The Role of PAM in Pancreatic Beta Cell Dysfunction and Diabetes

Shahana Sengupta

Department of Biochemistry

St Cross. College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity Term 2019

Memorandum

The work in this thesis is the original work of the author and any work contributed from others has been listed below. All experiments have been performed under the supervision of Professor Anna Gloyn and Dr. Benoit Hastoy at the Oxford Centre for Diabetes, Endocrinology and Metabolism (OCDEM), with input from Professor Roger Cox and Dr. Elizabeth Bentley at the MRC Harwell Institute. This project has been funded by the Wellcome Trust.

The sample preparation and data analysis for the long read (Pacific Bioscience) sequencing performed on human islets (chapter 3), was performed by Dr. Carla Burrows (OCDEM) and Dr. Anne Ndungu (Wellcome Centre for Human Genetics) respectively. The functional validation of the PAM knockout (KO) EndoC-βH1 cell line (chapter 4) was performed in collaboration with Antje Grotz (OCDEM), who generated this cell line by performing CRISPR-Cas9 editing. The characterisation of R36S-PAM (chapter 4) was carried out in collaboration with Lori Bonnycastle (NIH), Francis Collins (NIH), Hanna Huopio (University of Eastern Finland) and Markku Laakso (University of Eastern Finland). The *in silico* and functional characterisation of *PAM* coding variants (chapter 5) were identified through exome sequencing of 20,791 Type 2 diabetes cases and 24,440 controls as published (Flannick, Mercader et al. 2019).

Ultrastructure and localisation studies involving transmission electron microscopy (chapter 3 and 4) were performed with the assistance of Dr. Anne Clark.

No part of this thesis has been submitted for any other degree at this or other university.

This thesis is approximately 46,000 words.

Acknowledgments

Over the past 4 years, I have had the opportunity to meet such inspiring and awesome people. The completion of this thesis would not have been possible without the support of so many more people that mentioned here. I am so grateful to all of my friends, family and colleagues.

First, I would like to thank my supervisors Professor Anna Gloyn, Professor Roger Cox and Dr. Benoit Hastoy for the opportunity to do this DPhil, which has been a true rollercoaster ride that I am so happy I chose to undertake. I especially appreciate the support, guidance and confidence even when I had doubts in myself. I thank Prof Anna Gloyn for the opportunities to take part in international collaborations and exciting projects and for the invaluable support when I moved my base to OCDEM.

I am very grateful to Benoit for his endless guidance and support (even in 3 hour meetings), patience when I asked incessant questions and was reluctant to pH anything. To Anne, who has been a great source of knowledge, discipline and support, and someone who I look up to as a scientist. I have really enjoyed our scientific discussions and have learnt so much from both of you, including the 'proper' ways to set up EM experiments.

I would not have had such a fulfilling experience without the comradery, laughter and banter provided by my fellow DPhil students. To my OCDEM crew: Antje, Pippa, Claire, Katia and Joely for our constant ranting, western room crying, gin tasting, formals and foodie adventures and Choya consumption (for which we should all be ambassadors). To former students Mahdieh and Mahesh, who have been so supportive, kind and welcoming. To Hannah and Milan: for the support and friendship since we started together 4 years ago, it's crazy how fast it has gone.

I thank all my research colleagues at both MRC Harwell and OCDEM, for the ongoing support, kindness and laughter along the way. To Liz, Ying, Rajesh, Yixing, Sam, Tom, Nor, Rebecca, Vibe, Nicole, Emily, Fernando, Grace, Mandy, Sameena Robert, Carla, Nellie, Nik, Agata, Marta, Jason, Anubha, Mark, Sarah, Tom and Alex.

I am also very grateful to all the friends who have supported me. Those at college, especially Laura, Martin, Lewis, Vicky, Monica, Harry and Melissa; for all the hilarious adventures we've had from our first Feast together during MT 2015 to weddings, movies and all the awesome bits in between. To friends made before and during my Oxford adventure: Liv, Rob, Rosie, Ash, Ann, Charlotte, Kathleen, Lucinda, Sarah, Hannah, Jess, Hannah, Nicola, Sedona, Amelia, Nisha and Jess for all the chats, meet ups and any excuse to escape Oxford when I needed.

And to my family, without whom I could not have achieved anywhere as near as much as I have and who deserve this success as much I do. To my parents, who have always loved and supported me (and provided more food than I could ever consume) and to my sister, who is just as weird as I am and always been a great distraction from DPhil stresses. I have so thoroughly enjoyed my time in Oxford, sad to see it end, but ready for the next adventure!

The Role of PAM in Pancreatic Beta Cell Dysfunction and Diabetes

Shahana Sengupta, St. Cross College
Submitted in candidature for the degree of Doctor of Philosophy, Trinity Term 2019

Genome wide association studies (GWAS) have been critical to the identification of >400 signals at around 200 loci involved in Type 2 Diabetes (T2D) risk. However, the functional interrogation at these loci has been slow. Coding alleles in *PAM,* D563G (minor allele frequency, MAF=4.98%) and S539W (MAF=0.65%), were identified in association with increased Type 2 diabetes (T2D) risk and reduced insulinogenic index. These were recently shown to mediate risk through alterations to PAM amidating activity and expression in the β cell. Peptidylglycine α-amidating monooxygenase (PAM) was also recently shown to function in mediating insulin secretion and content. The work in this thesis has explored the role for PAM in pancreatic beta (β) cells and the cellular mechanisms by which its dysfunction alters diabetes risk across a range of diabetes phenotypes ranging from classical T2D through early onset diabetes pathogenesis.

First, I characterised the role for PAM in human pancreatic β cells using an authentic cell model, EndoC-βH1 cells. Examination of endogenous PAM expression identified the predominant isoform to be membrane-integral, which localised to insulin-containing secretory granules and was post-translationally cleaved to generate 2 further isoforms. Gene silencing and PAM catalytic inhibition studies identified that auxiliary proteins present in the secretory granule, such as chromogranin A, are amidated by PAM and that this may impact their stability and localisation. In addition, reduced intracellular insulin content mediated through impaired PAM activity was shown to be mediated through reduced *INS* gene expression.

Next, I determined the impact of a rare coding allele (p.R36S) on PAM and β cell function, identified in a proband with early onset diabetes mellitus. I compared the T2D-associated low frequency allele, S539W, which results in a loss of *PAM* function, with this novel allele to explore how these alleles might differ in their effect on β cells. Comparison of the cellular dysfunction between these two alleles identified that R36S-PAM caused β cell loss in contrast to the T2D-risk allele S539W which impaired insulin secretion.

Finally, I established a strategy to characterise rare variants identified in a gene-level signal for T2D risk in *PAM.* A comprehensive functional characterisation is needed to fully elucidate the relationship between altered *PAM* function and T2D risk. A pilot study of 5 of the 35 variants underlying the association identified a further *PAM* allele which caused reduced protein expression.

In summary, my work has uncovered the complex biology pertaining to PAM function in β cells and distinct mechanisms by which dysfunction may lead to diabetes of differing clinical severities. A more comprehensive understanding of these mechanisms may potentially inform diagnosis and drug development for diabetes along the spectrum of clinical severity.

Common abbreviations

ATP	Adenosine triphosphate
CgA	Chromogranin A
CPE	Carboxypeptidase E
CRISPR	Clustered regularly interspaced short palindromic repeats
DM	Diabetes mellitus
ER	Endoplasmic reticulum
GoF	Gain of function
GSIS	Glucose-stimulated insulin secretion
GWAS	Genome-wide association studies
IAPP	Islet amyloid polypeptide
ISG	Immature secretory granule
kDa	Kilo Dalton
KD	Knockdown
KO	Knockout
LDCV	Large dense core vesicle
LoF	Loss of function
MAF	Minor allelic frequency
MODY	Maturity onset diabetes of the young
NGS	Next generation sequencing
OR	Odds ratio
PAL	Peptidyl-α-hydroxyglycine α-amidating lyase
PAM	Peptidylglycine α-amidating monooxygenase
PC	Prohormone convertase
PHM	peptidylglycine α-hydroxylating monooxygenase
(P)NDM	(Permanent) Neonatal diabetes mellitus
PTV	Protein truncating variant
SDM	Site directed mutagenesis
SEM	Standard error of the mean
SNP	Single nucleotide polymorphism
T2D	Type 2 diabetes
TGN	*Trans*-Golgi network
TMEM	Transmembrane domain
UPR	Unfolded protein response
WES	Whole exome sequencing
WGS	Whole genome sequencing
WT	Wild type

Table of Contents

CHAPTER 1

General Introduction

1.1 TYPE 2 DIABETES (T2D)

1.1.1 The global epidemic

T2D is a complex, polygenic condition, characterised by impaired glucose-stimulated insulin secretion (GSIS) and reduced insulin sensitivity resulting in chronic hyperglycaemia (Stumvoll, Goldstein et al. 2005, Chatterjee, Khunti et al. 2017). T2D currently affects more than 425 million people globally and this is set to rise over the coming years (IDF 2017). This increased incidence has coincided with an increase in obesity. Lifestyle intervention with reported weight control has shown a reduction in the incidence of T2D (Tuomilehto, Lindstrom et al. 2001, Knowler, Barrett-Connor et al. 2002). Patients with T2D have an increased risk of developing complications, such as lower limb amputation, retinopathy, neuropathy and cardiovascular disease (Yau, Rogers et al. 2012). It is predicted that almost half of patients may have already developed a complication at the time of diagnosis for T2D (Spijkerman, Dekker et al. 2003). The cost of diabetes and its associated complications consumes ~10% of the annual NHS budget (ADA 2010, Diabetes-UK 2014, WHO 2016).

1.1.2 Lack of effective therapies for T2D

Available and effective therapies for T2D are limited, often non-specific and are not disease modifying (Del Prato, Penno et al. 2009). Since there are multiple aetiologies for T2D, using one regime of treatment does not benefit all patients in the same way. This will often lead to patients being prescribed multiple medications and lifestyle intervention to treat the chronic hyperglycaemia.

