelli, *MicroRNA biogenesis: Regulating the Regulators.* Crit Rev Biochem Mol Biol, 2013. **48**(1): p. 51-68.

276. Han, J., et al., *Posttranscriptional crossregulation between Drosha and DGCR8.* Cell, 2009. **136**(1): p. 75-84.

277. Kadener, S., et al., *Genome-wide identification of targets of the drosha-pasha/DGCR8 complex.* Rna, 2009. **15**(4): p. 537-45.

278. Lund, E., et al., *Nuclear export of microRNA precursors.* Science, 2004. **303**(5654): p. 95-8.

279. Bohnsack, M.T., K. Czaplinski, and D. Gorlich, *Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs.* Rna, 2004. **10**(2): p. 185-91.

280. Zeng, Y. and B.R. Cullen, *Structural requirements for pre-microRNA binding and nuclear export by Exportin 5.* Nucleic Acids Res, 2004. **32**(16): p. 4776-85.

281. Yi, R., et al., *Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs*, in *Genes Dev*. 2003. p. 3011-6.

282. Bernstein, E., et al., *Role for a bidentate ribonuclease in the initiation step of RNA interference.* Nature, 2001. **409**(6818): p. 363-6.

283. Hutvagner, G., et al., *A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA.* Science, 2001. **293**(5531): p. 834-8.

284. Kurzynska-Kokorniak, A., et al., *The many faces of Dicer: the complexity of the mechanisms regulating Dicer gene expression and enzyme activities.* Nucleic Acids Res, 2015. **43**(9): p. 4365-80.

285. Forman, J.J., A. Legesse-Miller, and H.A. Coller, *A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence*, in *Proc Natl Acad Sci U S A*. 2008. p. 14879-84.

286. Chakravarthy, S., et al., *Substrate-specific kinetics of Dicer-catalyzed RNA processing.* J Mol Biol, 2010. **404**(3): p. 392-402.

287. Heo, I., et al., *Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA.* Mol Cell, 2008. **32**(2): p. 276-84.

288. Chendrimada, T.P., et al., *TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing.* Nature, 2005. **436**(7051): p. 740-4.

289. Höck, J. and G. Meister, *The Argonaute protein family.* Genome Biology, 2008. **9**(2): p. 1-8.

290. Ozata, D.M., et al., *PIWI-interacting RNAs: small RNAs with big functions.* Nature Reviews Genetics, 2018. **20**(2): p. 89-108.

291. Carmell, M.A., et al., *The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis.* Genes Dev, 2002. **16**(21): p. 2733-42.

292. Meister, G., *Argonaute Proteins: Functional Insights and Emerging Roles.* Nature reviews. Genetics, 2013. **14**(7).

293. M, J. and D. JA, *A three-dimensional view of the molecular machinery of RNA interference.* Nature, 2009. **457**(7228).

294. Chu, Y., et al., *Argonaute Binding Within 3'-untranslated Regions Poorly Predicts Gene Repression.* Nucleic acids research, 2020.

295. Liu, J., et al., *Argonaute2 Is the Catalytic Engine of Mammalian RNAi.* Science (New York, N.Y.), 2004. **305**(5689).

296. Meister, G., et al., *Human Argonaute2 Mediates RNA Cleavage Targeted by miRNAs and siRNAs.* Molecular cell, 2004. **15**(2).

297. Su, H., et al., *Essential and overlapping functions for mammalian Argonautes in microRNA silencing.* Genes Dev, 2009. **23**(3): p. 304-17.

298. Azuma-Mukai, A., et al., *Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing.* Proc Natl Acad Sci U S A, 2008. **105**(23): p. 7964-9.

299. Dueck, A., et al., *microRNAs associated with the different human Argonaute proteins.* Nucleic Acids Res, 2012. **40**(19): p. 9850-62.

300. Leung, A., et al., *Poly(ADP-ribose) regulates stress responses and microRNA activity in the cytoplasm.* Molecular cell, 2011. **42**(4).

301. Zeng, Y., et al., *Phosphorylation of Argonaute 2 at serine-387 Facilitates Its Localization to Processing Bodies.* The Biochemical journal, 2008. **413**(3).

302. Rüdel, S., et al., *Phosphorylation of human Argonaute proteins affects small RNA binding*, in *Nucleic Acids Res*. 2011. p. 2330-43.

303. Qi, Y., et al., *Distinct Catalytic and Non-Catalytic Roles of ARGONAUTE4 in RNA-directed DNA Methylation.* Nature, 2006. **443**(7114).

304. Zilberman, D., et al., *Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA Methylation Triggered by Inverted Repeats.* Current biology : CB, 2004. **14**(13).

305. Janowski, B., et al., *Involvement of AGO1 and AGO2 in Mammalian Transcriptional Silencing.* Nature structural & molecular biology, 2006. **13**(9).

306. Kim, D., et al., *Argonaute-1 Directs siRNA-mediated Transcriptional Gene Silencing in Human Cells.* Nature structural & molecular biology, 2006. **13**(9).

307. Ameyar-Zazoua, M., et al., *Argonaute Proteins Couple Chromatin Silencing to Alternative Splicing.* Nature structural & molecular biology, 2012. **19**(10).

308. Voinnet, O., *Origin, Biogenesis, and Activity of Plant microRNAs.* Cell, 2009. **136**(4).

309. Plotnikova, O., B. A, and S. M, *Comprehensive Analysis of Human microRNA-mRNA Interactome.* Frontiers in genetics, 2019. **10**.

310. Ni, W.J. and X.M. Leng, *Dynamic miRNA–mRNA paradigms: New faces of miRNAs*, in *Biochem Biophys Rep*. 2015. p. 337-41.

311. Grimson, A., et al., *MicroRNA targeting specificity in mammals: determinants beyond seed pairing.* Molecular cell, 2007. **27**(1).

312. Lewis, B., B. CB, and B. DP, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* Cell, 2005. **120**(1).

313. Brennecke, J., et al., *Principles of microRNA-target recognition.* PLoS biology, 2005. **3**(3).

314. Doench, J. and S. PA, *Specificity of microRNA target selection in translational repression.* Genes & development, 2004. **18**(5).

315. Filipowicz, W., B. SN, and S. N, *Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?* Nature reviews. Genetics, 2008. **9**(2).

316. W, F., B. SN, and S. N, *Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?* Nature reviews. Genetics, 2008. **9**(2).

317. Fabian, M., et al., *miRNA-mediated Deadenylation Is Orchestrated by GW182 Through Two Conserved Motifs That Interact With CCR4-NOT.* Nature structural & molecular biology, 2011. **18**(11).

318. Rehwinkel, J., et al., *A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing*, in *RNA*. 2005. p. 1640-7.

319. Humphreys, D., et al., *MicroRNAs Control Translation Initiation by Inhibiting Eukaryotic Initiation Factor 4E/cap and poly(A) Tail Function.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(47).

320. Pillai, R., et al., *Inhibition of Translational Initiation by Let-7 MicroRNA in Human Cells.* Science (New York, N.Y.), 2005. **309**(5740).

321. Kiriakidou, M., et al., *An mRNA m7G Cap Binding-Like Motif Within Human Ago2 Represses Translation.* Cell, 2007. **129**(6).

322. Chendrimada, T., et al., *MicroRNA Silencing Through RISC Recruitment of eIF6.* Nature, 2007. **447**(7146).

323. Olsen, P. and A. V, *The lin-4 Regulatory RNA Controls Developmental Timing in Caenorhabditis Elegans by Blocking LIN-14 Protein Synthesis After the Initiation of Translation.* Developmental biology, 1999. **216**(2).

324. Nottrott, S., S. MJ, and R. JD, *Human let-7a miRNA Blocks Protein Production on Actively Translating Polyribosomes.* Nature structural & molecular biology, 2006. **13**(12).

325. Petersen, C., et al., *Short RNAs Repress Translation After Initiation in Mammalian Cells.* Molecular cell, 2006. **21**(4).

326. Pillai, R., B. SN, and F. W, *Repression of Protein Synthesis by miRNAs: How Many Mechanisms?* Trends in cell biology, 2007. **17**(3).

327. Huntzinger, E. and E. Izaurralde, *Gene silencing by microRNAs: contributions of translational repression and mRNA decay.* Nat Rev Genet, 2011. **12**(2): p. 99-110.

328. Gurtan, A. and S. PA, *The Role of miRNAs in Regulating Gene Expression Networks.* Journal of molecular biology, 2013. **425**(19).

329. Landgraf, P., et al., *A Mammalian microRNA Expression Atlas Based on Small RNA Library Sequencing.* Cell, 2007. **129**(7).

330. Bartel, D.P., *MicroRNAs: target recognition and regulatory functions.* Cell, 2009. **136**(2): p. 215-33.

331. Tsang, J.S., M.S. Ebert, and A. van Oudenaarden, *Genome-wide dissection of microRNA functions and co-targeting networks using gene-set signatures.* Mol Cell, 2010. **38**(1): p. 140-53.

332. Ebert, M. and S. PA, *Roles for microRNAs in Conferring Robustness to Biological Processes.* Cell, 2012. **149**(3).

333. Hwang, H.W. and J.T. Mendell, *MicroRNAs in cell proliferation, cell death, and tumorigenesis*, in *Br J Cancer*. 2006. p. 776-80.

334. Farh, K., et al., *The Widespread Impact of Mammalian MicroRNAs on mRNA Repression and Evolution.* Science (New York, N.Y.), 2005. **310**(5755).

335. Ludwig, N., et al., *Distribution of miRNA expression across human tissues.* Nucleic acids research, 2016. **44**(8).

336. Bernstein, E., et al., *Dicer is essential for mouse development.* Nature genetics, 2003. **35**(3).

337. Bueno, M. and M. M, *MicroRNAs and the Cell Cycle.* Biochimica et biophysica acta, 2011. **1812**(5).

338. Malhas, A., S. NJ, and V. DJ, *The Nuclear Envelope Can Control Gene Expression and Cell Cycle Progression via miRNA Regulation.* Cell cycle (Georgetown, Tex.), 2010. **9**(3).

339. Yong, P. and C.M. Croce, *The role of MicroRNAs in human cancer.* Signal Transduction and Targeted Therapy, 2016. **1**(1): p. 1-9.

340. Hatfield, S., Shcherbata HR, Fischer KA, Nakahara K, Carthew RW et al. Stem cell division is regulated by the microRNA pathway. Nature 2005, et al., *Stem Cell Division Is Regulated by the microRNA Pathway.* Nature, 2005. **435**(7044).

341. Subramanian, S. and S. CJ, *MicroRNAs as Gatekeepers of Apoptosis.* Journal of cellular physiology, 2010. **223**(2).

342. Cimmino, A., et al., *miR-15 and miR-16 Induce Apoptosis by Targeting BCL2.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(39).

343. Singh, R. and S. N, *Downregulation of BCL2 by miRNAs augments drug-induced apoptosis--a combined computational and experimental approach.* Journal of cell science, 2012. **125**(Pt 6).

344. Zhu, S., et al., *MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1).* The Journal of biological chemistry, 2007. **282**(19).

345. Calin, G., et al., *Frequent Deletions and Down-Regulation of Micro- RNA Genes miR15 and miR16 at 13q14 in Chronic Lymphocytic Leukemia.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(24).

346. Zhang, B., et al., *microRNAs as Oncogenes and Tumor Suppressors.* Developmental biology, 2007. **302**(1).

347. Lu, J., et al., *MicroRNA Expression Profiles Classify Human Cancers.* Nature, 2005. **435**(7043).

348. Kumar, M., et al., *Impaired microRNA Processing Enhances Cellular Transformation and Tumorigenesis.* Nature genetics, 2007. **39**(5).

349. Chu, A.S. and J.R. Friedman, *A role for microRNA in cystic liver and kidney diseases*, in *J Clin Invest.* 2008. p. 3585-7.

350. Sun, E. and S. Y, *MicroRNAs: Small Molecules With Big Roles in Neurodevelopment and Diseases.* Experimental neurology, 2015. **268**.

351. Martino, S., et al., *MicroRNA Implications across Neurodevelopment and Neuropathology.* J Biomed Biotechnol, 2009. **2009**.

352. Dorn, G., *MicroRNAs in Cardiac Disease.* Translational research : the journal of laboratory and clinical medicine, 2011. **157**(4).

353. Van Rooij, E., et al., *A Signature Pattern of Stress-Responsive microRNAs That Can Evoke Cardiac Hypertrophy and Heart Failure.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(48).

354. Ardekani, A.M. and M.M. Naeini, *The Role of MicroRNAs in Human Diseases*, in *Avicenna J Med Biotechnol.* 2010. p. 161-79.

355. Kim, M. and X. Zhang, *The Profiling and Role of miRNAs in Diabetes Mellitus.* J Diabetes Clin Res, 2019. **1**(1): p. 5-23.

356. Kaviani, M., et al., *The Role of microRNAs in Islet β-Cell Development.* Cell biology international, 2016. **40**(12).

357. Conrad, E., S. R, and H. CS, *Revealing Transcription Factors During Human Pancreatic β Cell Development.* Trends in endocrinology and metabolism: TEM, 2014. **25**(8).

358. Maestro, M.A., et al., *Hnf6 and Tcf2 (MODY5) are linked in a gene network operating in a precursor cell domain of the embryonic pancreas.* Human Molecular Genetics, 2020. **12**(24): p. 3307-3314.

359. David, A.C., et al., *Transcriptional control of mammalian pancreas organogenesis.* Cellular and Molecular Life Sciences, 2013. **71**(13): p. 2383-2402.

360. Lynn, F., et al., *MicroRNA Expression Is Required for Pancreatic Islet Cell Genesis in the Mouse.* Diabetes, 2007. **56**(12).

361. Kredo-Russo, S., et al., *Pancreas-enriched miRNA Refines Endocrine Cell Differentiation.* Development (Cambridge, England), 2012. **139**(16).

362. Kalis, M., et al., *Beta-cell Specific Deletion of Dicer1 Leads to Defective Insulin Secretion and Diabetes Mellitus.* PloS one, 2011. **6**(12).

363. Murtaza, S.K., M.G. Martin, and A. Bhushan, *Dicer1 Is Required to Repress Neuronal Fate During Endocrine Cell Maturation.* 2013.

364. Melkman-Zehavi, T., et al., *miRNAs Control Insulin Content in Pancreatic β-Cells via Downregulation of Transcriptional Repressors.* The EMBO journal, 2011. **30**(5).

365. Martinez-Sanchez, A., M.S. Nguyen-Tu, and G.A. Rutter, *DICER Inactivation Identifies Pancreatic β-Cell "Disallowed" Genes Targeted by MicroRNAs*, in *Mol Endocrinol*. 2015. p. 1067-79.

366. Wong, W., et al., *Non-Coding RNA in Pancreas and β-Cell Development.* Non-coding RNA, 2018. **4**(4).

367. Wei, R., et al., *Dynamic Expression of microRNAs During the Differentiation of Human Embryonic Stem Cells Into Insulin-Producing Cells.* Gene, 2013. **518**(2).

368. Poy, M.N., et al., *miR-375 maintains normal pancreatic α- and β-cell mass*, in *Proc Natl Acad Sci U S A*. 2009. p. 5813-8.

369. Correa-Medina, M., et al., *MicroRNA miR-7 Is Preferentially Expressed in Endocrine Cells of the Developing and Adult Human Pancreas.* Gene expression patterns : GEP, 2009. **9**(4).

370. Keller, D., et al., *Characterization of Pancreatic Transcription Factor Pdx-1 Binding Sites Using Promoter Microarray and Serial Analysis of Chromatin Occupancy.* The Journal of biological chemistry, 2007. **282**(44).

371. Kloosterman, W., et al., *Targeted Inhibition of miRNA Maturation With Morpholinos Reveals a Role for miR-375 in Pancreatic Islet Development.* PLoS biology, 2007. **5**(8).

372. Bravo-Egana, V., et al., *Quantitative Differential Expression Analysis Reveals Mir-7 As Major Islet MicroRNA.* Biochem Biophys Res Commun, 2008. **366**(4): p. 922-6.

373. Joglekar, M., J. VM, and H. AA, *Expression of Islet-Specific microRNAs During Human Pancreatic Development.* Gene expression patterns : GEP, 2009. **9**(2).

374. Nieto, M., et al., *Antisense miR-7 Impairs Insulin Expression in Developing Pancreas and in Cultured Pancreatic Buds:.* http://dx.doi.org/10.3727/096368911X612521, 2012.

375. Esguerra, J., et al., *MicroRNAs in Islet Hormone Secretion.* Diabetes, obesity & metabolism, 2018. **20 Suppl 2**.

376. Nielsen, D., et al., *Control of Insulin Gene Expression in Pancreatic Beta-Cells and in an Insulin-Producing Cell Line, RIN-5F Cells. I. Effects of Glucose and Cyclic AMP on the Transcription of Insulin mRNA.* The Journal of biological chemistry, 1985. **260**(25).

377. Latreille, M., et al., *MicroRNA-7a Regulates Pancreatic β Cell Function.* The Journal of clinical investigation, 2014. **124**(6).

378. Ahmad, Z., et al., *Pax6 Inactivation in the Adult Pancreas Reveals Ghrelin as Endocrine Cell Maturation Marker*, in *PLoS One*. 2015.

379. Abdelfattah, E.O., et al., *miR-375 Targets 3'-Phosphoinositide–Dependent Protein Kinase-1 and Regulates Glucose-Induced Biological Responses in Pancreatic β-Cells.* 2008.

380. Poy, M., et al., *A Pancreatic Islet-Specific microRNA Regulates Insulin Secretion.* Nature, 2004. **432**(7014).

381. Plaisance, V., et al., *MicroRNA-9 Controls the Expression of Granuphilin/Slp4 and the Secretory Response of Insulin-producing Cells.* 2006.

382. Roggli, E., et al., *Changes in MicroRNA Expression Contribute to Pancreatic β-Cell Dysfunction in Prediabetic NOD Mice*, in *Diabetes*. 2012. p. 1742-51.

383. Lovis, P., G. S, and R. R, *Regulation of the Expression of Components of the Exocytotic Machinery of Insulin-Secreting Cells by microRNAs.* Biological chemistry, 2008. **389**(3).

384. Jing, G., et al., *Thioredoxin-interacting Protein Promotes Islet Amyloid Polypeptide Expression through miR-124a and FoxA2.* 2014.

385. Baroukh, N., et al., *MicroRNA-124a regulates Foxa2 expression and intracellular signaling in pancreatic beta-cell lines.* The Journal of biological chemistry, 2007. **282**(27).

386. Masanari, M., et al., *Noc2 is essential in normal regulation of exocytosis in endocrine and exocrine cells.* 2004.

387. Kasai, K., et al., *Rab27a mediates the tight docking of insulin granules onto the plasma membrane during glucose stimulation*, in *J Clin Invest*. 2005. p. 388-96.

388. Bagge, A., D. CM, and D. LT, *Syntaxin-1a Is a Direct Target of miR-29a in Insulin-Producing β-cells.* Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme, 2013. **45**(6).

389. Esguerra, J.L.S., et al., *Differential Glucose-Regulation of MicroRNAs in Pancreatic Islets of Non-Obese Type 2 Diabetes Model Goto-Kakizaki Rat*, in *PLoS One*. 2011.

390. Tomas, A., et al., *Munc 18-1 and Granuphilin Collaborate During Insulin Granule Exocytosis.* Traffic (Copenhagen, Denmark), 2008. **9**(5).

391. Gulyás-Kovács, A., et al., *Munc18-1: Sequential Interactions With the Fusion Machinery Stimulate Vesicle Docking and Priming.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2007. **27**(32).

392. ZW, Z., et al., *MicroRNA-19b downregulates insulin 1 through targeting transcription factor NeuroD1.* FEBS letters, 2011. **585**(16).

393. Pullen, T., et al., *miR-29a and miR-29b contribute to pancreatic beta-cell-specific silencing of monocarboxylate transporter 1 (Mct1).* Molecular and cellular biology, 2011. **31**(15).

394. Hennessy, E., et al., *Identification of microRNAs with a role in glucose stimulated insulin secretion by expression profiling of MIN6 cells.* Biochemical and biophysical research communications, 2010. **396**(2).

395. Rottiers, V. and A.M. Näär, *MicroRNAs in metabolism and metabolic disorders.* Nature Reviews Molecular Cell Biology, 2012. **13**(4): p. 239-250.

396. Guay, C. and R. R, *Role of Islet microRNAs in Diabetes: Which Model for Which Question?* Diabetologia, 2015. **58**(3).

397. Nesca, V., et al., *Identification of Particular Groups of microRNAs That Positively or Negatively Impact on Beta Cell Function in Obese Models of Type 2 Diabetes.* Diabetologia, 2013. **56**(10).

398. Roggli, E., et al., *Involvement of microRNAs in the Cytotoxic Effects Exerted by Proinflammatory Cytokines on Pancreatic Beta-Cells.* Diabetes, 2010. **59**(4).

399. Delic, D., et al., *Characterization of Micro-RNA Changes during the Progression of Type 2 Diabetes in Zucker Diabetic Fatty Rats.* International journal of molecular sciences, 2016. **17**(5).

400. Tattikota, S., et al., *Argonaute2 Mediates Compensatory Expansion of the Pancreatic β Cell.* Cell metabolism, 2014. **19**(1).

401. Jacovetti, C., et al., *MicroRNAs Contribute to Compensatory β Cell Expansion During Pregnancy and Obesity.* The Journal of clinical investigation, 2012. **122**(10).

402. Tattikota, S., et al., *miR-184 Regulates Pancreatic β-Cell Function According to Glucose Metabolism.* The Journal of biological chemistry, 2015. **290**(33).

403. Kameswaran, V., et al., *Epigenetic Regulation of the DLK1-MEG3 microRNA Cluster in Human Type 2 Diabetic Islets.* Cell metabolism, 2014. **19**(1).

404. Kusari, J., et al., *Analysis of the Gene Sequences of the Insulin Receptor and the Insulin-Sensitive Glucose Transporter (GLUT-4) in Patients With Common-Type Non-Insulin-Dependent Diabetes Mellitus.* The Journal of clinical investigation, 1991. **88**(4).

405. Højlund, K., et al., *Partial Rescue of in Vivo Insulin Signalling in Skeletal Muscle by Impaired Insulin Clearance in Heterozygous Carriers of a Mutation in the Insulin Receptor Gene.* Diabetologia, 2006. **49**(8).

406. Hashimoto, N. and T. T, *Role of miRNAs in the Pathogenesis and Susceptibility of Diabetes Mellitus.* Journal of human genetics, 2017. **62**(2).

407. Wirsing, A., et al., *A Systematic Analysis of the 3'UTR of HNF4A mRNA Reveals an Interplay of Regulatory Elements Including miRNA Target Sites.* PloS one, 2011. **6**(11).

408. Van de Bunt, M., et al., *The miRNA Profile of Human Pancreatic Islets and Beta-Cells and Relationship to Type 2 Diabetes Pathogenesis.* PloS one, 2013. **8**(1).

409. Saunders, M.A., H. Liang, and W.H. Li, *Human polymorphism at microRNAs and microRNA target sites*, in *Proc Natl Acad Sci U S A*. 2007. p. 3300-5.

410. Elfaki, I., et al., *Potential Impact of MicroRNA Gene Polymorphisms in the Pathogenesis of Diabetes and Atherosclerotic Cardiovascular Disease*, in *J Pers Med*. 2019.

411. Assmann, T., et al., *Polymorphisms in Genes Encoding miR-155 and miR-146a Are Associated With Protection to Type 1 Diabetes Mellitus.* Acta diabetologica, 2017. **54**(5).

412. Kaidonis, G., et al., *A Single-Nucleotide Polymorphism in the MicroRNA-146a Gene Is Associated With Diabetic Nephropathy and Sight-Threatening Diabetic Retinopathy in Caucasian Patients.* Acta diabetologica, 2016. **53**(4).

413. Guay, C. and R. Regazzi, *Circulating microRNAs as novel biomarkers for diabetes mellitus.* Nature Reviews Endocrinology, 2013. **9**(9): p. 513-521.

414. M, K. and Z. X, *The Profiling and Role of miRNAs in Diabetes Mellitus.* Journal of diabetes and clinical research, 2019. **1**(1).

415. Zhang, T., et al., *Plasma miR-126 Is a Potential Biomarker for Early Prediction of Type 2 Diabetes Mellitus in Susceptible Individuals.* BioMed research international, 2013. **2013**.

416. Rezk, N., S. NA, and S. MS, *Role of MicroRNA 126 in Screening, Diagnosis, and Prognosis of Diabetic Patients in Egypt.* IUBMB life, 2016. **68**(6).

417. Zampetaki, A., et al., *Plasma microRNA Profiling Reveals Loss of Endothelial miR-126 and Other microRNAs in Type 2 Diabetes.* Circulation research, 2010. **107**(6).

418. Olivieri, F., et al., *MiR-21-5p and miR-126a-3p Levels in Plasma and Circulating Angiogenic Cells: Relationship With Type 2 Diabetes Complications.* Oncotarget, 2015. **6**(34).

419. Karolina, D., et al., *MicroRNA 144 Impairs Insulin Signaling by Inhibiting the Expression of Insulin Receptor Substrate 1 in Type 2 Diabetes Mellitus.* PloS one, 2011. **6**(8).

420. Zhu, H. and L. SW, *Identification of microRNA Biomarkers in Type 2 Diabetes: A Meta-Analysis of Controlled Profiling Studies.* Diabetologia, 2015. **58**(5).

421. Fu, X., et al., *MicroRNA-26a Regulates Insulin Sensitivity and Metabolism of Glucose and Lipids.* The Journal of clinical investigation, 2015. **125**(6).

422. Lin, X., et al., *MiR-155 Enhances Insulin Sensitivity by Coordinated Regulation of Multiple Genes in Mice*, in *PLoS Genet*. 2016.

423. Herrera, B.M., et al., *Global microRNA expression profiles in insulin target tissues in a spontaneous rat model of type 2 diabetes*, in *Diabetologia*. 2010. p. 1099-109.

424. Satake, E., et al., *Circulating miRNA Profiles Associated With Hyperglycemia in Patients With Type 1 Diabetes*, in *Diabetes*. 2018. p. 1013-23.

425. Titze-de-Almeida, R., D. C, and T.-d.-A. SS, *The Race of 10 Synthetic RNAi-Based Drugs to the Pharmaceutical Market.* Pharmaceutical research, 2017. **34**(7).

426. Li, Z. and R. TM, *Therapeutic Targeting of microRNAs: Current Status and Future Challenges.* Nature reviews. Drug discovery, 2014. **13**(8).

427. Huynh, C., et al., *Efficient in vivo microRNA targeting of liver metastasis.* Oncogene, 2011. **30**(12).

428. Trang, P., et al., *Regression of murine lung tumors by the let-7 microRNA.* Oncogene, 2010. **29**(11).

429. Gallant-Behm, C., et al., *A MicroRNA-29 Mimic (Remlarsen) Represses Extracellular Matrix Expression and Fibroplasia in the Skin.* The Journal of investigative dermatology, 2019. **139**(5).

430. Belgardt, B., et al., *The microRNA-200 Family Regulates Pancreatic Beta Cell Survival in Type 2 Diabetes.* Nature medicine, 2015. **21**(6).