Various types of medications are available, each targeting a different system in an attempt to normalise blood glucose. Sulphonylureas promote closure of K_{ATP} channels in pancreatic β cells thereby stimulating insulin secretion (Ashcroft 1996). Metformin, a first line therapy for T2D, works to increase insulin sensitivity through the inhibition of gluconeogenesis and glycogenolysis in the liver, the former is elevated ~3x in T2D patients (Hundal, Krssak et al. 2000, Scheen and Paquot 2013). This is orchestrated through multiple mechanisms including activation of AMP-activated protein kinase (AMPK), suppression of hepatic glucose production and decreased glucose absorption in the gastrointestinal tract (Fantus and Brosseau 1986, Zhou, Myers et al. 2001, Collier, Bruce et al. 2006, Miller, Chu et al. 2013). Incretin mimetics including glucagon-like peptide-1 receptor agonists (GLP1-RA; e.g. Exenatide) and dipeptidyl peptidase-4 antagonists (DPP4, e.g. sitagliptin) work by promoting the insulinotropic effects of GLP1 and inhibiting its catabolism respectively (Iltz, Baker et al. 2006). GLP1-RAs, which are considered to be safer and provide better glycaemic control than DPP4i, function by multiple modes of action: slowing of gastric emptying, the simultaneous reduction in glucagon and increase in insulin levels (Madsbad, Kielgast et al. 2011, Brunton 2014). Thiazolidinediones (TZDs) are an alternative second-line therapy with similar efficacy to metformin and sulfonylureas (Goldberg, Kendall et al. 2005, Deeg, Buse et al. 2007). TZDs target the peroxisome proliferator-activated receptor gamma (PPARɣ) receptor, which modulates the uptake of low density lipoprotein (LDL) (Tontonoz, Nagy et al. 1998). This means that TZDs, such as pioglitazone, can provide vascular benefits (Dormandy, Charbonnel et al. 2005). Negative health implications for TZDs include elevated risk of myocardial infarction, oedema and bone fractures (Nissen and Wolski 2007, Kahn, Zinman et al. 2008).

The efficacy of oral anti-hyperglycaemic agents has been shown to decline over time (Nathan, Buse et al. 2009). In combination with the progressive loss of β cell secretory function and increasing insulin resistance in T2D, many patients ultimately advance to insulin therapy (Gastaldelli, Ferrannini et al. 2004). Insulin is an effective, but invasive form of treatment, which increases the risk of hypoglycaemia and weight gain, and still does not address the underlying cellular defect or improve long term disease outcomes.

Further understanding of the disease and its underlying pathogenesis is needed to help inform drug development. To do this, a comprehensive understanding of the naturally occurring genetic variation underlying diabetes risk in humans and the resultant functional mechanisms surrounding pathogenesis and disease is needed. In addition, performing stratification of patient groups will make use of available therapies in a more targeted manner. This will increase the efficacy of treatment and reduce the prevalence of complications (McCarthy 2017, Mayor 2019).

T2D accounts for ~90% of diagnosed cases, but there are other forms of diabetes in the spectrum (IDF 2017). Type 1 diabetes (T1D) is the most common cause of diabetes in children and polygenic in nature (IDF 2017). It is caused by an autoimmunity-associated loss of pancreatic β cells, which drives insulin deficiency and hyperglycaemia (Gepts 1965, Eisenbarth 1986). The presence of autoantibodies (e.g. insulin, glutamate decarboxylase 2 (GAD65)) serve as biomarkers for the presence of autoimmunity, and can appear before the onset of symptoms (Bonifacio, Yu et al. 2010, Ilonen, Hammais et al. 2013, Krischer, Lynch et al. 2015). Individuals with specific

human leukocyte antigen (HLA) genotypes (e.g. HLA-DR and HLA-DQ), which encode major histocompatibility complex (MHC) proteins are considered to be at an elevated risk for T1D (Rewers, Bugawan et al. 1996, Nejentsev, Sjoroos et al. 1999).

1.2 MONOGENIC DIABETES

Monogenic diabetes, such as neonatal diabetes mellitus (NDM) and maturity onset diabetes of the young (MODY), result from mutations in a single gene and usually have more severe clinical phenotypes with an earlier onset than T2D (often within childhood or young adulthood).

1.2.1 Neonatal diabetes mellitus (NDM)

Neonatal diabetes is characterised by a diagnosis before 6 months of age, although cases presenting <9 months have also been reported (Shield 2000, Edghill, Dix et al. 2006, Rubio-Cabezas, Flanagan et al. 2012). Neonatal onset diabetes diagnosed at <6 months is unlikely to be autoimmune in aetiology as shown by human leukocyte antigen (HLA) genotyping (Edghill, Dix et al. 2006). NDM is genetically heterogeneous, with many genes identified for both syndromic and non-syndromic sub-types (Table 1.2.1.1). However, 18% of NDM cases have an unknown genetic aetiology (De Franco, Flanagan et al. 2015). Disease pathogenesis can currently be segregated into three main categories: defective β cell development, impaired β cell function, and β cell destruction leading to reduced β cell mass (Rubio-Cabezas and Ellard 2013).

Autosomal dominant cases are most commonly caused by mutations in the genes encoding the K_{ATP} channel subunits, *KCNJ11* and *ABCC8*, which are thought to

contribute 40-60% of all NDM cases (Hattersley and Ashcroft 2005, Flanagan, Edghill et al. 2006, Greeley, Naylor et al. 2011). These patients can be transferred from insulin therapy to high dose oral sulphonylureas, which modulates the closure of the K_{ATP} channel thereby stimulating insulin secretion (Pearson, Flechtner et al. 2006, Rafiq, Flanagan et al. 2008). This improves glycaemic control and reduces the risk of hypoglycaemia (Rafiq, Flanagan et al. 2008, Klupa, Skupien et al. 2010). Heterozygous mutations in *INS,* encoding insulin, are the second most common aetiology of NDM (Stoy, Edghill et al. 2007). Risk is mediated via the impairment of normal insulin biosynthesis, which results in ER stress and β cell death (Stoy, Edghill et al. 2007, Colombo, Porzio et al. 2008, Boesgaard, Pruhova et al. 2010). Mutations in transcription factors (e.g. *GATA6*, *GATA4*) regulating normal β cell development *in utero* can also result in NDM, although these are usually part of clinical syndromes with multiple phenotypes in different tissues (Nicolino, Claiborn et al. 2010, Allen, Flanagan et al. 2011, De Franco, Shaw-Smith et al. 2013).

The leading cause of autosomal recessive NDM are mutations in *EIF2AK3*, encoding protein kinase R-like endoplasmic reticulum kinase (PERK), which causes Wolcott-Rallison syndrome and is the most common cause of permanent NDM (PNDM) in consanguineous families (Delepine, Nicolino et al. 2000, Rubio-Cabezas, Patch et al. 2009). Loss of PERK, a key transmembrane mediator in the unfolded protein response (UPR) of β cells, results in non-autoimmune β cell destruction by apoptosis (Harding, Zeng et al. 2001, Zhang, McGrath et al. 2002). Wolfram syndrome is another type of autosomal recessive NDM (Strom, Hortnagel et al. 1998). This is characterised by diabetes insipidus, juvenile diabetes mellitus, optic atrophy and deafness (DIDMOAD)

(Fraser and Gunn 1977). Wolfram syndrome is caused by mutations in *WFS1* encoding wolframin, which under physiological conditions protects against ER stress through inhibition of ATF6 activation (Inoue, Tanizawa et al. 1998). Dysfunction in Wolfram syndrome has also been shown to be mediated through ER stress and non-autoimmune destruction of β cells (Fonseca, Fukuma et al. 2005, Kakiuchi, Ishiwata et al. 2006, Yamada, Ishihara et al. 2006). Mutations in non-coding regulatory elements have also been identified as a cause of NDM, although not as frequently as coding mutations. Coding alleles in *PTF1A* and non-coding alleles in a distal *PTF1A* enhancer have been identified in cases of pancreatic agenesis (Sellick, Barker et al. 2004, Weedon, Cebola et al. 2014, Houghton, Swift et al. 2016). There has also been the identification of non-T1D autoimmune NDM cases (e.g. *STAT3*, *LRBA, FOXP3*), which are diagnosed before 6 months of age (d'Hennezel, Bin Dhuban et al. 2012, Flanagan, Haapaniemi et al. 2014, Johnson, De Franco et al. 2017, Johnson, De Franco et al. 2019).

Gene	Inheritance	NDM [a]	Clinical presentation [b]	Reference(s)
PLAG1/ZAC	Imprinted	Transient	Syndromic	(Temple, James et al. 1995)
ZFP57	Recessive	Transient	Syndromic	(Mackay, Callaway et al. 2008)
CNOT1	Dominant	Permanent	Syndromic	(De Franco, Watson et al. 2019)
PDX1	Recessive	Permanent	Syndromic	(Stoffers, Zinkin et al. 1997)
PTF1A	Recessive	Permanent	Syndromic	(Sellick, Barker et al. 2004)
HNF1B	Dominant	Transient	Syndromic	(Horikawa, Iwasaki et al. 1997)
GATA4	Recessive	Permanent	Syndromic	(Shaw-Smith, De Franco et al. 2014)
RFX6	Recessive	Permanent	Syndromic	(Smith, Qu et al. 2010)

(Table continued overleaf)

NKX2-2	Recessive	Permanent	Syndromic	(Flanagan, De Franco et al. 2014)
MNX1/HLXB9	Recessive	Permanent	Non-syndromic	(Bonnefond, Vaillant et al. 2013)
GATA6	Dominant	Permanent	Syndromic	(Allen, Flanagan et al. 2011)
GLIS3	Recessive	Permanent	Syndromic	(Senee, Chelala et al. 2006)
NEUROG3	Recessive	Permanent	Syndromic	(Rubio-Cabezas, Jensen et al. 2011)
NEUROD1	Recessive	Permanent	Syndromic	(Rubio-Cabezas, Minton et al. 2010)
PAX6	Recessive	Permanent	Syndromic	(Solomon, Pineda-Alvarez et al. 2009)
KCNJ11	Dominant	Both	Both	(Gloyn, Pearson et al. 2004, Flanagan, Patch et al. 2007)
ABCC8	Both	Both	Both	(Babenko, Polak et al. 2006)
INS	Recessive	Both	Non-syndromic	(Garin, Edghill et al. 2010)
GCK	Recessive	Permanent	Non-syndromic	(Njolstad, Sovik et al. 2001)
SLC2A2	Recessive	Permanent	Fanconi-Bickel syndrome	(Santer, Schneppenheim et al. 1997)
SLC19A2	Recessive	Permanent	Roger's syndrome	(Yesilkaya, Bideci et al. 2009)
INS	Dominant	Permanent	Non-syndromic	(Stoy, Edghill et al. 2007)
EIF2AK3	Recessive	Permanent	Wolcott-Rallison Syndrome	(Delepine, Nicolino et al. 2000)
IER3IP1	Recessive	Permanent	Syndromic	(Poulton, Schot et al. 2011)
FOXP3	Recessive, X-linked	Permanent	IPEX syndrome	(Chatila, Blaeser et al. 2000)
WFS1	Recessive	Permanent	Wolfram syndrome	(Inoue, Tanizawa et al. 1998)

Table 1.2.1.1 Genetic aetiology of neonatal diabetes mellitus (NDM). Genes are segregated by the predominant mechanism driving disease pathogenesis. Upper panel (white) indicates loci resulting in abnormal pancreatic development, middle panel (pale grey) are loci that impact β cell function and the lower panel (dark grey) are loci that culminate in β cell destruction. [a] indicates the duration of the NDM present, e.g. whether it is permanent or transient. [b] Clinical presentation refers to whether these loci usually

(Table continued overleaf)

present as syndromic or non-syndromic (isolated) NDM. Syndrome names have been provided if known.