431. Trajkovski, M., et al., *MicroRNAs 103 and 107 Regulate Insulin Sensitivity.* Nature, 2011. **474**(7353).

432. Frost, R. and O. EN, *Control of Glucose Homeostasis and Insulin Sensitivity by the Let-7 Family of microRNAs.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(52).

433. Tsukita, S., et al., *MicroRNAs 106b and 222 Improve Hyperglycemia in a Mouse Model of Insulin-Deficient Diabetes via Pancreatic β-Cell Proliferation.* EBioMedicine, 2017. **15**.

434. Jordan, S., et al., *Obesity-induced Overexpression of miRNA-143 Inhibits Insulin-Stimulated AKT Activation and Impairs Glucose Metabolism.* Nature cell biology, 2011. **13**(4).

435. Baumann, V. and W. J, *miRNA-based therapies: strategies and delivery platforms for oligonucleotide and non-oligonucleotide agents.* Future medicinal chemistry, 2014. **6**(17).

436. Petri, R. and J. Jakobsson, *Identifying miRNA Targets Using AGO-RIPseq.* Methods Mol Biol, 2018. **1720**: p. 131-140.

437. Tan, S.M., et al., *Sequencing of captive target transcripts identifies the network of regulated genes and functions of primate-specific miR-522.* Cell Rep, 2014. **8**(4): p. 1225-39.

438. Agarwal, V., et al., *Predicting effective microRNA target sites in mammalian mRNAs.* 2015.

439. Yue, D., L. H, and H. Y, *Survey of Computational Algorithms for MicroRNA Target Prediction.* Current genomics, 2009. **10**(7).

440. Helwak, A. and D. Tollervey, *Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH).* Nat Protoc, 2014. **9**(3): p. 711-28.

441. Broughton, J.P., et al., *Pairing beyond the Seed Supports MicroRNA Targeting Specificity.* Mol Cell, 2016. **64**(2): p. 320-333.

442. Baek, D., et al., *The impact of microRNAs on protein output.* Nature, 2008. **455**(7209): p. 64-71.

443. Pinzón, N., et al., *microRNA Target Prediction Programs Predict Many False Positives.* Genome research, 2017. **27**(2).

444. Rojo Arias, J.E. and V. Busskamp, *Challenges in microRNAs' targetome prediction and validation*, in *Neural Regen Res*. 2019. p. 1672-7.

445. Garzon, R., M. G, and C. CM, *Targeting microRNAs in cancer: rationale, strategies and challenges.* Nature reviews. Drug discovery, 2010. **9**(10).

446. Steinkraus, B., T. M, and F. TA, *Tiny giants of gene regulation: experimental strategies for microRNA functional studies.* Wiley interdisciplinary reviews. Developmental biology, 2016. **5**(3).

447. Licatalosi, D., et al., *HITS-CLIP yields genome-wide insights into brain alternative RNA processing.* Nature, 2008. **456**(7221).

448. Chi, S., et al., *Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps.* Nature, 2009. **460**(7254).

449. Hafner, M., et al., *Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP.* Cell, 2010. **141**(1).

450. Hafner, M., et al., *Genome-wide identification of miRNA targets by PAR-CLIP.* Methods (San Diego, Calif.), 2012. **58**(2).

451. Farazi, T., et al., *Identification of distinct miRNA target regulation between breast cancer molecular subtypes using AGO2-PAR-CLIP and patient datasets.* Genome biology, 2014. **15**(1).

452. Broughton, J. and P. AE, *Identifying Argonaute binding sites in Caenorhabditis elegans using iCLIP.* Methods (San Diego, Calif.), 2013. **63**(2).

453. Godlewski, J., et al., *MicroRNA-451 Regulates LKB1/AMPK Signaling and Allows Adaptation to Metabolic Stress in Glioma Cells.* Mol Cell, 2010. **37**(5): p. 620-32.

454. Chen, H., et al., *Micro-RNA-195 and -451 regulate the LKB1/AMPK signaling axis by targeting MO25.* PloS one, 2012. **7**(7).

455. Zhao, G., et al., *miR-148b functions as a tumor suppressor in pancreatic cancer by targeting AMPKα1.* Molecular cancer therapeutics, 2013. **12**(1).

456. Yi, Y., et al., *Transcriptional suppression of AMPKα1 promotes breast cancer metastasis upon oncogene activation.* Proceedings of the National Academy of Sciences of the United States of America, 2020. **117**(14).

457. Zhang, Y., D. G, and F. S, *MicroRNA-301a modulates doxorubicin resistance in osteosarcoma cells by targeting AMP-activated protein kinase alpha 1.* Biochemical and biophysical research communications, 2015. **459**(3).

458. Martinez-Sanchez, A., et al., *MiR-184 Expression Is Regulated by AMPK in Pancreatic Islets.* FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 2018. **32**(5).

459. Nagosa, S., et al., *microRNA-184 Induces a Commitment Switch to Epidermal Differentiation.* Stem cell reports, 2017. **9**(6).

460. Iovino, N., P. A, and G. U, *miR-184 has multiple roles in Drosophila female germline development.* Developmental cell, 2009. **17**(1).

461. Tao, P., et al., *Expression of miR-664 and miR-184 on proliferation, apoptosis and migration of osteosarcoma cells.* Oncology letters, 2019. **17**(2).

462. Zhu, H., et al., *miR-184 Inhibits Tumor Invasion, Migration and Metastasis in Nasopharyngeal Carcinoma by Targeting Notch2.* Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology, 2018. **49**(4).

463. Yu, J., et al., *MicroRNA-184 antagonizes microRNA-205 to maintain SHIP2 levels in epithelia.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(49).

464. Yu, J., et al., *MicroRNA-205 promotes keratinocyte migration via the lipid phosphatase SHIP2.* FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 2010. **24**(10).

465. Nicolas, M., et al., *Notch1 functions as a tumor suppressor in mouse skin.* Nature genetics, 2003. **33**(3).

466. Prentki, M. and N. CJ, *Islet Beta Cell Failure in Type 2 Diabetes.* The Journal of clinical investigation, 2006. **116**(7).

467. Nomura, T., et al., *MeCP2-dependent repression of an imprinted miR-184 released by depolarization.* Human Molecular Genetics, 2008. **17**(8): p. 1192-1199.

468. Morita, S., et al., *MiR-184 regulates insulin secretion through repression of Slc25a22.* PeerJ, 2013. **1**.

469. de Candia, P., et al., *A unique plasma microRNA profile defines type 2 diabetes progression*, in *PLoS One*. 2017.

470. Emmrich, S., et al., *miR-99a/100~125b tricistrons regulate hematopoietic stem and progenitor cell homeostasis by shifting the balance between TGFβ and Wnt signaling.* Genes Dev, 2014. **28**(8): p. 858-74.

471. Xie, B., et al., *miRCancer: a microRNA-cancer association database constructed by text mining on literature.* Bioinformatics (Oxford, England), 2013. **29**(5).

472. Ooi, A., et al., *MicroRNA-125b expands hematopoietic stem cells and enriches for the lymphoid-balanced and lymphoid-biased subsets.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(50).

473. Bousquet, M., et al., *MicroRNA miR-125b causes leukemia.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(50).

474. Lin, K.Y., et al., *miR-125b, a target of CDX2, regulates cell differentiation through repression of the core binding factor in hematopoietic malignancies.* J Biol Chem, 2011. **286**(44): p. 38253-63.

475. Bloomston, M., et al., *MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis.* JAMA, 2007. **297**(17).

476. Sui, M., et al., *Upregulation of miR-125b is associated with poor prognosis and trastuzumab resistance in HER2-positive gastric cancer*, in *Exp Ther Med*. 2017. p. 657-63.

477. Vriens, M., et al., *MicroRNA expression profiling is a potential diagnostic tool for thyroid cancer.* Cancer, 2012. **118**(13).

478. Ferracin, M., et al., *miR-125b targets erythropoietin and its receptor and their expression correlates with metastatic potential and ERBB2/HER2 expression.* Molecular cancer, 2013. **12**(1).

479. Scott, G., et al., *Coordinate suppression of ERBB2 and ERBB3 by enforced expression of micro-RNA miR-125a or miR-125b.* The Journal of biological chemistry, 2007. **282**(2).

480. Guan, Y., et al., *MiR-125b targets BCL3 and suppresses ovarian cancer proliferation.* International journal of cancer, 2011. **128**(10).

481. Liu, L., et al., *miR-125b suppresses the proliferation and migration of osteosarcoma cells through down-regulation of STAT3.* Biochemical and biophysical research communications, 2011. **416**(1-2).

482. Calin, G.A., et al., *Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers*, in *Proc Natl Acad Sci U S A*. 2004. p. 2999-3004.

483. J, M., et al., *Establishment of a pancreatic beta cell line that retains glucose-inducible insulin secretion: special reference to expression of glucose transporter isoforms.* Endocrinology, 1990. **127**(1).

484. Ravassard, P., et al., *A genetically engineered human pancreatic β cell line exhibiting glucose-inducible insulin secretion.* The Journal of clinical investigation, 2011. **121**(9).

485. Buenaventura, T., et al., *Agonist-induced membrane nanodomain clustering drives GLP-1 receptor responses in pancreatic beta cells.* PLoS biology, 2019. **17**(8).

486. Andrews, S. *FastQC: a quality control tool for high throughput sequence data. Available online at: FastQC: A Quality Control Tool for High Throughput Sequence Data [Online].* 2010; Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

487. Patro, R., et al., *Salmon provides fast and bias-aware quantification of transcript expression.* Nature methods, 2017. **14**(4).

488. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biology, 2014. **15**(12): p. 1-21.

489. Anders, S., et al., *Count-based differential expression analysis of RNA sequencing data using R and Bioconductor.* Nature protocols, 2013. **8**(9).

490. Kurata, J. and L. RJ, *MicroRNA-focused CRISPR-Cas9 library screen reveals fitness-associated miRNAs.* RNA (New York, N.Y.), 2018. **24**(7).

491. Kanehisa, M. and G. S, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic acids research, 2000. **28**(1).

492. Hunter, R., et al., *Mechanism of Action of Compound-13: An α1-Selective Small Molecule Activator of AMPK*, in *Chem Biol*. 2014. p. 866-79.

493. Xiao, B., et al., *Structural Basis of AMPK Regulation by Small Molecule Activators.* Nature communications, 2013. **4**.

494. Weitzel, R., et al., *Reduced methyl-CpG protein binding contributing to miR-184 expression in umbilical cord blood CD4 + T-cells.* Leukemia, 2010. **25**(1): p. 169-172.

495. Marson, A., et al., *Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells.* Cell, 2008. **134**(3).

496. Buenrostro, J., et al., *ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide.* Curr Protoc Mol Biol, 2015. **109**: p. 21 29 1-9.

497. Vietri, R.M., et al., *Comparative Hi-C Reveals That CTCF Underlies Evolution of Chromosomal Domain Architecture.* Cell reports, 2015. **10**(8).

498. Somi, K., N.-K. Yu, and B.-K. Kaang, *CTCF as a multifunctional protein in genome regulation and gene expression.* Experimental & Molecular Medicine, 2015. **47**(6).

499. Phillips, J. and C. VG, *CTCF: master weaver of the genome.* Cell, 2009. **137**(7).

500. Furlan-Magaril, M., et al., *An insulator embedded in the chicken α-globin locus regulates chromatin domain configuration and differential gene expression.* Nucleic acids research, 2011. **39**(1).

501. Recillas-Targa, F., et al., *Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(10).

502. Ong, C. and C. VG, *CTCF: an architectural protein bridging genome topology and function.* Nature reviews. Genetics, 2014. **15**(4).

503. Ohlsson, R., R. R, and L. V, *CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease.* Trends in genetics : TIG, 2001. **17**(9).

504. Rhee, H. and P. BF, *Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution.* Cell, 2011. **147**(6).

505. Nakahashi, H., et al., *A genome-wide map of CTCF multivalency redefines the CTCF code.* Cell reports, 2013. **3**(5).

506. Chen, H., et al., *Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome.* PloS one, 2012. **7**(7).

507. A, M.-S., et al., *MiR-184 Expression Is Regulated by AMPK in Pancreatic Islets.* FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 2018. **32**(5).

508. Tsui, S., D. W, and L. L, *CCCTC-binding Factor Mediates Effects of Glucose on Beta Cell Survival.* Cell proliferation, 2014. **47**(1).

509. Wang, R., et al., *CCCTC-binding factor controls its own nuclear transport via regulating the expression of importin 13.* Molecules and cells, 2013. **35**(5).

510. EM, K., et al., *Functional phosphorylation sites in the C-terminal region of the multivalent multifunctional transcriptional factor CTCF.* Molecular and cellular biology, 2001. **21**(6).

511. Kidder, B.L., G. Hu, and K. Zhao, *ChIP-Seq: technical considerations for obtaining high-quality data.* Nature Immunology, 2011. **12**(10): p. 918-922.

512. Lu, T., et al., *The Polycomb-Dependent Epigenome Controls β Cell Dysfunction, Dedifferentiation, and Diabetes.* Cell metabolism, 2018. **27**(6).

513. Ibarra Urizar, A., et al., *Beta-cell dysfunction induced by non-cytotoxic concentrations of Interleukin-1β is associated with changes in expression of beta-cell maturity genes and associated histone modifications.* Molecular and cellular endocrinology, 2019. **496**.

514. Cvetesic, N., et al., *SLIC-CAGE: high-resolution transcription start site mapping using nanogram-levels of total RNA.* Genome research, 2018. **28**(12).

515. Filippova, G., et al., *An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes.* Molecular and cellular biology, 1996. **16**(6).

516. Jaafar, R., et al., *mTORC1 to AMPK switching underlies β-cell metabolic plasticity during maturation and diabetes.* The Journal of clinical investigation, 2019. **129**(10).

517. Takeshi, S., et al., *Involvement of CTCF in transcription regulation of EGR1 at early G1 phase as an architecture factor.* Scientific Reports, 2019. **9**(1): p. 1-12.

518. Sekiya, T., et al., *Mitotic phosphorylation of CCCTC-binding factor (CTCF) reduces its DNA binding activity.* FEBS open bio, 2017. **7**(3).

519. Obenauer, J., C. LC, and Y. MB, *Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs.* Nucleic acids research, 2003. **31**(13).

520. Miao, F., et al., *Profiles of epigenetic histone post-translational modifications at type 1 diabetes susceptible genes.* The Journal of biological chemistry, 2012. **287**(20).

521. Liu, C., et al., *Epigenetic regulation of miR-184 by MBD1 governs neural stem cell proliferation and differentiation.* Cell stem cell, 2010. **6**(5).

522. Messeguer, X., et al., *PROMO: detection of known transcription regulatory elements using species-tailored searches.* Bioinformatics (Oxford, England), 2002. **18**(2).

523. Gearing, L., et al., *CiiiDER: A tool for predicting and analysing transcription factor binding sites.* PloS one, 2019. **14**(9).

524. Sakamoto, S., et al., *The NF90-NF45 complex functions as a negative regulator in the microRNA processing pathway.* Molecular and cellular biology, 2009. **29**(13).

525. Dini Modigliani, S., et al., *An ALS-associated mutation in the FUS 3'-UTR disrupts a microRNA-FUS regulatory circuitry.* Nature communications, 2014. **5**.

526. Kim, N., et al., *AMPKα2 translocates into the nucleus and interacts with hnRNP H: implications in metformin-mediated glucose uptake.* Cellular signalling, 2014. **26**(9).

527. Shaham, L., et al., *MiR-125 in Normal and Malignant Hematopoiesis.* Leukemia, 2012. **26**(9).

528. Sun, Y.M., K.Y. Lin, and Y.Q. Chen, *Diverse functions of miR-125 family in different cell contexts.* J Hematol Oncol, 2013. **6**: p. 6.

529. Shi, X.B., et al., *An androgen-regulated miRNA suppresses Bak1 expression and induces androgen-independent growth of prostate cancer cells.* Proc Natl Acad Sci U S A, 2007. **104**(50): p. 19983-8.

530. Tili, E., et al., *Modulation of miR-155 and miR-125b levels following lipopolysaccharide/TNF-α stimulation and their possible roles in regulating the response to endotoxin shock.* Journal of Immunology, 2007. **179**(8): p. 5082-5089.

531. Zhou, R., et al., *Binding of NF-kappaB p65 subunit to the promoter elements is involved in LPS-induced transactivation of miRNA genes in human biliary epithelial cells*, in *Nucleic Acids Res*. 2010. p. 3222-32.

532. Zhou, R., et al., *NF-kappaB p65-dependent transactivation of miRNA genes following Cryptosporidium parvum infection stimulates epithelial cell immune responses.* PLoS Pathog, 2009. **5**(12): p. e1000681.

533. Ottaviani, S., et al., *TGF-β induces miR-100 and miR-125b but blocks let-7a through LIN28B controlling PDAC progression.* Nature communications, 2018. **9**(1).

534. David, C.J., et al., *TGF-beta Tumor Suppression through a Lethal EMT.* Cell, 2016. **164**(5): p. 1015-30.

535. Dhawan, S., et al., *Inhibition of TGF-beta Signaling Promotes Human Pancreatic beta-Cell Replication.* Diabetes, 2016. **65**(5): p. 1208-18.

536. Sjoholm, A. and C. Hellerstrom, *TGF-beta stimulates insulin secretion and blocks mitogenic response of pancreatic beta-cells to glucose.* Am J Physiol, 1991. **260**(5 Pt 1): p. C1046-51.

537. Lin, H.M., et al., *Transforming growth factor-beta/Smad3 signaling regulates insulin gene transcription and pancreatic islet beta-cell function.* J Biol Chem, 2009. **284**(18): p. 12246-57.

538. Lin, H., et al., *AMPK Inhibits the Stimulatory Effects of TGF-β on Smad2/3 Activity, Cell Migration, and Epithelial-to-Mesenchymal Transition*, in *Mol Pharmacol*. 2015. p. 1062-71.

539. Lim, J., et al., *AMP-activated protein kinase inhibits TGF-β-induced fibrogenic responses of hepatic stellate cells by targeting transcriptional coactivator p300.* Journal of cellular physiology, 2012. **227**(3).

540. Mishra, R., et al., *AMP-activated protein kinase inhibits transforming growth factor-beta-induced Smad3-dependent transcription and myofibroblast transdifferentiation.* The Journal of biological chemistry, 2008. **283**(16).

541. Xin, X., et al., *Transforming growth factor-β in stem cells and tissue homeostasis.* Bone Research, 2018. **6**(1): p. 1-31.

542. Creugny, A., F. A, and P. S, *Regulation of primary microRNA processing.* FEBS letters, 2018. **592**(12).

543. Rane, S.G., J.H. Lee, and H.M. Lin, *Transforming growth factor-beta pathway: role in pancreas development and pancreatic disease.* Cytokine Growth Factor Rev, 2006. **17**(1-2): p. 107-19.

544. Suzuki, T., et al., *TGF-β Signaling Regulates Pancreatic β-Cell Proliferation through Control of Cell Cycle Regulator p27 Expression*, in *Acta Histochem Cytochem*. 2013. p. 51-8.

545. El-Gohary, Y., et al., *Smad signaling pathways regulate pancreatic endocrine development.* Dev Biol, 2013. **378**(2): p. 83-93.

546. Sanvito, F., et al., *TGF-beta 1 influences the relative development of the exocrine and endocrine pancreas in vitro.* Development, 1994. **120**(12): p. 3451-62.

547. Zhang, Y.-Q., et al., *Involvement of Smad proteins in the differentiation of pancreatic AR42J cells induced by activin A.* Diabetologia, 1999. **42**(6): p. 719-727.

548. Lee, J.-H., et al., *Protection from β-cell apoptosis by inhibition of TGF-β/Smad3 signaling.* Cell Death & Disease, 2020. **11**(3): p. 1-15.

549. Wu, H., et al., *Differential regulation of mouse pancreatic islet insulin secretion and Smad proteins by activin ligands.* Diabetologia, 2013. **57**(1): p. 148-156.

550. Kong, W., et al., *MicroRNA-155 is regulated by the transforming growth factor beta/Smad pathway and contributes to epithelial cell plasticity by targeting RhoA.* Molecular and cellular biology, 2008. **28**(22).

551. Chung, A., et al., *miR-192 mediates TGF-beta/Smad3-driven renal fibrosis.* Journal of the American Society of Nephrology : JASN, 2010. **21**(8).

552. Qin, W., et al., *TGF-β/Smad3 signaling promotes renal fibrosis by inhibiting miR-29.* Journal of the American Society of Nephrology : JASN, 2011. **22**(8).

553. Kato, M., et al., *TGF-beta activates Akt kinase through a microRNA-dependent amplifying circuit targeting PTEN.* Nature cell biology, 2009. **11**(7).

554. Sun, Q., et al., *Transforming growth factor-beta-regulated miR-24 promotes skeletal muscle differentiation.* Nucleic acids research, 2008. **36**(8).

555. Gao, Y., et al., *Role of TGF-β/Smad Pathway in the Transcription of Pancreas-Specific Genes During Beta Cell Differentiation.* Frontiers in cell and developmental biology, 2019. **7**.

556. Bai, C., et al., *MicroRNA-34c acts as a bidirectional switch in the maturation of insulin-producing cells derived from mesenchymal stem cells.* Oncotarget, 2017. **8**(63).

557. Bai, C., et al., *MicroRNAs can effectively induce formation of insulin-producing cells from mesenchymal stem cells.* Journal of tissue engineering and regenerative medicine, 2017. **11**(12).

558. Van de Laar, I., et al., *Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis.* Nature genetics, 2011. **43**(2).

559. Zhang, Y., *Non-Smad Signaling Pathways of the TGF-β Family.* Cold Spring Harbor perspectives in biology, 2017. **9**(2).

560. Bertero, A., et al., *The SMAD2/3 interactome reveals that TGFβ controls m 6 A mRNA methylation in pluripotency.* Nature, 2018. **555**(7695).

561. Rosa, A., et al., *miR-373 is regulated by TGFβ signaling and promotes mesendoderm differentiation in human Embryonic Stem Cells.* Developmental biology, 2014. **391**(1).

562. Oida, T. and W. HL, *Depletion of TGF-β from fetal bovine serum.* Journal of immunological methods, 2010. **362**(1-2).

563. Haiya, W., et al., *Differential regulation of mouse pancreatic islet insulin secretion and Smad proteins by activin ligands.* Diabetologia, 2013. **57**(1): p. 148-156.

564. Riboulet-Chavey, A., et al., *Inhibition of AMP-activated protein kinase protects pancreatic beta-cells from cytokine-mediated apoptosis and CD8+ T-cell-induced cytotoxicity.* Diabetes, 2008. **57**(2).

565. Nicole, R., *Reverse ChIP.* Nature Methods, 2020. **6**(3): p. 187-187.

566. Ben-Hamo, R. and E. S, *MicroRNA regulation of molecular pathways as a generic mechanism and as a core disease phenotype.* Oncotarget, 2015. **6**(3).

567. Hsieh, W., et al., *Investigating microRNA-target interaction-supported tissues in human cancer tissues based on miRNA and target gene expression profiling.* PloS one, 2014. **9**(4).

568. Wang, Y., Z. G, and J. Y, *The Emerging Roles of miR-125b in Cancers.* Cancer management and research, 2020. **12**.

569. Le, M., et al., *MicroRNA-125b is a novel negative regulator of p53.* Genes & development, 2009. **23**(7).

570. Nishida, N., et al., *MicroRNA miR-125b is a prognostic marker in human colorectal cancer.* International journal of oncology, 2011. **38**(5).

571. Rasheed, Z., et al., *MicroRNA-125b-5p regulates IL-1β induced inflammatory genes via targeting TRAF6-mediated MAPKs and NF-κB signaling in human osteoarthritic chondrocytes.* Scientific Reports, 2019. **9**(1): p. 1-13.

572. Valmiki, S., et al., *miR-125b and miR-223 Contribute to Inflammation by Targeting the Key Molecules of NFκB Pathway.* Frontiers in medicine, 2020. **6**.

573. Villeneuve., L.M., et al., *Enhanced Levels of microRNA-125b in Vascular Smooth Muscle Cells of Diabetic db/db Mice Lead to Increased Inflammatory Gene Expression by Targeting the Histone Methyltransferase Suv39h1.* 2010.

574. Klein, D., et al., *MicroRNA expression in alpha and beta cells of human pancreatic islets.* PloS one, 2013. **8**(1).

575. Yu, C., Y. CY, and R. ZL, *MicroRNA-125b-5p improves pancreatic β-cell function through inhibiting JNK signaling pathway by targeting DACT1 in mice with type 2 diabetes mellitus.* Life sciences, 2019. **224**.

576. Subramanian, A., et al., *Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.* 2005.

577. Mootha, V.K., et al., *PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.* Nature Genetics, 2003. **34**(3): p. 267-273.

578. Wang, W., et al., *Anti-Argonaute RIP-Chip shows that miRNA transfections alter global patterns of mRNA recruitment to microribonucleoprotein complexes.* RNA (New York, N.Y.), 2010. **16**(2).

579. Perconti, G., et al., *RIP-Chip analysis supports different roles for AGO2 and GW182 proteins in recruiting and processing microRNA targets.* BMC bioinformatics, 2019. **20**(Suppl 4).

580. Goff, L., et al., *Ago2 immunoprecipitation identifies predicted microRNAs in human embryonic stem cells and neural precursors.* PloS one, 2009. **4**(9).

581. Xia, H., et al., *MiR-125b expression affects the proliferation and apoptosis of human glioma cells by targeting Bmf.* Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology, 2009. **23**(4-6).

582. YX, F., et al., *MicroRNA-125b inhibits cell proliferation and induces cell apoptosis in esophageal squamous cell carcinoma by targeting BMF.* Oncology reports, 2018. **40**(1).

583. Mészáros, G., et al., *Lysosomes in nutrient signalling: A focus on pancreatic β-cells.* Diabetes, obesity & metabolism, 2018. **20 Suppl 2**.

584. Goginashvili, A., et al., *Insulin granules. Insulin secretory granules control autophagy in pancreatic β cells.* Science (New York, N.Y.), 2015. **347**(6224).

585. Pasquier, A., et al., *Lysosomal degradation of newly formed insulin granules contributes to β cell failure in diabetes.* Nature Communications, 2019. **10**(1): p. 1-14.

586. Braulke, T. and B. JS, *Sorting of lysosomal proteins.* Biochimica et biophysica acta, 2009. **1793**(4).

587. Coutinho, M., P. MJ, and A. S, *Mannose-6-phosphate pathway: a review on its role in lysosomal function and dysfunction.* Molecular genetics and metabolism, 2012. **105**(4).

588. Perera, R. and Z. R, *The Lysosome as a Regulatory Hub.* Annual review of cell and developmental biology, 2016. **32**.

589.   Settembre, C., et al., *Signals from the lysosome: a control centre for cellular clearance and energy metabolism.* Nature reviews. Molecular cell biology, 2013. **14**(5).

590.   Kuna, R., et al., *Glucagon-like peptide-1 receptor-mediated endosomal cAMP generation promotes glucose-stimulated insulin secretion in pancreatic β-cells.* American journal of physiology. Endocrinology and metabolism, 2013. **305**(2).

591.   Gong, H., et al., *Evaluation of candidate reference genes for RT-qPCR studies in three metabolism related tissues of mice after caloric restriction.* Scientific reports, 2016. **6**.