1.2.2 Maturity onset diabetes of the young (MODY)

MODY constitutes a group of conditions which are defined by their autosomal dominant inheritance and comprise ~1-2% of all diabetes cases (Shields, Hicks et al. 2010, Juszczak, Pryse et al. 2016). The most common genetic causes are mutations in *GCK*, *HNF1A*, *HNF4A* and *HNF1B*, with other causes listed in Table 1.2.2.1.

Gene	Incidence (%)[a]	Reference(s)
HNF4A	5-10	(Fajans, Bell et al. 2001)
GCK	30-50	(Froguel, Zouali et al. 1993, Pearson, Velho et al. 2001)
HNF1A	30-60	(Stride, Ellard et al. 2005)
PDX1	<1	(Wright, Metzger et al. 1993, Stoffers, Ferrer et al. 1997, Fajans, Bell et al. 2010)
HNF1B	<5	(Montoli, Colussi et al. 2002, Bellanne-Chantelot, Chauveau et al. 2004, Ulinski, Lescure et al. 2006, Faguer, Decramer et al. 2011)
NEUROD1	<1	(Malecki, Jhala et al. 1999, Kristinsson, Thorolfsdottir et al. 2001)
CEL	<1	(Raeder, Johansson et al. 2006, Johansson, Torsvik et al. 2011)
PAX4	<1	(Mauvais-Jarvis, Smith et al. 2004, Plengvidhya, Kooptiwut et al. 2007)

INS	<1	(Edghill, Flanagan et al. 2008, Meur, Simon et al. 2010)
ABCC8	<1	(Bowman, Flanagan et al. 2012)
KCNJ11	<1	(Bonnefond, Philippe et al. 2012, Liu, Nagashima et al. 2013)

Table 1.2.2.1 Genetic aetiology of MODY. Genes are listed as per the order of the MODY sub-type classification. [a] Incidence of mutations in this gene as a % of total MODY cases.

MODY cases caused by mutations in *GCK*, affect 1:1000 individuals and are often asymptomatic. Cases present as stable, mild fasting hyperglycaemia from birth and do not require active treatment (Froguel, Zouali et al. 1993, Pearson, Velho et al. 2001, Chakera, Spyer et al. 2014, Steele, Shields et al. 2014). *GCK* encodes glucokinase, which catalyses the conversion of glucose to glucose-6-phosphate and is therefore a key component of glucose-stimulated insulin secretion (GSIS) (Matschinsky, Glaser et al. 1998). HNF1A-MODY cases comprise 30-60% of MODY incidence and have an early disease onset (25-30 years of age) (Colclough, Bellanne-Chantelot et al. 2013). These individuals are treated with low dose oral sulfonylureas (Yamagata, Oda et al. 1996, Pearson, Starkey et al. 2003, Hattersley, Bruining et al. 2006, Raile, Schober et al. 2015). *HNF1A* encodes a transcription factor that is involved in β cell development and regulates genes involved in glucose transport and lipid metabolism in both the liver and pancreas (Odom, Zizlsperger et al. 2004, Galan, Garcia-Herrero et al. 2011).

1.2.3 Genetics of monogenic diabetes

Since the aetiology of monogenic diabetes can be heterogeneous, identification of the cause can be critical to the efficacy of the treatment that individual receives

(Hattersley, Bruining et al. 2006). Monogenic disease is largely attributed to rare variants with large effect sizes, so it was historically easier to detect these variants in family and linkage studies and to characterise the associated cellular dysfunction as rare alleles can provide greater biological insight (Bonnefond and Froguel 2015). However, this approach is limited by the accuracy of the clinical information provided at the time of referral, since the genes prioritised for sequencing are often associated with specific clinical presentations (De Franco, Flanagan et al. 2015). This is especially true for the identification of autosomal recessive alleles in consanguineous families, which are usually performed using homozygosity mapping (El-Shanti, Lidral et al. 2000, Wakeling, Laver et al. 2019). This relies on the enrichment of pathogenic disease-causing mutations, which allows for the prioritisation of pathogenic variants in Mendelian disease (Botstein and Risch 2003, Hoskinson, Dubuc et al. 2017, Wakeling, Laver et al. 2019). A homozygous mutation in a recessive disease gene, which segregates in the child through both the maternal and paternal lines, allows for easy detection of the homozygosity block that carries the causal gene (Lander and Botstein 1987). Homozygosity maps can now be generated directly from next-generation sequencing (NGS) datasets, providing the possibility of testing all known monogenic disease genes in one assay (Ellard, Lango Allen et al. 2013, Belkadi, Pedergnana et al. 2016).

Autosomal dominant mutations were previously identified through candidate association studies but are now performed using whole exome (WES) and genome (WGS) sequencing in outbred populations (Allen, Flanagan et al. 2011, Flanagan, Haapaniemi et al. 2014, De Franco, Flanagan et al. 2017). Several NGS strategies have

been developed for the genetic testing of monogenic diabetes (Ellard, Lango Allen et al. 2013, Gao, Liu et al. 2014, De Franco, Flanagan et al. 2015). Trio-based exome sequencing strategies can be useful to identify dominant causal variants in an affected proband, and provide the opportunity to identify *de novo* causal mutations if the parents are unaffected (De Franco, Flanagan et al. 2017).

There is an overlap in the genetic aetiology of monogenic diabetes and T2D. Several loci implicated in monogenic diabetes, such as *HNF1A* and *KCNJ11*, have been shown to harbour T2D-associated variants (Gloyn, Pearson et al. 2004, Stanescu, Hughes et al. 2012, Estrada, Aukrust et al. 2014, Phani, Guddattu et al. 2014). As such, characterisation of variants identified in patients with monogenic diabetes, in known T2D risk loci, are often utilised to gain insight into the underlying biology and the capacity in which a particular protein functions.

1.3 GENETICS OF T2D

1.3.1 Genetics is a risk factor for T2D

Type 2 diabetes has a well-documented familial pattern of incidence, suggesting the involvement of genetics in the precipitation of disease (Kahn, Soeldner et al. 1969, Zimmet 1982, Meigs, Cupples et al. 2000). In addition, examination of the concordance rate for T2D amongst monozygotic and dizygotic twins, demonstrated elevated risk for monozygotic twins (70%) compared with dizygotic twins (20-30%) further exemplifying the contribution of genetics to T2D risk in cases where environmental risk is comparable (Newman, Selby et al. 1987, Kaprio, Tuomilehto et al. 1992, Poulsen, Kyvik et al. 1999). Consequently, genetics can be utilised to uncover the naturally occurring

variation, which underlies T2D risk in man, as a method of gaining insight into its heritability and inherent polygenic nature.

1.3.2 Using genetics to identify risk loci for T2D

The first T2D risk loci *PPARG, KCNJ11* and *TCF7L2*, were identified through linkage analysis, which is based in the notion that areas of the genome located in close proximity are likely to be co-inherited (Altshuler, Hirschhorn et al. 2000, Gloyn, Weedon et al. 2003, Grant, Thorleifsson et al. 2006). Fine-mapping efforts at the *TCF7L2* locus, identified an intronic variant (rs7903146) which has the strongest association to date for T2D risk (OR=1.4) (Grant, Thorleifsson et al. 2006). Linkage analysis is a method with high success in detecting rare variants with large effect sizes, but has not been as successful at detecting variants associated with complex polygenic disease traits; this is due to its limited resolution across the genome and being underpowered to detect variants with small effect sizes (Guan, Pluzhnikov et al. 2008, Prasad and Groop 2015).

Candidate association studies were extensively used to identify variants in known loci (Barroso, Luan et al. 2003). The first locus to be reproducibly associated was *PPARG,* in which the minor allele at p.Pro12Ala was protective and associated with reduced BMI and increased insulin sensitivity (Deeb, Fajas et al. 1998, Altshuler, Hirschhorn et al. 2000). The T2D-associated E23K allele in *KCNJ11* was shown to increase T2D risk (Gloyn, Hashim et al. 2001, Florez, Burtt et al. 2004). Activating mutations in *KCNJ11* are observed in patients with NDM, *KCNJ11* encodes 4 of the inwardly rectifying Kir6.2 subunits of K_{ATP} channels which are involved in the triggering pathway of insulin

secretion. The K-ATP channel is the target of sulfonylurea drugs (e.g. tolbutamide), which bind to the SUR1 subunits encoded by *ABCC8* (Ashcroft 1996, Gloyn, Pearson et al. 2004). Both *PPARG* and *KCNJ11* are loci with large effect sizes that have been successfully replicated in follow-up studies and meta-analyses.

Genome wide association studies (GWAS) have provided a means of screening the genome for common variants in T2D in an unbiased manner, utilising catalogues of common variants inherited in haplotypes. The genetic risk for T2D has been predicted to be composed of the cumulative effect of multiple common variants (minor allelic frequency, MAF>5%) as per the common disease-common variant (CDCV) hypothesis (Hemminki, Försti et al. 2008). *SLC30A8* was one of the first loci to be identified in association with T2D risk using GWAS (Sladek, Rocheleau et al. 2007). The next locus to be identified was *CDKAL1,* in which a cluster of highly associated SNPs were identified by several research groups within the same year (Saxena, Voight et al. 2007, Scott, Mohlke et al. 2007, Zeggini, Weedon et al. 2007). This GWAS era was immediately followed by studies that imputed existing or new GWAS data into meta-analyses of more than 50,000 individuals (Voight, Scott et al. 2010). This helped to identify 12 further novel loci, including those that had only previously been reported in association with monogenic disease.

GWAS signals may harbour multiple loci and as such, it can be difficult to identify the causal gene or effector transcript without performing further fine mapping of these association signals. For example, the CardioMetabochip was used to genotype a meta-

analysis of known GWAS, which led to the identification of a further 8 risk loci for T2D (Morris, Voight et al. 2012).

Trans-ethnic GWAS have enabled a better discovery of complex trait-associated loci, allowing for deeper and more comprehensive interrogation into the genetic architecture for T2D risk spanning different ancestries (Mahajan, Go et al. 2014). Heritability for T2D is predicted to range between 25-80% as estimated by family and twin-based studies (Poulsen, Kyvik et al. 1999, Meigs, Cupples et al. 2000). However, common variants identified through GWAS capture only 10% of this heritability (Stancakova and Laakso 2016). This is thought to be due to the small effect sizes that common variants carry (Manolio, Collins et al. 2009). To resolve this, many studies have focused on trying to identify lower frequency variants with larger effect sizes, in the hope of explaining some of this missing heritability (Fuchsberger, Flannick et al. 2016).

To date, more than 400 signals at 200 loci influencing T2D risk have been identified using large scale sequencing technologies, with many of the variants mediating risk through β cell dysfunction (Morris, Voight et al. 2012, Fuchsberger, Flannick et al. 2016). However, the majority of association signals map to non-coding regions of the genome, making identification of an effector or functional transcript difficult. Some studies have utilised coding variants in known T2D risk loci, allowing identification of the effector transcript and a more defined approach to uncover the functional mechanisms underlying risk (e.g. HNF1A) (Mahajan, Sim et al. 2015, Fuchsberger, Flannick et al. 2016).

Exome sequencing, which by definition focuses on screening the transcribed genome, allows for the identification of coding and splicing variants, which are naturally accompanied by a known functional transcript. This approach has been utilised by many groups to promote the identification of not only coding variation but also of lower frequency alleles that may carry larger effect sizes for T2D risk or associated traits.

1.3.3 Gene burden signals for disease risk

An alternative hypothesis for the genetic aetiology of complex traits is known as the common disease rare variant hypothesis (CDRV), which postulates that complex traits are caused by multiple rare variants that carry modest to high penetrance of their minor alleles (Reich and Lander 2001, Smith and Lusis 2002, Iyengar and Elston 2007). This variation in the penetrance of the minor allele is termed (high) allelic heterogeneity and can be exploited to explore the full functional capacity of a gene and its association with disease risk (Li and Leal 2008). Whilst next generation sequencing (NGS) has revolutionised the ability to identify low frequency variation in disease risk and complex traits, use of single-variant analysis is underpowered to identify rare variants (Slager, Huang et al. 2000, Gorlov, Gorlova et al. 2008). Approaches where multiple risk variants are collapsed into one statistic, such as gene-level association signals, are one method to circumvent this and restore power (Cohen, Pertsemlidis et al. 2006, Price, Kryukov et al. 2010).

1.3.4 Allelic series

Characterising the relationship between protein perturbation and clinical effect via the use of naturally occurring coding variation, can be a powerful method to determine the gene-dose effect on T2D risk and therefore assess whether drug development is a feasible option. One approach for this is to characterise the impact of multiple (coding) variants on protein and cellular function in an allelic series. As in the case of the *SLC30A8* locus, investigating the impact of 12 protein-truncating variants (PTVs) provided the insight that loss of function (LoF) alleles in *SLC30A8* were protective and could mediate a 60% reduction in T2D risk (Flannick, Thorleifsson et al. 2014). This relationship could not have been ascertained from characterisation of just one variant.

1.4 IDENTIFICATION OF T2D-ASSOCIATED VARIANTS IN *PAM*

1.4.1 Independently associated variants in *PAM*

Peptidylglycine α-amidating monooxygenase (PAM) catalyses carboxyamidation, which is key to the biosynthesis of over 50% of mammalian neuroendocrine peptides, and is widely expressed in neuroendocrine tissues.

Coding variants in *PAM* were identified through exome array analysis in association with insulinogenic index, an indicator of β cell function in the METSIM study, a cohort of Finnish males (Huyghe, Jackson et al. 2013). These findings were replicated in a whole genome sequencing study conducted by the deCODE consortium, which identified 23 additional variants in the *PAM-PPIP5K2* locus in the Icelandic population, one of which, S539W, was associated with T2D risk (MAF=0.65%, OR=1.42) (Steinthorsdottir, Thorleifsson et al. 2014).

The incidence of two independent risk-associated alleles in *PAM* (D563G and S539W; r^2<0.01) makes *PAM* the more likely effector gene than *PPIP5K2* to be driving T2D risk via defects in β cell function (Eipper and Mains 1988, Eipper, Milgram et al. 1993, Steinthorsdottir, Thorleifsson et al. 2014). Both genes have relevant cellular functions that could play roles in β cells: PAM is an established modifier of neuroendocrine peptides. *PPIP5K2* encodes Diphosphoinositol pentakisphosphate kinase 2, which functions to generate inositol pyrophosphates and is involved in cellular homeostasis and signalling. Variants in *PAM* have also been identified in association with anthropometric traits, such as height: the minor allele at D563G was associated with reduced stature (β =-0.027, OR=0.97) (Marouli, Graff et al. 2017).

1.4.2 Functional characterisation of independent T2D-associated variants in *PAM*

Characterisation of the T2D-associated variants, D563G and S539W, identified defects in PAM protein function in pancreatic β cells and in insulin secretion (Thomsen, Raimondo et al. 2018). Both alleles mediate T2D risk through a loss of PAM function: D563G results in reduced amidation activity and S539W mediates mislocalisation and haploinsufficiency (FIG. 1.3.2.1).

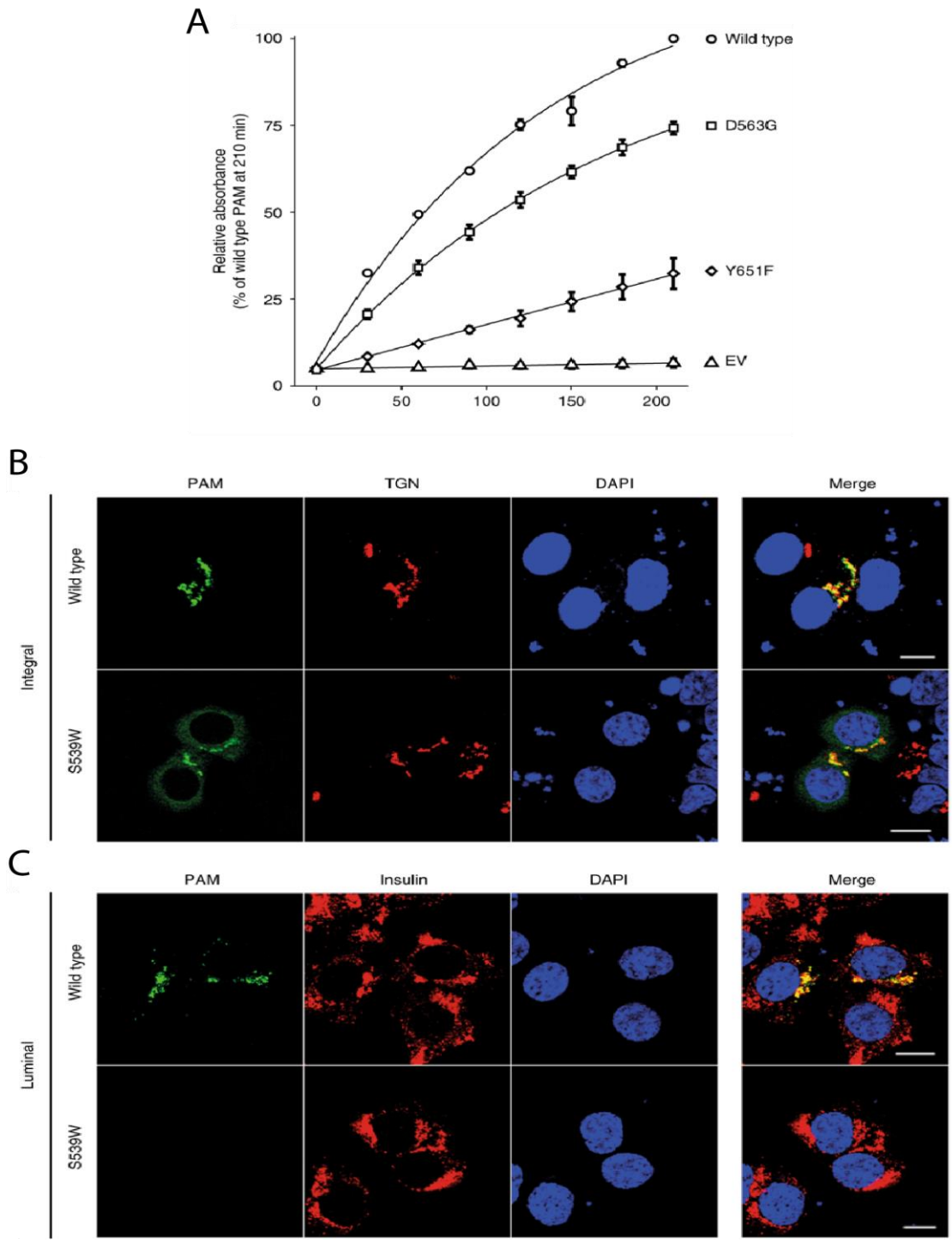

FIG. 1.3.2.1 T2D-associated variants result in loss PAM function. Figure is adapted from Thomsen et al, 2018. (**A**) Recombinant D563G-PAM exhibits reduced amidation activity as measured by an *in vitro* amidation assay. Y617F-PAM has been included as an internal negative control and all variants are normalised to WT-PAM activity at 210min. Recombinant PAM was generated from stable HEK293 cells expressing each variant. Membrane-integral (**B**) S539W-PAM is mislocalised to the ER and luminal PAM (**C**) was not detected as examined by immunofluorescence in EndoC-βH1 cells. PAM: peptidylglycine α-amidating monooxygenase; TGN: *trans*-Golgi network.

1.5 PEPTIDYLGLYCINE ALPHA-AMIDATING MONOOXYGENASE

1.5.1 Current understanding of function

Currently, PAM is the only known mammalian enzyme to catalyse the carboxy(α)-amidation of glycine-extended neuroendocrine peptides, substrates which are predicted to constitute almost half of all mammalian neuroendocrine peptides (Eipper and Mains 1988, Merkler 1994, Sforca, Oyama et al. 2004). This occurs post-translationally and maximises the biological potency and stability of the substrate (Shalev, Mor et al. 2002, Sforca, Oyama et al. 2004). Complete carboxyamidation is mediated by the sequential activity of the two catalytic domains present in the bifunctional transcripts of PAM: peptidylglycine α-hydroxylating monooxygenase (PHM) and peptidyl-α-hydroxyglycine α-amidating lyase (PAL) (Eipper, Green et al. 1992). Expression of monofunctional PAM is also possible, where isoforms may contain either the PHM or PAL domain only, either via splicing of these specific transcripts or by post-translational endoproteolytic cleavage at the dibasic motif in the linker domain (FIG. 1.5.4.1) (Ouafik, Stoffers et al. 1992, Vishwanatha, Bäck et al. 2014). In neuroendocrine cells, PAM is trafficked through the secretory pathway, to the large dense core vesicles. Mature LDCVs have an intraluminal pH (5.5), which is optimal for PAM to mediate its primary function: carboxyamidation of peptide hormones (El Meskini, Galano et al. 2001, Sobota, Ferraro et al. 2006). In the pancreatic β cells, insulin is not amidated by PAM but other factors within in the granules, proposed to be involved in insulin secretion, are (islet amyloid polypeptide, IAPP; chromogranin A, CgA).