592.   Agarwal, V., et al., *Predicting effective microRNA target sites in mammalian mRNAs.* eLife, 2015. **4**.

593.   Svoboda, P., *A toolbox for miRNA analysis.* FEBS letters, 2015. **589**(14).

594.   Martinez-Sanchez, A. and M. CL, *MicroRNA Target Identification-Experimental Approaches.* Biology, 2013. **2**(1).

595.   Linsley, P., et al., *Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression.* Molecular and cellular biology, 2007. **27**(6).

596.   Khan, A., et al., *Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs.* Nature biotechnology, 2009. **27**(6).

597.   Fan, Y., et al., *MicroRNA-125b inhibits cell proliferation and induces cell apoptosis in esophageal squamous cell carcinoma by targeting BMF.* Oncology reports, 2018. **40**(1).

598.   Huang, T., P. SM, and P. A, *Proteomics for understanding miRNA biology.* Proteomics, 2013. **13**(3-4).

599.   Bhattacharyya, S., et al., *Relief of microRNA-mediated translational repression in human cells subjected to stress.* Cell, 2006. **125**(6).

600.   Chen, G., et al., *Discordant protein and mRNA expression in lung adenocarcinomas.* Molecular & cellular proteomics : MCP, 2002. **1**(4).

601.   Griffin, T., et al., *Complementary profiling of gene expression at the transcriptome and proteome levels in Saccharomyces cerevisiae.* Molecular & cellular proteomics : MCP, 2002. **1**(4).

602.   Chase, A., L. E, and S. C, *Torsin ATPases: Harnessing Dynamic Instability for Function.* Frontiers in molecular biosciences, 2017. **4**.

603.   Rose, A., B. RS, and S. C, *Torsins: not your typical AAA+ ATPases.* Critical reviews in biochemistry and molecular biology, 2015. **50**(6).

604.   Ofman, R., L. S, and W. RJ, *Etherphospholipid biosynthesis and dihydroxyactetone-phosphate acyltransferase: resolution of the genomic organization of the human gnpat gene and its use in the identification of novel mutations.* Biochemical and biophysical research communications, 2001. **281**(3).

605.   Malheiro, A.R., T.F.d. Silva, and P. Brites, *Plasmalogens and fatty alcohols in rhizomelic chondrodysplasia punctata and Sjögren-Larsson syndrome.* Journal of Inherited Metabolic Disease, 2014. **38**(1): p. 111-121.

606.   Braverman, N. and M. AB, *Functions of plasmalogen lipids in health and disease.* Biochimica et biophysica acta, 2012. **1822**(9).

607.   Weir, J.M., et al., *Plasma lipid profiling in a large population-based cohort.* 2013.

608.   Acehan, D., et al., *Cardiac and skeletal muscle defects in a mouse model of human Barth syndrome.* The Journal of biological chemistry, 2011. **286**(2).

609.   Soustek, M., et al., *Characterization of a transgenic short hairpin RNA-induced murine model of Tafazzin deficiency.* Human gene therapy, 2011. **22**(7).

610.   Phoon, C., et al., *Tafazzin knockdown in mice leads to a developmental cardiomyopathy with early diastolic dysfunction preceding myocardial noncompaction.* Journal of the American Heart Association, 2012. **1**(2).

611.   Chen, S., H. Q, and G. ML, *Loss of tafazzin in yeast leads to increased oxidative stress during respiratory growth.* Molecular microbiology, 2008. **68**(4).

612. Dudek, J., et al., *Cardiac-specific succinate dehydrogenase deficiency in Barth syndrome.* EMBO molecular medicine, 2016. **8**(2).

613. Wang, G., et al., *Modeling the mitochondrial cardiomyopathy of Barth syndrome with induced pluripotent stem cell and heart-on-chip technologies.* Nature medicine, 2014. **20**(6).

614. Tondera, D., et al., *The mitochondrial protein MTP18 contributes to mitochondrial fission in mammalian cells.* Journal of cell science, 2005. **118**(Pt 14).

615. Morita, M., et al., *mTOR Controls Mitochondrial Dynamics and Cell Survival via MTFP1.* Molecular cell, 2017. **67**(6).

616. Tondera, D., et al., *Knockdown of MTP18, a novel phosphatidylinositol 3-kinase-dependent protein, affects mitochondrial morphology and induces apoptosis.* The Journal of biological chemistry, 2004. **279**(30).

617. Duroux-Richard, I., et al., *miR-125b controls monocyte adaptation to inflammation through mitochondrial metabolism and dynamics.* Blood, 2016. **128**(26).

618. Giroud, M., et al., *miR-125b affects mitochondrial biogenesis and impairs brite adipocyte formation and function.* Molecular metabolism, 2016. **5**(8).

619. Herzig, S. and S. RJ, *AMPK: guardian of metabolism and mitochondrial homeostasis.* Nature reviews. Molecular cell biology, 2018. **19**(2).

620. Bergeron, R., et al., *Chronic activation of AMP kinase results in NRF-1 activation and mitochondrial biogenesis.* American journal of physiology. Endocrinology and metabolism, 2001. **281**(6).

621. O'Neill, H., et al., *AMP-activated protein kinase (AMPK) beta1beta2 muscle null mice reveal an essential role for AMPK in maintaining mitochondrial content and glucose uptake during exercise.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(38).

622. Wang, C. and Y. R, *Cell biology: Form follows function for mitochondria.* Nature, 2016. **530**(7590).

623. Egan, D., et al., *Phosphorylation of ULK1 (hATG1) by AMP-activated protein kinase connects energy sensing to mitophagy.* Science (New York, N.Y.), 2011. **331**(6016).

624. Gerber, P. and R. GA, *The Role of Oxidative Stress and Hypoxia in Pancreatic Beta-Cell Dysfunction in Diabetes Mellitus.* Antioxidants & redox signaling, 2017. **26**(10).

625. Ludwig, T., et al., *Targeted disruption of the mouse cation-dependent mannose 6-phosphate receptor results in partial missorting of multiple lysosomal enzymes.* EMBO J, 1993. **12**(13): p. 5225-35.

626. Raza, M., et al., *Mucolipidosis types II and III and non-syndromic stuttering are associated with different variants in the same genes.* European journal of human genetics : EJHG, 2016. **24**(4).

627. De Araujo, M., et al., *Lysosomal size matters.* Traffic (Copenhagen, Denmark), 2020. **21**(1).

628. Jang, M., et al., *AMPK contributes to autophagosome maturation and lysosomal fusion.* Scientific reports, 2018. **8**(1).

629. Marasco, M. and L. AK, *β-Cell Autophagy in Diabetes Pathogenesis.* Endocrinology, 2018. **159**(5).

630. Huppertz, I., et al., *iCLIP: protein-RNA interactions at nucleotide resolution.* Methods (San Diego, Calif.), 2014. **65**(3).

631. Van Nostrand, E., et al., *Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP).* Nature methods, 2016. **13**(6).

632. Sugimoto, Y., et al., *Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions.* Genome biology, 2012. **13**(8).

633. I, H., et al., *iCLIP: protein-RNA interactions at nucleotide resolution.* Methods (San Diego, Calif.), 2014. **65**(3).

634. Polyak, K., *The p27Kip1 tumor suppressor gene: Still a suspect or proven guilty?* Cancer cell, 2006. **10**(5).

635. Chu, I., H. L, and S. JM, *The Cdk inhibitor p27 in human cancer: prognostic potential and relevance to anticancer therapy.* Nature reviews. Cancer, 2008. **8**(4).

636. Cheng, M., et al., *The p21(Cip1) and p27(Kip1) CDK 'inhibitors' are essential activators of cyclin D-dependent kinases in murine fibroblasts.* The EMBO journal, 1999. **18**(6).

637. Fujita, N., et al., *Akt-dependent phosphorylation of p27Kip1 promotes binding to 14-3-3 and cytoplasmic localization.* The Journal of biological chemistry, 2002. **277**(32).

638. Nakayama, K. and N. K, *Ubiquitin ligases: cell-cycle control and cancer.* Nature reviews. Cancer, 2006. **6**(5).

639. Martínez-Sánchez, A. and G. F, *Regulation of p27(kip1) mRNA expression by microRNAs.* Progress in molecular and subcellular biology, 2010. **50**.

640. Cuesta, R., M.-S. A, and G. F, *miR-181a regulates cap-dependent translation of p27(kip1) mRNA in myeloid cells.* Molecular and cellular biology, 2009. **29**(10).

641. Oboshi, W., et al., *MicroRNA-150 suppresses p27 Kip1 expression and promotes cell proliferation in HeLa human cervical cancer cells.* Oncology letters, 2020. **20**(5).

642. Visone, R., et al., *MicroRNAs (miR)-221 and miR-222, both overexpressed in human thyroid papillary carcinomas, regulate p27Kip1 protein levels and cell cycle.* Endocrine-related cancer, 2007. **14**(3).

643. Kishore, S., et al., *A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins.* Nature Methods, 2011. **8**(7): p. 559-564.

644. Friedersdorf, M.B. and J.D. Keene, *Advancing the functional utility of PAR-CLIP by quantifying background binding to mRNAs and lncRNAs.* Genome Biology, 2014. **15**(1): p. 1-16.

645. Schraivogel, D., et al., *Importin-β facilitates nuclear import of human GW proteins and balances cytoplasmic gene silencing protein levels.* Nucleic Acids Research, 2015. **43**(15): p. 7447-7461.

646. Alló, M., et al., *Control of alternative splicing through siRNA-mediated transcriptional gene silencing.* Nature structural & molecular biology, 2009. **16**(7).

647. Grishok, A., S. JL, and S. PA, *Transcriptional silencing of a transgene by RNAi in the soma of C. elegans.* Genes & development, 2005. **19**(6).

648. Matzke, M. and B. JA, *RNAi-mediated pathways in the nucleus.* Nature reviews. Genetics, 2005. **6**(1).

649. Martienssen, R., Z. M, and G. DB, *RNA interference and heterochromatin in the fission yeast Schizosaccharomyces pombe.* Trends in genetics : TIG, 2005. **21**(8).

650. Alcid, E. and T. T, *Expansion of antisense lncRNA transcriptomes in budding yeast species since the loss of RNAi.* Nature structural & molecular biology, 2016. **23**(5).

651. Carissimi, C., et al., *ARGONAUTE2 cooperates with SWI/SNF complex to determine nucleosome occupancy at human Transcription Start Sites.* Nucleic acids research, 2015. **43**(3).

652. Miguel-Escalada, I., et al., *Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes.* Nature genetics, 2019. **51**(7).

653. Helwak, A., et al., *Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding.* Cell, 2013. **153**(3): p. 654-65.

654. Cokorinos, E., et al., *Activation of Skeletal Muscle AMPK Promotes Glucose Disposal and Glucose Lowering in Non-human Primates and Mice.* Cell metabolism, 2017. **25**(5).

655. Schanen, B. and L. X, *Transcriptional regulation of mammalian miRNA genes.* Genomics, 2011. **97**(1).

656. Kefas, B., et al., *AICA-riboside induces apoptosis of pancreatic beta cells through stimulation of AMP-activated protein kinase.* Diabetologia, 2003. **46**(2).

657. Lee, Y., et al., *β-cell autophagy: Mechanism and role in β-cell dysfunction.* Molecular metabolism, 2019. **27S**(Suppl).

658. Orci, L., et al., *Insulin, not C-peptide (proinsulin), is present in crinophagic bodies of the pancreatic B-cell.* The Journal of cell biology, 1984. **98**(1).

659. Müller, A., et al., *A 4D view on insulin secretory granule turnover in the β-cell.* Diabetes, obesity & metabolism, 2017. **19 Suppl 1**.

660. Binger, K.J., et al., *Atp6ap2 deletion causes extensive vacuolation that consumes the insulin content of pancreatic β cells.* 2019.

661. Røder, M., et al., *Disproportionately elevated proinsulin levels reflect the degree of impaired B cell secretory capacity in patients with noninsulin-dependent diabetes mellitus.* The Journal of clinical endocrinology and metabolism, 1998. **83**(2).

662. Ward, W., et al., *Disproportionate elevation of immunoreactive proinsulin in type 2 (non-insulin-dependent) diabetes mellitus and in experimental insulin resistance.* Diabetologia, 1987. **30**(9).

663. Parenti, G., M. DL, and B. A, *The rapidly evolving view of lysosomal storage diseases.* EMBO molecular medicine, 2021. **13**(2).

664. Maechler, P. and W. CB, *Mitochondrial function in normal and diabetic beta-cells.* Nature, 2001. **414**(6865).

665. Malaisse, W., et al., *The stimulus-secretion coupling of glucose-induced insulin release. XXXV. The links between metabolic and cationic events.* Diabetologia, 1979. **16**(5).

666. Haythorne, E., et al., *Diabetes causes marked inhibition of mitochondrial metabolism in pancreatic β-cells.* Nature communications, 2019. **10**(1).

667. Ma, Z.A., Z. Zhao, and J. Turk, *Mitochondrial Dysfunction and β-Cell Failure in Type 2 Diabetes Mellitus.* Experimental Diabetes Research, 2011. **2012**.

668. Sakai, K., et al., *Mitochondrial reactive oxygen species reduce insulin secretion by pancreatic beta-cells.* Biochemical and biophysical research communications, 2003. **300**(1).

669. Ranieri, M., et al., *Mitochondrial Fusion Proteins and Human Diseases.* Neurology Research International, 2013. **2013**.

670. Yu, T., R. JL, and Y. Y, *Increased production of reactive oxygen species in hyperglycemic conditions requires dynamic change of mitochondrial morphology.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(8).

671. Miyazaki, J., et al., *Establishment of a pancreatic beta cell line that retains glucose-inducible insulin secretion: special reference to expression of glucose transporter isoforms.* Endocrinology, 1990. **127**(1).