Much of the characterisation of the processing and function of PAM has been performed in AtT20 cells, a mouse-derived cell line from a pituitary tumour (Kumar, Mains et al. 2016). As a result, there are likely species- and tissue-specific differences in how PAM is expressed, processed and trafficked between rodent models described here and the characterisation of human PAM presented later in my results.

1.5.2 Conditions for carboxyamidation (α-amidation)

Amidating activity by PAM was initially observed in 1982, when Bradbury et al observed that amidation of a synthetic peptidylglycine substrate in porcine pituitary required the precursor molecules to contain a C-terminal glycine. This suggested the possibility that this mechanism of α-amidation (hereafter referred to as carboxyamidation) required a specific enzyme and could not be performed by a simple transaminase reaction (Bradbury, Finnie et al. 1982). This was followed by the identification of two distinct enzymatic domains responsible for mediating carboxyamidation and the concept of PAM being a bifunctional enzyme: encompassing these activities (Katopodis, Ping et al. 1990, Perkins, Husten et al. 1990). Since then, PAM has been extensively studied in the anterior pituitary, with particular emphasis on elucidating its role in the biosynthesis of proopiomelanocortin (POMC) and related peptides. (Kumar, Mains et al. 2016).

Like other enzymatic reactions, carboxyamidation by PAM requires optimal conditions, which have been summarised in FIG 1.4.2.1. Key co-factors for catalysis include copper, ascorbate and molecular oxygen; factors that are shared with a similar copper-dependent monooxygenase called dopamine β-monooxygenase (DβM) that catalyses

the conversion of dopamine to norepinephrine (Stewart and Klinman 1988). Many of the studies elucidating PAM function have utilised the structural and catalytic similarity with DβM.

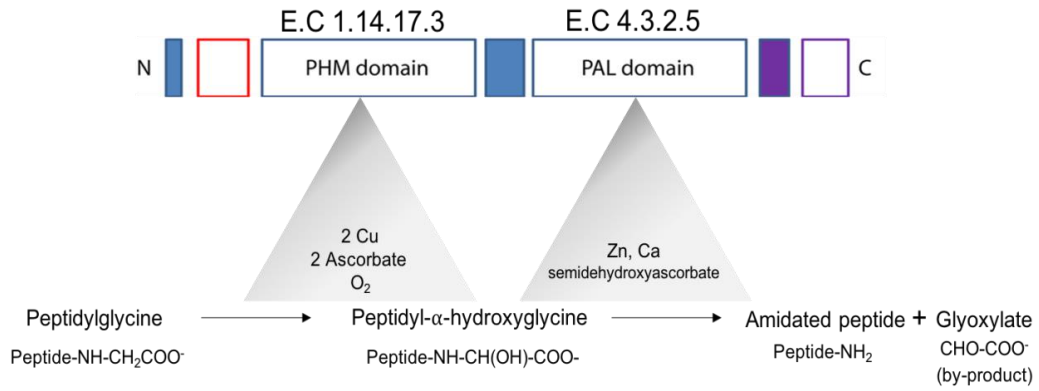

FIG. 1.4.2.1 Summary of the α-amidation reaction catalysed by PAM
PAM catalyses the α-amidation of neuroendocrine peptides that terminate in a C-terminal glycine. The reaction occurs via sequential action of its 2 catalytic domains: peptidylglycine α-hydroxylating monooxygenase (PHM) and peptidyl α-hydroxyglycine α-amidating lyase (PAL) using the enzymatic conditions summarised above. The enzyme commission (EC) numbers refer to the classification for a specific enzyme-catalysed reaction. Since PAM contains 2 catalytically distinct domains, there are 2 EC numbers associated with PAM.

1.5.3 Complex architecture of *PAM* locus

PAM has multiple layers of complexity, spanning from the multiple varying types of mRNA transcripts to tissue-specific expression of these transcripts. There are 14 protein-coding mRNA isoforms for *PAM* and 11 further processed transcripts (FIG

1.5.3.1). This multi-factorial nature of complexity leads to a larger possibility in different PAM isoforms, which each may possess different activities in the cell.

1.5.4 Post-translational processing and targeting of PAM

In addition to the complex architecture of *PAM* at mRNA transcript level, the protein undergoes extensive post-translational modification. The functional domains present in PAM are capable of influencing not only function, but also the trafficking of PAM. Endoproteolytic cleavage occurs at dibasic motifs (Lys-Arg, KR or Arg-Arg, RR) that are scattered along the PAM protein (FIG. 1.5.4.1). Cleavage is proposed to occur in secretory granules by a combination of different endoproteolytic enzymes, e.g. PC1/3.

PAM contains an N-terminal ER signal peptide, which is cleaved by signal peptidases soon after PAM enters the endoplasmic reticulum (von Heijne 1983). The proregion, which sits immediately after the ER signal peptide cleavage site and before the first catalytic domain (PHM), is important to both the secretion of the soluble, luminal isoform and also assists in the correct folding of PAM (Mains, Milgram et al. 1995). Proregions in similar proteins (e.g. subtilisin) are crucial for the secretion of the enzyme and can act in chaperone-like functions to assist with the folding process (Zhu, Ohta et al. 1989) (Shinde and Inouye 1993).

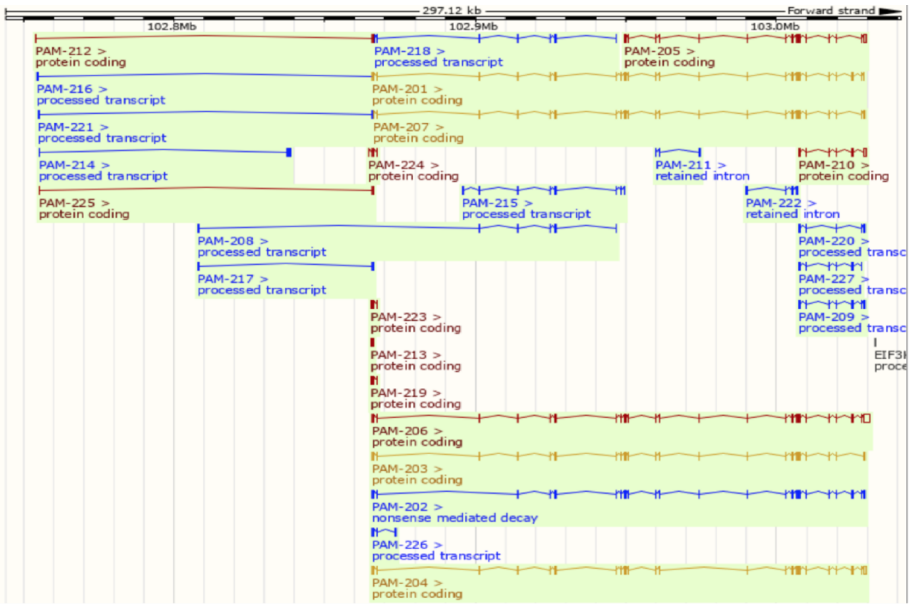

FIG. 1.5.3.1 Complex architecture of *PAM* in humans
PAM encodes multiple processed and protein-coding transcripts of varying sizes and transcriptional start sites (TSS) present. Of these, 14 are protein-coding transcripts, with the longest encoding a 974 amino acid isoform. Figure is adapted from Ensembl genome browser release 95, assembly GRCh38.p12, accessed 26th February 2019.

A histidine-rich linker region called exon A is located between the 2 catalytic domains, which is predicted to act as a pH sensor and influence the trafficking of PAM (Vishwanatha, Bäck et al. 2014). This function is mediated by a cluster of histidine residues, which due to their imidazole side chains can be easily protonated and deprotonated (Ludwig, Vanek et al. 2003). Having an intramolecular pH sensor is of particular relevance to PAM. Carboxyamidation occurs at pH5.5, following the acidification process in the maturation of secretory granules. The linker region also contains a cleavage site (contained within exon 14 in humans) allowing for the generation of monofunctional isoforms of PAM (Milgram, Johnson et al. 1992, Vishwanatha, Bäck et al. 2014). Prohormone convertases, such as those critical for the

conversion of proinsulin to insulin, and α-secretases can cleave PAM at the linker domain to generate monofunctional isoforms of PAM (Rajagopal, Stone et al. 2010, Vishwanatha, Bäck et al. 2014). This is an important first step in the solubilisation of the luminal domains of PAM or prior to cleavage at other more C-terminal sites in the protein (Rajagopal, Stone et al. 2010). In particular, cleavage within the transmembrane domain at the C-terminal side of the PAL domain, catalysed by γ-secretase, cannot occur without the luminal cleavage by either prohormone convertases or α-secretase at the linker domain (Rajagopal, Stone et al. 2010).

The transmembrane domain at the C-terminal side of PAL has been shown to anchor into the membranes of secretory granules (Milgram, Johnson et al. 1992). In addition, characterisation of the cytosolic domain, which immediately follows the transmembrane domain; suggests it is important in the trafficking of PAM through both the secretory and endocytic pathways (Milgram, Mains et al. 1996). Cleavage at the dibasic motif after the transmembrane domain can yield a soluble cytosolic fragment known as sf-CD, which can translocate to the perinuclear region in pituitary cells, following either stimulation by secretagogues or endocytosis of PAM (Rajagopal, Stone et al. 2010). The cytosolic domain also contains multiple sites for phosphorylation, which affects the endocytic trafficking of PAM (Yun, Milgram et al. 1995, Steveson, Keutmann et al. 1999). Whilst endoproteolytic cleavage of PAM occurs largely in the granules, the extent to which cleavage occurs in terms of isoforms generated will dictate the functional capacity for PAM at stages beyond cleavage.

FIG. 1.5.4.1 Non-catalytic domains are important for processing of PAM
The longest *PAM* transcript encodes a protein with multiple functional domains and five dibasic cleavage sites (navy stars), where further derivative PAM isoforms are generated through endoproteolytic cleavage in secretory granules. Key domains include the N-ter ER signal peptide (grey), the proregion (lilac), intra-catalytic linker domain Exon A (mint), C-ter transmembrane domain (cobalt) and a cytoplasmic region of multiple microdomains (pale blue). PHM: peptidylglycine α-hydroxylating monooxygenase; PAL: peptidyl-α-hydroxyglycine α-amidating lyase; TMEM: transmembrane domain.

Phosphorylation of the cytosolic domain of PAM by the kinase P-CIP2 and interactions with Rho GTP/GDP exchange factor Kalirin promote the reorganisation of the actin cytoskeleton and allow modulation of regulated secretion as well as holding a role in secretory granule maturation (Alam, Steveson et al. 2001, Ferraro, Ma et al. 2007). In addition, overexpression of the membrane-integral isoform in AtT20 cells resulted in disruption to granule maturation and reduced POMC content, suggesting that the Cter region of PAM may have roles in maintaining hormone content in stored granules (Ciccotosto, Schiller et al. 1999). This is proposed to occur through key interactors of the Cter such as Kalirin, which at high expression levels of PAM may be sequestered, thereby inhibiting their function in granule storage (Alam, Caldwell et al. 1996). Kalirin was shown to impact proinsulin and insulin secretion in β cells through signalling via the GLP-1 receptor (Dufurrena, Back et al. 2018). This indicates a non-catalytic role for PAM in trafficking and granule maturation over and above that which has been heavily described as part of neuroendocrine peptide biosynthesis.

The processing mechanisms described here were characterised in rodent pituitary tissue and likely tissue-specific. Therefore, they will differ from those present in pancreatic β cells. This should be taken into consideration when I present my data in the next chapters, which has been collected from human β (EndoC-βH1) cells.

1.6 THE PANCREATIC ISLET

1.6.1 Cellular composition

The islets of Langerhans reside in the parenchyma of the pancreas, and are composed of multiple cell types of varying functions (Da Silva Xavier 2018). Human islets are made up of 60% β cells (insulin) and 30% α cells (glucagon), with the remaining 10% comprising δ cells (somatostatin), γ cells (pancreatic polypeptide, PP) and ε- cells (ghrelin) (Brissova, Fowler et al. 2005, Ichii, Inverardi et al. 2005, Cabrera, Berman et al. 2006, Ionescu-Tirgoviste, Gagniuc et al. 2015). Whilst there are distinctions in the cellular functions of each islet cell, all are involved in the biosynthesis and secretion of their specific hormone and contribute widely to both glucose homeostasis and systemic homeostasis.

1.6.2 The secretory pathway of β cells

The secretory pathway is comprised of multiple organelles and cellular processes that ensure that products destined for regulated and/or constitutive secretion are correct and appropriately available. These processes can be segregated into 3 main stages: folding and processing of proteins at the endoplasmic reticulum (ER), granule biogenesis and maturation and exocytosis (FIG. 1.6.2) and are summarised in detail below.

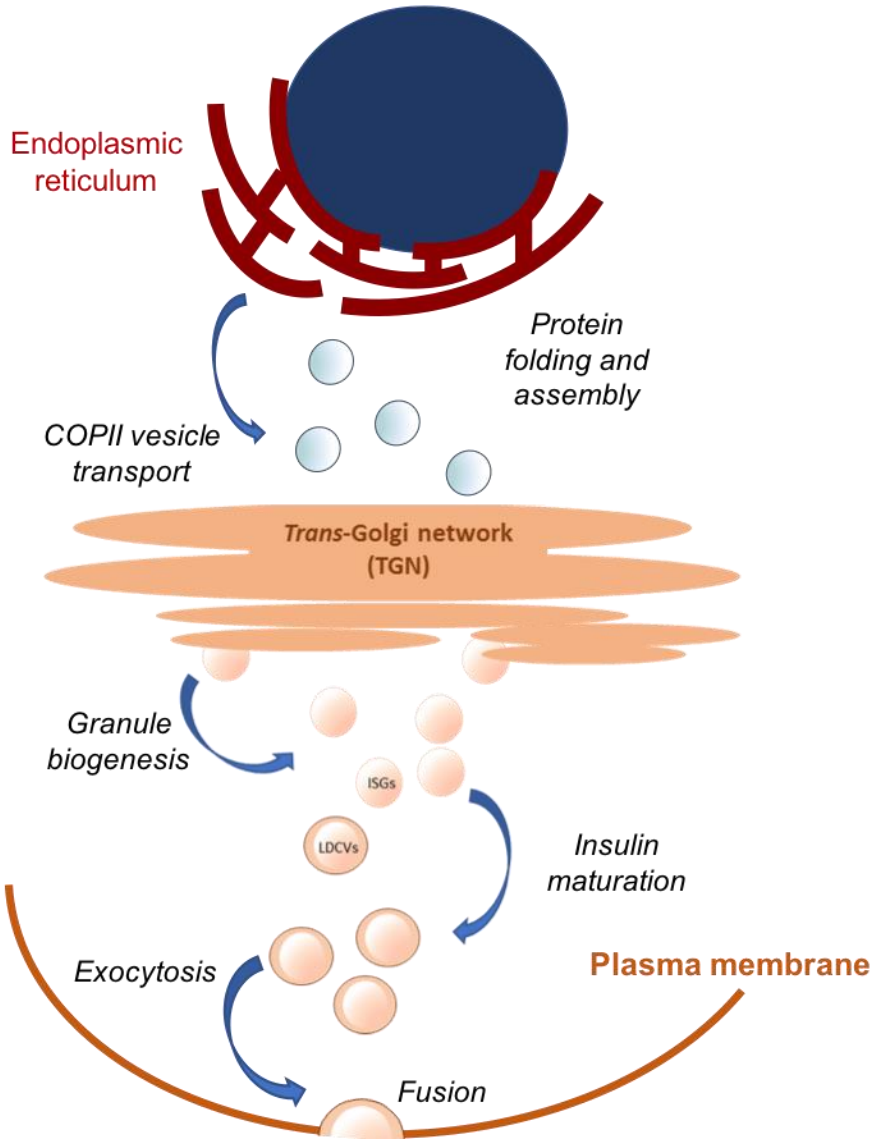

FIG. 1.6.2 The secretory pathway in pancreatic β cells. Translation of the nascent polypeptide leads to folding and assembly, which occurs in the endoplasmic reticulum (ER). From here, assembled proteins are delivered to the Golgi in COPII-vesicles, where they are then packaged into granules. Mature vesicles undergo exocytosis upon stimulation (e.g. glucose). Exocytosis is a multistep process, which culminates with the fusion of vesicular membrane with the plasma membrane and the release of insulin into the extracellular space of β cells.

1.6.2.1 Endoplasmic reticulum (ER)

Newly translated secretory proteins are targeted initially to the ER, where their folding is aided by resident chaperone proteins and post-translational modifications such as N-linked glycosylation (Kowarik, Kung et al. 2002, Jonikas, Collins et al. 2009). Insulin is expressed initially as preproinsulin, a 110 amino acid polypeptide, which is targeted to the lumen of the ER. Here, its N-terminal signal peptide is cleaved to generate proinsulin (Dodson and Steiner 1998). Disulphide isomerases link distal inter-chain (A7-B7 and A19-B20) and intra-chain (A6-A11) residues together by disulphide bonds in the ER (Blundell and Humbel 1980, Chang, Choi et al. 2003). The oxidative environment is ideal for the formation of disulfide proteins, which are performed by protein disulfide isomerases (PDI) (Jessop, Watkins et al. 2009, Fass 2010). The folding of membrane proteins and soluble proteins are handled differently by the ER, through the interaction of specific chaperones (Woolhead, McCormick et al. 2004). The processing and folding mediated in the ER is important to ensure correct packaging later on in the secretory pathway, and is governed by quality control mechanisms to ensure that this process is performed accordingly (Steiner and James 1992, Huang and Arvan 1995).

One of these mechanisms is known as the unfolded protein response (UPR) and is induced following the accumulation of misfolded proteins in the ER lumen (Onn and Ron 2010). The UPR is inherently physiological, functioning to facilitate normal proteastasis and prevent undue stress to the cell. Accumulation of misfolded proteins can result in the formation of aggregates and oxidative stress, which may lead to cellular dysfunction and stress (Haynes, Titus et al. 2004). The UPR is mediated through three transmembrane sensors, activating transcription factor 6 (ATF6), protein kinase

R-like ER kinase (PERK) and inositol-requiring enzyme 1 (IRE1), which reside in the ER membrane ready to be activated following stimuli presenting in the ER lumen and can initiate a number of downstream cellular mechanisms (FIG. 1.6.2.1.1) (Haze, Yoshida et al. 1999, Harding, Zhang et al. 2000, Tirasophon, Lee et al. 2000, Ron and Walter 2007).

During non-stress conditions, these sensors are held in inactive complexes with chaperone proteins (e.g. BiP), in a model known as 'competitive deprivation'; this prevents premature activation of the UPR (Bertolotti, Zhang et al. 2000). When a misfolded protein enters the ER lumen, these chaperones dissociate from these inactive sensor complexes and bind the misfolded proteins to stabilise them, freeing up the sensors for activation and progression to downstream effector cascades. For the most part, these adaptive mechanisms are able to resolve the blockade in normal protein synthesis, either by stabilisation of the misfolded protein or clearance by degradation (Rutkowski and Kaufman 2007, Fonseca, Burcin et al. 2009).

IRE1 is activated through oligomerisation and trans autophosphorylation by its own kinase domain (Lee, Dey et al. 2008). This results in the exposure and activation of its ribonuclease domain, which cleaves XBP1 to generate spliced XBP1 (XBP1s) (Calfon, Zeng et al. 2002). Following phosphorylation, XBP1s is translocated to the nucleus, where it upregulates expression of adaptive UPR genes (e.g. chaperones) (Lee, Iwakoshi et al. 2003, Lee, Sun et al. 2011).

PERK activation results in the phosphorylation of translation initiation factor 2 (eIF2a), which in turn results in attenuation of mRNA translation and reduces the nascent traffic entering the ER (Walter and Ron 2011).

Upon activation, ATF6 is packaged into vesicles and delivered to the Golgi where it undergoes cleavage, generating ATF6(N) via proteolytic cleavage by S1P and S2P (Haze, Yoshida et al. 1999, Ye, Rawson et al. 2000, Schindler and Schekman 2009). This cytosolic fragment ATF6(N) is then trafficked to the nucleus, where it activates transcription of UPR chaperone proteins, such as BiP (Adachi, Yamamoto et al. 2008).

Proteins that are unable to pass the quality control checkpoints are cleared from the ER via ER-associated degradation (ERAD), a process which facilitates degradation through the ubiquitin-proteasome to restore ER homeostasis (Meusser, Hirsch et al. 2005). There are essential communications that occur following UPR induction to mediate activation of ERAD (Travers, Patil et al. 2000, Hoseki, Ushioda et al. 2010). The delivery of these proteins to the proteasome required ubiquitination by E3 ubiquitin ligases, such as HRD1 (Gauss, Jarosch et al. 2006, Horn, Hanna et al. 2009, Schulz, Avci et al. 2017).