672. Ishihara, H., et al., *Pancreatic beta cell line MIN6 exhibits characteristics of glucose metabolism and glucose-stimulated insulin secretion similar to those of normal islets.* Diabetologia, 1993. **36**(11).

673. Wei, L., et al., *MiR-125b-2 knockout increases high-fat diet-induced fat accumulation and insulin resistance.* Scientific reports, 2020. **10**(1).

674. Moss, N. and S. L, *mRNA Processing: An Emerging Frontier in the Regulation of Pancreatic β Cell Function.* Frontiers in genetics, 2020. **11**.

675. Jeffery, N., et al., *Cellular stressors may alter islet hormone cell proportions by moderation of alternative splicing patterns.* Human molecular genetics, 2019. **28**(16).

Appendix A: Supplementary Material

Supplementary Table 1: list of miR-125b direct targets identified using IP-RNA/T-RNA ratio

	Gene.name	FC_T-RNA	padj_T-RNA	FC_IP-RNA	padj_IP-RNA	IP-RNA/ T-RNA	TargetScan	RNAhybrid
1	Taz	0.9	2.13E-01	2.2	1.31E-05	2.5	x	
2	M6pr	0.8	5.50E-03	2	1.31E-11	2.5	x	
3	Tor2a	0.7	1.55E-03	1.8	8.19E-05	2.5	x	
4	Mtfp1	0.8	1.80E-02	1.9	3.34E-04	2.4	x	
5	Tnks1bp1	0.9	6.07E-01	2.2	2.42E-06	2.4		CDS
6	Gnpat	0.9	4.40E-01	2.1	7.87E-16	2.3		CDS
7	Gm47356	0.9	6.02E-01	2	2.47E-03	2.3		3'UTR
8	Trnp1	1	9.68E-01	2.2	1.03E-09	2.2		CDS
9	Rhot2	1	9.82E-01	2.1	2.82E-07	2.1	x	
10	Lsm14b	0.9	6.11E-01	1.8	3.61E-05	2		CDS
11	Mlxipl	0.9	7.43E-01	1.8	3.43E-05	2		3'UTR
12	Lrrc10b	0.8	2.92E-02	1.5	5.56E-02	2	x	
13	Trabd	0.9	4.16E-01	1.7	6.82E-04	2		CDS
14	Ceacam1	0.7	1.86E-03	1.3	6.08E-01	1.9	x	
15	Cpne5	0.7	1.58E-02	1.3	6.13E-01	1.9	x	
16	Abtb1	0.8	2.42E-01	1.6	2.50E-02	1.9	x	
17	Nectin2	0.8	1.14E-01	1.6	4.95E-02	1.9		CDS
18	Gm10125	0.9	8.18E-01	1.7	3.02E-02	1.9		
19	Pex11g	0.8	6.01E-02	1.4	4.86E-01	1.9		
20	Dusp23	0.7	6.32E-03	1.3	7.68E-01	1.9	x	
21	Kcnc2	0.6	4.77E-06	1.1	9.99E-01	1.9	x	
22	Guk1	1	8.96E-01	1.9	3.43E-05	1.9		
23	Mfge8	0.7	1.33E-02	1.3	7.71E-01	1.9	x	
24	Rnf113a2	1.1	4.94E-01	2	6.60E-04	1.8		
25	Stc1	0.8	1.42E-01	1.4	4.06E-01	1.8	x	
26	Src	0.7	1.08E-02	1.3	8.52E-01	1.8	x	
27	Nrm	0.7	3.92E-07	1.2	9.63E-01	1.8	x	
28	Pgap3	0.7	2.08E-04	1.3	8.02E-01	1.8	x	
29	Coch	1	9.54E-01	1.8	3.84E-03	1.8		
30	Rabl6	0.9	3.31E-01	1.6	1.37E-03	1.8	x	
31	Apoo	1.1	6.89E-01	1.9	3.84E-03	1.8		
32	Cyth1	0.9	5.14E-01	1.7	6.47E-04	1.8	x	
33	Gm17167	0.8	4.45E-01	1.5	3.64E-01	1.8		
34	Cd34	0.8	1.23E-01	1.3	6.62E-01	1.8	x	
35	Cbx7	0.8	1.30E-01	1.4	3.79E-01	1.8	x	
36	Brms1	0.8	1.10E-01	1.4	4.28E-01	1.8	x	

37	Grk4	0.9	7.87E-01	1.6	1.75E-01	1.7		
38	Rnf44	0.8	3.38E-01	1.5	2.50E-02	1.7	x	
39	Pstpip2	0.7	2.19E-03	1.2	9.99E-01	1.7		
40	Syvn1	0.9	3.93E-01	1.5	2.06E-05	1.7	x	
41	Agtrap	0.9	6.01E-01	1.5	2.58E-01	1.7	x	
42	Fam83h	0.7	1.12E-03	1.3	5.43E-01	1.7	x	
43	Them6	0.6	9.79E-04	1.1	9.99E-01	1.7	x	
44	Nudt16	0.8	2.34E-01	1.5	4.02E-01	1.7		
45	Vipr1	0.9	4.99E-01	1.5	4.55E-04	1.7	x	
46	Lypla2	0.9	2.04E-01	1.5	3.68E-02	1.7	x	
47	Gm43960	1	9.38E-01	1.8	2.94E-02	1.7		
48	G6pc3	0.9	5.17E-01	1.5	1.14E-01	1.7		
49	Dis3l2	0.8	1.17E-02	1.4	6.40E-01	1.7	x	
50	Ccdc50	0.8	7.34E-02	1.4	2.33E-01	1.7		
51	Sirt6	0.8	1.87E-02	1.4	6.07E-01	1.7	x	
52	Dnajc14	0.8	1.62E-03	1.4	3.29E-01	1.7	x	
53	Mlycd	0.8	2.50E-03	1.3	7.75E-01	1.7	x	
54	Gm45844	0.9	7.50E-01	1.5	1.96E-01	1.7		
55	Aftph	1	7.47E-01	1.6	1.80E-02	1.7		
56	Dph7	0.9	7.74E-01	1.6	2.09E-01	1.7		
57	Ptpn5	0.7	1.11E-02	1.2	9.99E-01	1.7		
58	Pgp	1	9.80E-01	1.7	3.84E-03	1.7	x	
59	Car8	0.6	3.82E-08	1	9.99E-01	1.7		
60	AU040320	0.8	1.92E-05	1.3	5.33E-01	1.7	x	
61	Rpl35	1.1	8.39E-01	1.7	3.10E-03	1.7		
62	Ppp4r3a	1	7.89E-01	1.6	3.10E-03	1.6	x	
63	Ptpn1	0.8	4.63E-04	1.2	8.05E-01	1.6	x	
64	Sfi1	1	9.54E-01	1.6	1.46E-01	1.6		
65	P2ry6	0.7	2.00E-09	1.1	9.99E-01	1.6		
66	Impdh1	0.8	3.24E-03	1.3	3.79E-01	1.6		
67	St14	0.8	8.75E-02	1.3	7.52E-01	1.6	x	
68	Rps6-ps4	1	9.57E-01	1.7	7.78E-02	1.6		
69	Qsox2	0.8	2.65E-02	1.4	9.57E-02	1.6	x	
70	Gm38025	1	8.90E-01	1.6	2.43E-01	1.6		
71	Acads	0.8	1.83E-01	1.3	4.41E-01	1.6	x	
72	D130040H23Rik	0.8	2.80E-01	1.3	6.73E-01	1.6	x	
73	Gm47102	0.8	4.42E-01	1.4	4.28E-01	1.6		
74	Tlnrd1	0.8	9.73E-02	1.3	5.33E-01	1.6		
75	Creld1	0.8	1.89E-01	1.4	5.17E-01	1.6	x	
76	Sox12	0.6	9.37E-06	1	9.99E-01	1.6	x	

77	Anp32a	1	9.08E-01	1.6	5.55E-03	1.6		
78	Entpd3	0.9	6.04E-01	1.5	7.78E-02	1.6	x	
79	Lfng	0.8	7.54E-02	1.2	8.45E-01	1.6	x	
80	6030458C11Rik	0.8	8.61E-05	1.2	6.84E-01	1.6	x	
81	Klc2	0.9	5.36E-01	1.4	1.10E-01	1.6	x	
82	Igfbp7	0.7	2.19E-03	1.1	9.99E-01	1.6		
83	Ppt2	0.7	9.35E-05	1.1	9.99E-01	1.6	x	
84	Zfp111	1	8.42E-01	1.5	1.11E-01	1.6	x	
85	Phldb2	0.9	4.55E-01	1.4	3.75E-01	1.6	x	
86	4933421O10Rik	0.8	3.50E-02	1.3	8.00E-01	1.6		
87	Rtn2	1	8.96E-01	1.5	3.72E-02	1.6		
88	Slc37a1	0.7	4.49E-07	1.1	9.99E-01	1.6		
89	Dedd2	0.9	6.82E-01	1.4	2.86E-01	1.6		
90	Lactb	0.8	2.73E-02	1.2	9.99E-01	1.6	x	
91	Dgcr8	0.9	6.18E-01	1.5	3.75E-01	1.6		
92	Dgcr8	0.9	6.18E-01	1.5	3.75E-01	1.6		
93	Gm42785	0.9	6.12E-01	1.4	3.79E-01	1.6		
94	Gm28756	1	9.12E-01	1.5	3.03E-01	1.6		
95	Gm21244	0.9	7.63E-01	1.5	3.90E-01	1.6		
96	Gm15097	1	9.39E-01	1.5	2.58E-01	1.6		
97	4933411E06Rik	0.9	6.20E-01	1.4	5.02E-01	1.6		
98	Mtus1	0.8	5.22E-02	1.3	9.17E-01	1.6	x	
99	Ppp6r1	1	9.60E-01	1.6	8.38E-04	1.6		
100	Fen1	0.9	5.34E-01	1.4	7.17E-02	1.6		
101	Gm44578	1	9.84E-01	1.6	2.29E-01	1.6		
102	Gcnt1	0.8	4.36E-02	1.3	5.01E-01	1.6	x	
103	Pop5	0.9	4.48E-01	1.4	2.43E-01	1.6	x	
104	Gm5860	0.8	2.62E-01	1.3	8.28E-01	1.6		
105	Gm43498	1	9.94E-01	1.6	1.83E-01	1.6		
106	Ap1s3	0.9	7.05E-01	1.5	3.90E-01	1.6	x	
107	Cdh22	0.8	1.80E-01	1.3	6.66E-01	1.6	x	
108	Rce1	0.8	8.06E-02	1.3	6.90E-01	1.6		
109	Gm45159	0.9	8.20E-01	1.5	3.98E-01	1.6		
110	Klhl21	0.8	2.86E-01	1.3	6.93E-01	1.6	x	
111	Atg4d	0.8	4.05E-01	1.3	6.13E-01	1.6	x	
112	Fgfr1	0.7	1.21E-02	1.2	9.99E-01	1.6	x	
113	Zfp488	0.9	7.34E-01	1.5	4.28E-01	1.6	x	
114	Olfr1116	1	9.69E-01	1.5	2.58E-01	1.6		
115	Neurl3	0.7	5.11E-03	1.1	9.99E-01	1.6		
116	Ccdc162	0.9	4.99E-01	1.4	5.82E-01	1.6		

117	Rabl2	0.8	9.79E-02	1.3	8.78E-01	1.6	x	
118	Stk11ip	0.8	1.86E-03	1.3	9.00E-01	1.6		
119	Mrpl53	1	9.58E-01	1.5	1.05E-01	1.6		
120	C1qtnf1	0.7	2.36E-02	1.1	9.99E-01	1.6	x	
121	Frem2	0.7	5.02E-02	1.2	9.99E-01	1.6		
122	Ecd	0.9	1.53E-01	1.3	4.90E-01	1.6		
123	Zfp385a	0.7	9.75E-03	1.1	9.99E-01	1.6	x	
124	Vps29	1.1	5.36E-01	1.7	5.98E-04	1.6	x	
125	Ano3	0.7	7.66E-03	1.1	9.99E-01	1.6		
126	Spata6	1	8.69E-01	1.5	3.58E-01	1.6		
127	Olfr338	0.9	5.49E-01	1.4	6.03E-01	1.6		
128	Efcab14	0.9	3.01E-01	1.4	3.64E-01	1.6	x	
129	Klf16	0.9	1.52E-01	1.3	5.94E-01	1.6	x	
130	Znhit1	1	8.85E-01	1.5	3.53E-01	1.6	x	
131	Gm48443	0.9	4.04E-01	1.4	5.33E-01	1.6		
132	Churc1	1	9.86E-01	1.5	2.34E-01	1.6		
133	Cep19	0.9	5.23E-01	1.4	5.33E-01	1.6		
134	9930111J21Rik2	0.9	7.85E-01	1.4	3.75E-01	1.6	x	
135	Cnot6	0.8	1.68E-01	1.3	1.44E-01	1.6		
136	Pgpep1	0.9	6.10E-01	1.4	3.85E-01	1.5	x	
137	Npc1	0.8	5.64E-04	1.2	9.77E-01	1.5		
138	Hdac5	0.8	1.51E-01	1.2	9.42E-01	1.5		
139	Ctbp1	0.8	5.11E-03	1.3	4.30E-01	1.5	x	
140	Olfr365	1	9.54E-01	1.5	3.53E-01	1.5		
141	Cchcr1	1	9.36E-01	1.5	3.28E-01	1.5		
142	Olfr552	1	9.40E-01	1.5	3.45E-01	1.5		
143	Sp110	0.9	6.09E-01	1.3	6.48E-01	1.5		
144	Lingo1	0.9	3.12E-01	1.3	4.74E-01	1.5	x	
145	Ppp6r3	0.9	6.18E-01	1.4	2.34E-02	1.5	x	
146	Wdr25	0.9	5.51E-01	1.4	5.33E-01	1.5	x	
147	Prpf18	0.9	3.22E-01	1.4	4.43E-01	1.5		
148	2810039B14Rik	0.9	4.96E-01	1.3	7.08E-01	1.5		
149	Rab17	0.7	5.11E-03	1.1	9.99E-01	1.5	x	
150	Mob2	0.9	3.91E-01	1.4	4.86E-01	1.5	x	
151	1700082M22Rik	0.9	6.20E-01	1.3	7.17E-01	1.5		
152	Cadm1	0.9	3.85E-01	1.3	3.66E-01	1.5		
153	Triap1	0.9	6.87E-01	1.4	3.82E-01	1.5	x	
154	Olfr1328	0.9	7.15E-01	1.4	3.66E-01	1.5		
155	Gm4117	0.9	6.09E-01	1.3	6.39E-01	1.5		
156	Nmnat1	0.9	4.08E-01	1.3	8.19E-01	1.5	x	