In pancreatic β cells, both IRE1 and PERK activation have been linked to proinsulin biosynthesis, particularly under stimulatory glucose conditions (Harding, Zhang et al. 2000, Lipson, Fonseca et al. 2006). For IRE1α, the predominant isoform in the pancreas, selective activation has been observed to result in enhanced protein folding capacities, without the added activation of pro-inflammatory and stress markers

(Tirasophon, Lee et al. 2000, Lipson, Fonseca et al. 2006). Following exposure to high glucose (16.7mM), IRE1α activation results in the elevated expression of protein disulfide isomerase P5, which is critical to the correct folding of proinsulin in the ER (Yoshida, Matsui et al. 2003, Lipson, Fonseca et al. 2006). PERK, conversely, is predicted to act as a negative regulator of insulin biosynthesis, as its expression has been observed to inhibit insulin biosynthesis at high glucose (Harding, Zhang et al. 2000, Lipson, Fonseca et al. 2006). Furthermore, PERK deficiency, which results in Wolcott-Rallison syndrome, results in syndromic permanent neonatal diabetes mellitus (PNDM) through reduced proliferation and impaired proinsulin trafficking (Delepine, Nicolino et al. 2000, Feng, Wei et al. 2009, Gupta, McGrath et al. 2010). These mechanisms allow for the increased expression of *INS* and secretory demand of the β cell during high glucose conditions, where the potential for ER stress due to error is high (Schuit, In't Veld et al. 1988).

In some cases, the accumulation of misfolded protein results in chronic activation of the UPR, which when unresolved can lead to ER stress (Gurzov, Ortis et al. 2009). In the Akita mouse model ($Ins2^{C96Y}$), misfolded proinsulin results in elevated CHOP and BiP, both markers of pathological ER stress (Wang, Takeuchi et al. 1999, Oyadomari, Koizumi et al. 2002). This mediates a dominant negative diabetes pathogenesis, where aggregation complexes consisting of both mutant and wild-type insulin form and are degraded leading to insufficient insulin secretion (Yoshioka, Kayo et al. 1997, Izumi, Yokota-Hashimoto et al. 2003). Prolonged activation of the UPR often results in cellular chaos, as the insult cannot be resolved through homeostatic UPR signalling or ERAD, which can result in a switch to a pro-apoptotic cellular fate. Hyperactivation of IRE1

drives proapoptotic signalling through ASK1, which results in phospho-activation of the JNK pathway and leads to apoptosis (Urano, Wang et al. 2000, Nishitoh, Matsuzawa et al. 2002). Chronic activation of PERK can result in CHOP, which can also lead to apoptosis (Zinszner, Kuroda et al. 1998, Marciniak, Yun et al. 2004). The precise mechanisms surrounding this switch of cellular fate are unclear; the involvement of the JNK pathway is a clear contributor to this cell fate (Szegezdi, Logue et al. 2006, Brown, Strudwick et al. 2016). β cell death resulting from ER stress is commonly associated with pathogenesis of Type 1 diabetes (T1D) but is also a central mechanism implicated in the pathogenesis of both neonatal diabetes and T2D. Cell death is usually mediated through apoptosis, which requires a cell-autonomous decision and activation of specific effectors (FIG 1.6.2.1.2).

Increased β cell death can result in a loss of β cell mass, which can be a central hallmark of diabetes (Chen, Cohrs et al. 2017). There is some controversy over the extent to which loss of β cell mass contributes to T2D pathogenesis (Butler, Janson et al. 2003). Many believe that defects in the secretory capacity of β cells are a more likely cause and predict that reduced β cell mass occurs secondarily to this (Sakuraba, Mizukami et al. 2002, Yoon, Ko et al. 2003). Both of these are potential outcomes that can result from varying attenuations in the same central pathways involved in β cell function (Donath, Ehses et al. 2005). Minor adjustments in how these pathways function can have a diverse impact on proliferation, cell survival and secretory pathways in the cell.

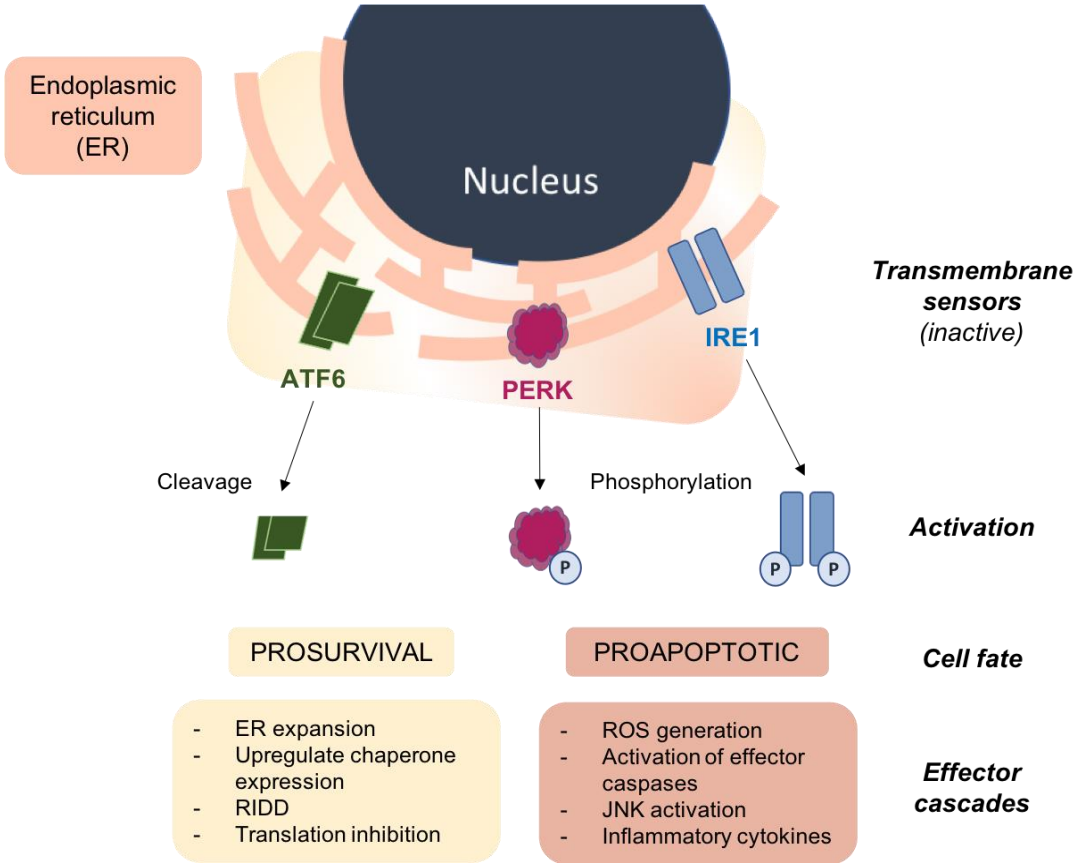

FIG. 1.6.2.1.1 Activation of the unfolded protein response (UPR) can result in varying cellular outcomes When a misfolded protein presents in the ER lumen, a cascade of physiological mechanisms are activated to resolve the associated dysfunction. These are mediated through 3 transmembrane sensors, which are activated in direct response to the misfolded protein in the ER lumen. Downstream signalling cascades can either be physiological (prosurvival) or pathological (proapoptotic) and each comprise multiple signalling cascades across the cell. IRE1: inositol-requiring enzyme 1; PERK: PRKR-like endoplasmic reticulum kinase; ATF6: activating transcription factor 6; RIDD: regulated IRE1-dependent decay (of nascent mRNA entering ER); JNK: c-Jun N-terminal kinase; ROS: reactive oxygen species.

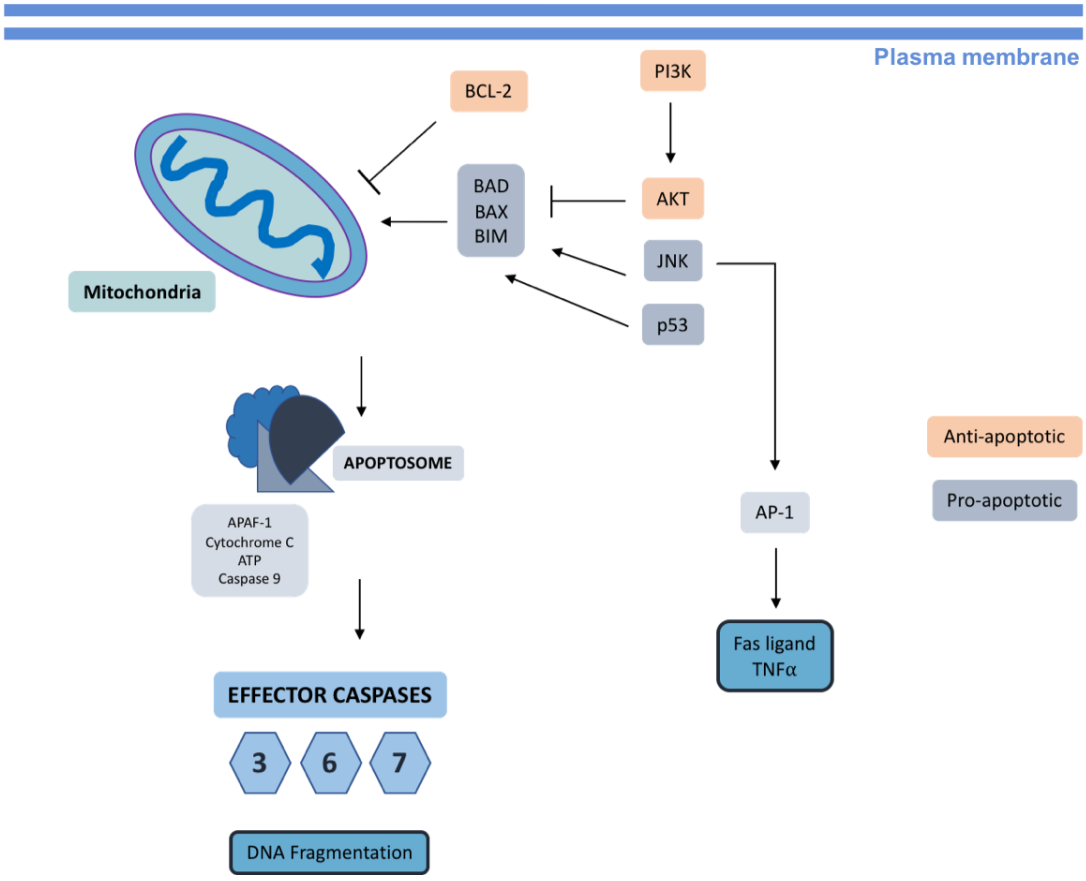

FIG. 1.6.2.1.2 The intrinsic pathway of apoptosis (programmed cell death)

Triggering signals within the cell, such as unresolvable ER stress, may activate pro-apoptotic pathways directing the cell fate toward death. Multiple pathways can lead to this decision, but the terminal cascade in initiating apoptosis is mediated through mitochondria, with effector caspases or inflammatory cytokines responsible for the final actions prior to cell death. APAF-1: apoptotic protease activating factor 1; AP-1: activator protein 1; TNFα: tumour necrosis factor α; PI3K: phosphoinositide 3-kinase; AKT: protein kinase B; JNK: c-Jun N-terminal kinase; BCL-2: B cell lymphoma 2.