157	Gm43577	0.9	8.64E-01	1.4	4.65E-01	1.5		
158	Gm42927	1	9.41E-01	1.5	3.28E-01	1.5		
159	Zfp697	0.9	5.97E-01	1.3	6.89E-01	1.5	x	
160	Bmf	0.6	1.80E-04	0.9	9.99E-01	1.5	x	
161	Gm45038	0.9	7.95E-01	1.4	4.01E-01	1.5		
162	C130073E24Rik	1	9.63E-01	1.5	3.60E-01	1.5		
163	E130309D02Rik	0.9	7.16E-01	1.4	3.66E-01	1.5	x	
164	Coro2a	0.7	2.98E-04	1	9.99E-01	1.5	x	
165	Mdga1	0.8	1.93E-01	1.2	9.99E-01	1.5	x	
166	Mogs	0.8	5.02E-02	1.2	8.57E-01	1.5		
167	4632415L05Rik	0.9	4.14E-01	1.3	7.13E-01	1.5	x	
168	Gm43190	1	9.60E-01	1.5	3.75E-01	1.5		
169	Prdm11	0.9	6.83E-01	1.4	6.48E-01	1.5	x	
170	Dhtkd1	1	9.86E-01	1.5	3.45E-01	1.5		
171	Vps4b	0.8	5.28E-02	1.2	9.99E-01	1.5	x	
172	Inpp4b	0.8	7.62E-02	1.2	9.99E-01	1.5		
173	Oraov1	0.9	3.69E-01	1.4	6.37E-01	1.5	x	
174	Gm43660	0.9	7.60E-01	1.4	4.19E-01	1.5		
175	Tcp11l2	0.7	1.90E-02	1.1	9.99E-01	1.5		
176	Olfr374	0.9	5.51E-01	1.3	7.86E-01	1.5		
177	Sned1	0.9	8.08E-01	1.4	3.48E-01	1.5	x	
178	Gm16892	1	9.25E-01	1.4	3.75E-01	1.5		
179	Akap2	1	9.44E-01	1.5	2.86E-01	1.5	x	
180	Kcnn2	0.7	5.25E-02	1.1	9.99E-01	1.5		

Supplementary Table 2: list of genes down-regulated upon miR-125b overexpression in MIN6

	Gene Name	FC	padj			Gene Name	FC	padj
1	Psme3	-6.3	3.60E^{-02}		43	Ostm1	-3.8	2.90E-02
2	Cnot9	-5.3	4.67E^{-02}		44	Isg20l2	-3.8	3.99E-02
3	Clcn4	-5.2	9.60E-02		45	Qsox2	-3.8	2.65E-02
4	Cep120	-5	3.10E-02		46	Naa40	-3.7	4.42E-02
5	Mcl1	-5	7.37E-02		47	Map3k11	-3.7	2.39E-02
6	Lhpp	-4.9	7.86E-02		48	Pak1	-3.7	5.31E-02
7	Map2k7	-4.8	9.27E-02		49	Twistnb	-3.7	3.60E-02
8	AU019823	-4.7	8.13E-02		50	Fhdc1	-3.7	4.42E-02
9	Fam193b	-4.7	8.06E-02		51	Bcl2l2	-3.6	2.75E-02
10	Esco1	-4.7	5.45E-02		52	Gtf2h3	-3.6	9.04E-02
11	Galnt11	-4.6	8.43E-02		53	Hspb6	-3.6	4.84E-02
12	Actl6a	-4.6	7.19E-02		54	Zdhhc1	-3.6	5.74E-02
13	Shc2	-4.6	5.46E-02		55	Tmem108	-3.6	2.87E-02
14	Sbno2	-4.6	5.79E-02		56	Sh2b3	-3.6	8.26E-02
15	Epb41l4b	-4.5	2.17E-03		57	Dynlt3	-3.5	2.35E-02
16	C2cd4c	-4.4	7.86E-02		58	BC017158	-3.5	1.00E-02
17	Glt8d1	-4.3	3.90E-02		59	Rabl2	-3.5	9.79E-02
18	Vwa9	-4.3	3.87E-03		60	Sirt6	-3.5	1.87E-02
19	Odf2	-4.3	3.18E-02		61	Myo1e	-3.5	4.21E-04
20	Nrip3	-4.3	3.52E-02		62	Smurf1	-3.5	1.25E-02
21	Trib1	-4.2	7.86E-02		63	4933421O10Rik	-3.4	3.50E-02
22	Tmem198	-4.2	6.76E-02		64	M6pr	-3.4	5.50E-03
23	Ctbp1	-4.2	5.11E-03		65	Amigo2	-3.4	6.41E-02
24	Ankrd40	-4.2	4.25E-02		66	Gpc1	-3.4	3.90E-02
25	Flad1	-4.2	7.63E-02		67	Ube3b	-3.4	2.87E-03
26	Lgals8	-4.2	1.16E-02		68	P2rx4	-3.4	4.82E-02
27	Vopp1	-4.1	5.41E-02		69	Plekho2	-3.4	9.27E-02
28	Eif1ad	-4.1	3.00E-02		70	Knop1	-3.4	1.05E-03
29	Ccdc50	-4	7.34E-02		71	Traf6	-3.3	9.79E-02
30	Mfap1a	-4	3.63E-02		72	St14	-3.3	8.75E-02
31	Blzf1	-4	3.83E-02		73	Nin	-3.3	3.31E-02
32	Nudt5	-4	2.39E-02		74	Csmd2	-3.3	3.23E-02
33	Fam81a	-3.9	8.13E-02		75	Rce1	-3.3	8.06E-02
34	Nemp2	-3.9	9.17E-02		76	Mesdc1	-3.3	9.73E-02
35	Bcam	-3.9	9.81E-02		77	Fam102a	-3.3	1.27E-02
36	Yipf6	-3.9	6.94E-02		78	Entpd4	-3.2	3.90E-02
37	Gsn	-3.9	8.21E-02		79	Tmub2	-3.2	1.04E-02
38	Mphosph6	-3.9	2.61E-02		80	Pcsk7	-3.2	5.36E-02
39	Rnpepl1	-3.9	2.91E-02		81	Cyb5r4	-3.2	8.89E-02
40	Usp12	-3.8	1.67E-02		82	Slc17a5	-3.2	3.04E-02
41	Dnajc14	-3.8	1.62E-03		83	Adam9	-3.2	2.36E-02
42	Arl2bp	-3.8	9.27E-02		84	Oat	-3.2	4.05E-02

85	Mpg	-3.2	5.11E-03		127	Ttc7b	-2.8	4.18E-02
86	Dis3l2	-3.2	1.17E-02		128	Fam35a	-2.8	8.68E-03
87	Enox1	-3.2	7.69E-02		129	Dnm1	-2.8	1.39E-02
88	Igsf1	-3.2	5.12E-03		130	Sestd1	-2.8	7.69E-03
89	Tgfb3	-3.2	6.70E-02		131	Esr1	-2.8	4.84E-02
90	A530058N18Rik	-3.2	1.21E-02		132	Cpne2	-2.8	2.95E-02
91	Slc25a15	-3.2	4.43E-02		133	Inpp4b	-2.8	7.62E-02
92	Pxn	-3.1	7.19E-02		134	Boc	-2.8	5.22E-02
93	Slc35d2	-3.1	8.13E-02		135	Mxd4	-2.8	2.43E-02
94	Smug1	-3.1	1.58E-02		136	Alg6	-2.8	2.09E-02
95	Tyk2	-3.1	1.44E-03		137	Gabbr1	-2.7	1.57E-02
96	Pomgnt1	-3.1	1.44E-03		138	Shtn1	-2.7	5.86E-03
97	Stk11ip	-3.1	1.86E-03		139	Napb	-2.7	2.96E-02
98	Nrep	-3.1	2.01E-02		140	Ipmk	-2.7	3.13E-03
99	Lpar2	-3.1	7.18E-02		141	Sp6	-2.7	7.34E-02
100	Gcnt1	-3.1	4.36E-02		142	Lrrc10b	-2.7	2.92E-02
101	Ccp110	-3.1	1.52E-02		143	Vmac	-2.7	2.56E-02
102	Vwa5b1	-3.1	2.22E-02		144	2310015A10Rik	-2.7	5.94E-02
103	Dock10	-3.1	6.02E-02		145	Mogs	-2.7	5.02E-02
104	Zbtb9	-3	4.01E-02		146	Npc1	-2.7	5.64E-04
105	Mtus1	-3	5.22E-02		147	Cpn1	-2.7	5.79E-04
106	Mfng	-3	6.76E-02		148	Gm17690	-2.7	5.18E-02
107	Rgs7bp	-3	3.80E-02		149	Nfkbiz	-2.7	9.17E-02
108	Uhmk1	-3	1.53E-03		150	Trim65	-2.6	7.59E-03
109	Chrm3	-3	9.66E-03		151	Dnaic2	-2.6	9.91E-02
110	Fam46a	-3	8.75E-02		152	Tap2	-2.6	4.60E-02
111	Chst11	-3	1.27E-02		153	Hdac9	-2.6	8.96E-02
112	Nup210	-3	7.86E-02		154	Pak6	-2.6	3.31E-02
113	Lgi2	-3	3.76E-02		155	Lfng	-2.6	7.54E-02
114	Rnf150	-2.9	2.15E-02		156	Lgalsl	-2.6	4.33E-02
115	Dazap2	-2.9	1.53E-03		157	Tmem71	-2.6	3.90E-02
116	Fam57a	-2.9	1.90E-02		158	Mapk12	-2.6	3.88E-02
117	Pip4k2a	-2.9	6.78E-03		159	Enpep	-2.6	1.65E-02
118	Parp9	-2.9	4.08E-03		160	Gc	-2.6	6.45E-02
119	Chfr	-2.9	1.44E-03		161	Ermp1	-2.6	4.32E-04
120	Lipa	-2.9	3.66E-02		162	Slc24a3	-2.6	9.66E-03
121	Rab31	-2.9	4.72E-03		163	Ctso	-2.6	8.94E-02
122	Jun	-2.9	5.91E-02		164	Prph	-2.6	6.83E-02
123	Ptprj	-2.9	7.91E-02		165	Gcnt2	-2.6	2.06E-04
124	Impdh1	-2.9	3.24E-03		166	Esyt1	-2.6	7.71E-03
125	Asah2	-2.9	3.37E-02		167	Mtfp1	-2.5	1.80E-02
126	Vps4b	-2.8	5.28E-02		168	Scarb1	-2.5	9.57E-04

169	Cadm4	-2.5	7.60E-03	211	Hsd17b11	-2.3	1.55E-03
170	Tes	-2.5	3.66E-02	212	Faim2	-2.3	6.94E-02
171	Sidt1	-2.5	5.14E-02	213	Fam83h	-2.3	1.12E-03
172	Ptpn1	-2.5	4.63E-04	214	Dnajc28	-2.3	5.13E-02
173	Mcfd2	-2.5	3.96E-02	215	Sbsn	-2.3	6.37E-02
174	Pex11g	-2.5	6.01E-02	216	Zfp385a	-2.2	9.75E-03
175	Ptp4a3	-2.5	2.24E-04	217	Aass	-2.2	5.63E-03
176	Col11a2	-2.5	7.86E-02	218	Glb1l2	-2.2	1.42E-03
177	Six1	-2.5	2.18E-02	219	Kcnn2	-2.2	5.25E-02
178	Mlycd	-2.5	2.50E-03	220	Rhbdl3	-2.2	1.27E-02
179	Vash2	-2.4	7.13E-02	221	Tmem91	-2.2	5.79E-02
180	Inhba	-2.4	8.06E-02	222	Tgfbr3l	-2.2	3.61E-02
181	Rbm41	-2.4	9.03E-02	223	Spry1	-2.2	5.79E-02
182	Bbs10	-2.4	1.12E-03	224	Ldb3	-2.2	4.04E-02
183	Brsk2	-2.4	3.60E-02	225	Krt90	-2.2	5.68E-02
184	Pitpnm3	-2.4	2.32E-02	226	Marveld1	-2.2	5.23E-02
185	Lactb	-2.4	2.73E-02	227	Robo3	-2.2	5.68E-02
186	Cfb	-2.4	2.67E-02	228	Pou6f1	-2.2	4.40E-02
187	Gpd1	-2.4	2.50E-02	229	Arnt2	-2.2	1.25E-02
188	Slc1a1	-2.4	9.60E-02	230	Zc3h12a	-2.2	4.69E-02
189	Fam131b	-2.4	8.99E-02	231	C1qtnf1	-2.2	2.36E-02
190	C130074G19Rik	-2.4	5.14E-02	232	Klhdc8b	-2.2	1.86E-02
191	Nipal2	-2.4	1.44E-03	233	Mfsd9	-2.2	3.97E-02
192	Tesc	-2.4	3.10E-02	234	Olfml3	-2.2	5.23E-02
193	Ehd3	-2.4	1.05E-03	235	Asic2	-2.2	3.70E-02
194	Slc16a6	-2.4	2.24E-04	236	Hs3st5	-2.2	1.71E-02
195	Fgfr1	-2.4	1.21E-02	237	Rasal1	-2.2	4.77E-02
196	Fam78b	-2.4	6.21E-02	238	Scn4a	-2.1	4.27E-02
197	Golm1	-2.4	1.43E-03	239	Tor2a	-2.1	1.55E-03
198	A930033H14Rik	-2.4	6.58E-02	240	B4galnt1	-2.1	1.94E-02
199	Lrrn3	-2.3	2.91E-02	241	Cnr1	-2.1	2.50E-03
200	Apmap	-2.3	4.32E-04	242	Cartpt	-2.1	1.75E-02
201	Dennd6a	-2.3	5.48E-04	243	Rab17	-2.1	5.11E-03
202	Nov	-2.3	8.13E-02	244	Opn3	-2.1	1.39E-03
203	Eif4e3	-2.3	1.30E-02	245	Pmepa1	-2.1	3.63E-02
204	Dusp6	-2.3	4.11E-02	246	Popdc3	-2.1	2.43E-02
205	Cbln2	-2.3	7.43E-02	247	Nxph1	-2.1	7.66E-03
206	Zfp217	-2.3	9.44E-03	248	Kcnq1	-2.1	3.31E-02
207	Pros1	-2.3	8.22E-03	249	Ica1l	-2	2.32E-02
208	Npepl1	-2.3	3.45E-02	250	Tmtc2	-2	1.55E-03
209	Frem2	-2.3	5.02E-02	251	Mfge8	-2	1.33E-02
210	Gm21685	-2.3	2.56E-04	252	Plppr1	-2	1.27E-02

253	Src	-2	1.08E-02
254	Rasd2	-2	1.00E-02
255	Gbp6	-2	3.10E-02
256	Tcp11l2	-2	1.90E-02
257	Nfatc1	-2	1.54E-02
258	Cyp1a1	-2	2.61E-02
259	Ptpn5	-2	1.11E-02
260	Dusp23	-2	6.32E-03
261	Phlda3	-2	2.32E-02
262	Slc40a1	-2	9.79E-04
263	Pgap3	-1.9	2.08E-04
264	Tex101	-1.9	1.50E-02
265	Anxa10	-1.9	1.27E-02
266	Cpne5	-1.9	1.58E-02
267	Clmn	-1.9	2.36E-03
268	Neurl3	-1.9	5.11E-03
269	Pstpip2	-1.9	2.19E-03
270	Gm5105	-1.9	2.88E-03
271	Arid3a	-1.9	1.61E-03
272	Slc5a1	-1.9	1.17E-02
273	Satb1	-1.9	3.43E-03
274	Mmp24	-1.9	1.59E-03
275	Tacr3	-1.8	7.53E-03
276	Slc4a11	-1.8	3.43E-03
277	9030617O03Rik	-1.8	7.71E-03
278	Ano3	-1.8	7.66E-03
279	D230017M19Rik	-1.8	3.70E-03
280	Ceacam1	-1.7	1.86E-03
281	Igfbp7	-1.7	2.19E-03
282	Coro2a	-1.6	2.98E-04
283	Them6	-1.6	9.79E-04
284	Ngfr	-1.5	3.66E-04
285	Bmf	-1.5	1.80E-04

Supplementary Table 3: DESeq2 code used to perform differential analysis on RNA-seq data

```
### R script that runs DEseq2 using miR125b OE vs C GP
## Requires two arguments:
## 1-the name of the outfile (outputName) in which the results returned by R
should be printed
## 2-the name of the directory (outfolder)in which the Salmon files are
located for the specific experiment - It will also be used as the directory
in which to put the plots produced
## 3-the p-value selected for experiment


print('Hello there! You made it, you are in R!')


##Tell it to look for arguments
args <- commandArgs(trailingOnly = TRUE)


##Define the outfile string
out <- paste(args[2],"/",args[1],sep="")


directory <- args[2]
adjpThreshold <- args[3]


library(DESeq2)
library(tximport)
library(readr)
library(tximportData)


print('Libraries were loaded succesfully')


print(file.path(directory,"sample_conditions.tsv"))
sampleInfo    <-    read.table(file.path(directory,"sample_conditions.tsv"),
header=TRUE, quote="", sep="\t")
head(sampleInfo)


print('sampleInfo generated')


sampleNames <- sampleInfo$sample_name
head(sampleNames)
```

```
print('sampleNames variables succesfully generated')


files <- file.path(directory,sampleNames,"quant.sf")
head(files)


names(files) <- sampleNames
head(files)


print('files and names variables succesfully generated')
print(file.path(directory,"tx2gene_mm10_cdna_ncrna.csv"))
tx2gene <- read.csv(file.path(directory,"tx2gene_mm10_cdna_ncrna.csv"))
print('variable tx2gene succesfully generated')


txi <- tximport(files, type="salmon", tx2gene=tx2gene)
head(txi$counts)


print('txfiles imported adequately')


rownames(sampleInfo) <- colnames(txi$counts)
head(sampleInfo)


ddsTxi <- DESeqDataSetFromTximport(txi, colData=sampleInfo, design=~experi-
ment + condition)


library(ggplot2)
library(RColorBrewer)
library(vsn)
library(pheatmap)
library (gplots)


print('libraries loaded')


dds <- DESeq(ddsTxi)


print('DESeq ran')


normCounts <- counts(dds,normalized=TRUE)
```

```
print('Printing the counts')
write.csv(as.data.frame(normCounts),file=paste(out,"_normalizedGene-
Counts.csv",sep=""))


print('Getting results')
res <- results(dds, contrast=c('condition',"miR_125","C"), alpha=adjpThresh-
old)


print('res looks like:')
print(summary(res))
head(res)


print('ordering and saving results')
resOrdered <- res[order(res$padj),]


print('resOrdered results written')


pdf(paste0(out,"_MAplot.pdf"))
plotMA(res,main='DESeq2',ylim=c(-2,2))
dev.off()


print ('MAplot written')


rld <- rlog(dds)
vsd <- varianceStabilizingTransformation(dds)
nt <- normTransform(dds)


pdf(paste0(out,"_variancePlots.pdf"))
par(mfrow=c(1,3))
print ('variancePlots  written')


notAllZero <- (rowSums(counts(dds))>0)


meanSdPlot(log2(counts(dds,normalized=TRUE)[notAllZero,] +1))


meanSdPlot(assay(rld[notAllZero,]))
```

```
meanSdPlot(assay(rld[notAllZero,]))
dev.off()
print(rld)
pdf(paste(out,'_PCA.pdf',sep=""))
 plotPCA(rld, intgroup=c('condition','experiment'))
dev.off()


select <- order(rowMeans(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:15]

df <- as.data.frame(colData(dds)[,c('condition','experiment')])

pdf(paste(out,'_heatmap_normalizedDDScounts.pdf',sep=""))
  pheatmap(assay(nt)[select,],cluster_rows=FALSE,show_rownames=FALSE,clus-
ter_cold=FALSE,annotation_col=df)
dev.off()


pdf(paste(out,'_heatmap_RLDcounts.pdf',sep=""))
  pheatmap(assay(rld)[select,],cluster_rows=FALSE,show_rownames=FALSE,clus-
ter_cold=FALSE,annotation_col=df)
dev.off()


pdf(paste(out,'_heatmap_VSDcounts.pdf',sep=""))
  pheatmap(assay(vsd)[select,],cluster_rows=FALSE,show_rownames=FALSE,clus-
ter_cold=FALSE,annotation_col=df)
dev.off()


colors <- colorRampPalette(brewer.pal(9,'Blues'))(255)


sampleDists <- dist(t(assay(rld)))
sampleDistMatrix <- as.matrix(sampleDists)

rownames(sampleDistMatrix) <- rld$condition
colnames(sampleDistMatrix) <- NULL
pdf(paste(out,'heatmap_distanceBetweenSamples.pdf',sep=""))
  pheatmap(sampleDistMatrix,clustering_distance_rows=sampleDists,cluster-
ing_distance_cols=sampleDists,col=colors)
dev.off()
```

```
res_noNA <- res[complete.cases(res),]
write.csv(as.data.frame(res_noNA),file=paste0(out,"_results_noNA.csv"))


print ("table with NA results filtered generated")


res_up <- res_noNA[res_noNA$log2FoldChange>0,]
signThreshold <- adjpThreshold


print ("res_up variable generated")


res_up_sign <- res_up[res_up$padj<signThreshold,]
res_up_sign <- res_up_sign[order(res_up_sign$padj),]
write.csv(as.data.frame(res_up_sign),file=paste0(out,"__up_signUp.csv"))


res_down <- res_noNA[res_noNA$log2FoldChange<0,]
signThreshold <- adjpThreshold


print ("res_down variable generated")


res_down_sign <- res_down[res_down$padj<signThreshold,]
res_down_sign <- res_down_sign[order(res_down_sign$padj),]
write.csv(as.data.frame(res_down_sign),file=paste0(out,"_down_signDown.csv")
)


print ("All done :)")



print ('MAplot written')


rld <- rlog(dds)
vsd <- varianceStabilizingTransformation(dds)
nt <- normTransform(dds)


pdf(paste0(out,"_variancePlots.pdf"))
par(mfrow=c(1,3))
print ('variancePlots  written')


notAllZero <- (rowSums(counts(dds))>0)
```

```
meanSdPlot(log2(counts(dds,normalized=TRUE)[notAllZero,] +1))


meanSdPlot(assay(rld[notAllZero,]))


meanSdPlot(assay(rld[notAllZero,]))
dev.off()
print(rld)
pdf(paste(out,'_PCA.pdf',sep=""))
 plotPCA(rld, intgroup=c('condition','experiment'))
dev.off()


select <- order(rowMeans(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:15]


df <- as.data.frame(colData(dds)[,c('condition','experiment')])


pdf(paste(out,'_heatmap_normalizedDDScounts.pdf',sep=""))
  pheatmap(assay(nt)[select,],cluster_rows=FALSE,show_rownames=FALSE,clus-
ter_cold=FALSE,annotation_col=df)
dev.off()


pdf(paste(out,'_heatmap_RLDcounts.pdf',sep=""))
  pheatmap(assay(rld)[select,],cluster_rows=FALSE,show_rownames=FALSE,clus-
ter_cold=FALSE,annotation_col=df)
dev.off()


pdf(paste(out,'_heatmap_VSDcounts.pdf',sep=""))
  pheatmap(assay(vsd)[select,],cluster_rows=FALSE,show_rownames=FALSE,clus-
ter_cold=FALSE,annotation_col=df)
dev.off()


colors <- colorRampPalette(brewer.pal(9,'Blues'))(255)


sampleDists <- dist(t(assay(rld)))
sampleDistMatrix <- as.matrix(sampleDists)


rownames(sampleDistMatrix) <- rld$condition
colnames(sampleDistMatrix) <- NULL
```

```
pdf(paste(out,'heatmap_distanceBetweenSamples.pdf',sep=""))
  pheatmap(sampleDistMatrix,clustering_distance_rows=sampleDists,cluster-
ing_distance_cols=sampleDists,col=colors)
dev.off()


res_noNA <- res[complete.cases(res),]
write.csv(as.data.frame(res_noNA),file=paste0(out,"_results_noNA.csv"))


print ("table with NA results filtered generated")


res_up <- res_noNA[res_noNA$log2FoldChange>0,]
signThreshold <- adjpThreshold


print ("res_up variable generated")


res_up_sign <- res_up[res_up$padj<signThreshold,]
res_up_sign <- res_up_sign[order(res_up_sign$padj),]
write.csv(as.data.frame(res_up_sign),file=paste0(out,"__up_signUp.csv"))


res_down <- res_noNA[res_noNA$log2FoldChange<0,]
signThreshold <- adjpThreshold


print ("res_down variable generated")


res_down_sign <- res_down[res_down$padj<signThreshold,]
res_down_sign <- res_down_sign[order(res_down_sign$padj),]
write.csv(as.data.frame(res_down_sign),file=paste0(out,"_down_signDown.csv")
)


print ("All done :)")
```

Appendix B: List of Publications

Published papers:

Mousavy Gharavy SN, Owen BM, Millership SJ, Chabosseau P, **Pizza G**, ...and Rutter GA. Sexually dimorphic roles for the type 2 diabetes-associated C2cd4b gene in murine glucose homeostasis. Diabetologia. 2021 Apr;64(4):850-864. doi: 10.1007/s00125-020-05350-x. Epub 2021 Jan 25. PMID: 33492421; PMCID: PMC7829492.

Published abstracts:

- Rebecca Cheung, **Grazia Pizza**, ...and Aida Martinez-Sanchez. *Regulation and Role of MiR-125b in β-cells.* EASD Islet Study Group & Beta-cell Workshop Apr 2019 (Poster)

- **Grazia Pizza***, Rebecca Cheung*, ...and Aida Martinez-Sanchez. *MiR-125b Is Regulated by Glucose via AMPK and Impairs ß-Cell Function.* Diabetes Jun 2019, 68 (Supplement 1) 2183-P; DOI: 10.2337/db19-2183-P (Poster)

- **Grazia Pizza***, Marie-Sophie Nguyen-Tu, ...and Aida Martinez-Sanchez. *Glucose regulates miR-184 via AMP-activated protein kinase (AMPK) in pancreatic β-cells.* Presented at Society for Endocrinology BES (2018), Glasgow, UK. Metabolism and Obesity abstract. (Poster). ***Awarded best Early Career poster presentation prize***.

- Aida Martinez-Sanchez, **Grazia Pizza**, ...and Guy Rutter. *Glucose regulates miR-125b expression via AMP-activated protein kinase (AMPK)*. Diabetic Med 35:49 (2018) For Diabetes UK (Poster)