1.6.2.2 Golgi Apparatus: secretory granule biogenesis

Correctly folded secretory proteins that have passed quality control checkpoints, are packaged into COPII-vesicles at ER exit sites and delivered to the Golgi apparatus via anterograde transport (Lee, Miller et al. 2004, Stagg, Gurkan et al. 2006, Bhattacharya, J et al. 2012). Here, they are trafficked through the cisternae to the *trans*-Golgi network (TGN), where biogenesis of secretory granules and budding occurs. Proinsulin is packaged into immature secretory granules (ISGs) along with components needed to maintain granule structure (e.g. chromogranin A, IAPP) and hormone modification enzymes (e.g. PAM, PC1/3, PC2, CPE) (Farquhar and Palade 1981). The secretory granule is a complex scaffold, which facilitates highly dynamic interactions between granule components, cofactors and enzymes and is sensitive to the microenvironment within the granular lumen.

Once packaged, immature granules containing proinsulin bud from the TGN surface and undergo maturation to generate large dense core vesicles (LDCVs) (FIG. 1.6.2.2.1). This is mediated by a V-type H^+-ATPase pump, which utilises ATP hydrolysis to actively pump protons into the granule lumen, utilising Cl^- as counterions in this process (Pace and Sachs 1982, Barg, Huang et al. 2001). Acidification (pH7 to pH5.5) is crucial for the endoproteolytic processing of prohormones and auxiliary proteins within the granular lumen, such as cleavage of PAM itself leading to the generation of luminal secreted PAM isoforms (Hutton 1982, Arvan and Castle 1998, Rajagopal, Stone et al. 2010). The conditions within LDCVs are optimal for carboxyamidation, which promotes the processing of auxiliary proteins such as islet amyloid polypeptide (IAPP) (Eipper and Mains 1988).

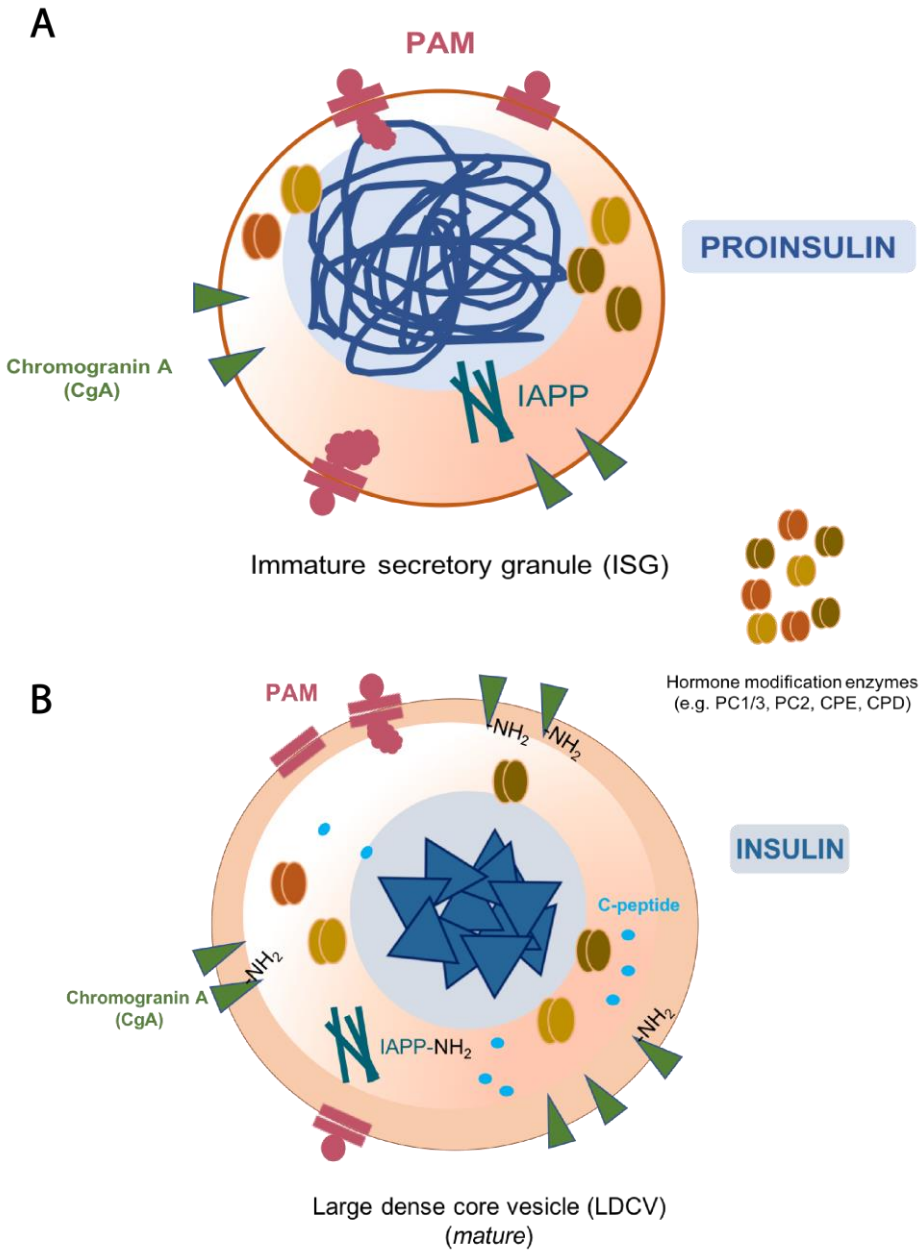

FIG. 1.6.2.2.1 Maturation of secretory granules in pancreatic β cells. Following the packaging of proinsulin into nascent granules at the TGN, immature secretory granules (ISGs) (**A**) undergo maturation and storage as large dense core vesicles (LDCVs) (**B**) prior to exocytosis following stimulus-coupling secretion. Whilst the internal composition remains similar, LDCVs can be distinguished from ISGs by the presence of a dense core containing mature insulin (blue triangles) and a halo at the periphery containing cleaved C-peptide fragments; as well as the acidic pH within the lumen following granule maturation. PAM is catalytically active in LDCVs, where carboxyamidation of granule components, e.g. IAPP occurs. These interactions are needed to ensure secretion of insulin occurs properly. PAM: peptidylglycine α-amidating monooxygenase; CgA: chromogranin A; IAPP: islet amyloid polypeptide.

Proinsulin is modified to insulin through successive cleavage by the calcium–dependent endoproteases PC2 and PC1/3, resulting in the release of C-peptide (FIG. 1.6.2.2.2) (Orci, Ravazzola et al. 1985, Davidson, Rhodes et al. 1988). Pulse-chase labelling indicates that only mature insulin and not proinsulin can progress to mature LDCVs, but that both species are sensitive to stimulus-coupled secretion (Orci, Ravazzola et al. 1985). An aggregation process, in which insulin hexamers form via the coordination of Zn^{2+} by 3 histidine residues present in the central axis of insulin, forcing insulin to become crystalline and form the so-called dense core of LDCVs (Hill, Dauter et al. 1991, Dodson and Steiner 1998). Upon secretion into the extracellular space, these hexamers dissociate to form monomers and are circulated via the bloodstream (Carpenter and Wilcox 2014).

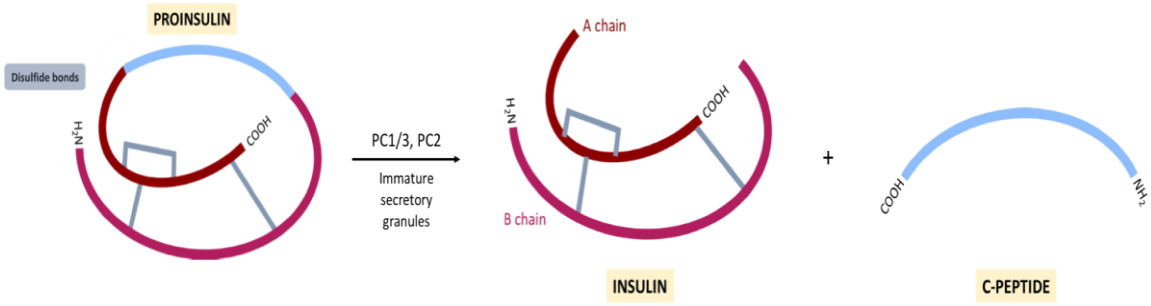

FIG. 1.6.2.2.2 Endoproteolytic conversion of proinsulin to insulin releases the internal C-peptide fragment. Following packaging at the TGN, proinsulin in ISGs undergoes maturation to yield mature insulin, which condenses into hexameric complexes in the dense proteinaceous core of LDCVs and C-peptide which localises to the halo. This conversion is catalysed by the sequential action of PC2 and PC1/3, which cleave at dibasic motifs on either side of the C-peptide. PC: prohormone convertase; C-peptide: connecting peptide.

1.5.3 Glucose-stimulated insulin secretion (GSIS)

The triggering pathway is one of two main components that make up the canonical pathway for glucose-stimulated insulin secretion (GSIS) (FIG. 1.5.3.1). In humans, entry of glucose is proposed to be mediated by the GLUT1 and GLUT2 transporters, which facilitate rapid entry of glucose into the cell owing to their low affinity and high capacity nature (Newgard and McGarry 1995, Henquin 2000, Henquin 2004, McCulloch, van de Bunt et al. 2011, van de Bunt and Gloyn 2012). Glucose is subsequently phosphorylated by glucokinase (GCK), and processed through oxidative glycolysis to generate pyruvate, which then proceeds into a multistep cascade to generate ATP (Iynedjian 1993, Matschinsky 1996). A subsequent elevation in the ATP: ADP ratio mediates the closure of the K_{ATP} channels via the Kir6.2 subunit (Aguilar-Bryan, Clement et al. 1998, Ashcroft and Gribble 1998, Seino 1999). This, in turn, results in the depolarisation of the plasma membrane and the opening of voltage-dependent calcium channels (VDCCs), causing a net influx of Ca^{2+} into the cell down the electrochemical gradient (Dean and Matthews 1970, Ashcroft and Rorsman 1989, Satin LS 1994). This mediates a rise in cytoplasmic free Ca^{2+}, which triggers the activation of exocytotic machinery and secretion of insulin-containing secretory granules into the extracellular space (Henquin 2000).

Activating mutations in either *KCNJ11* (encoding the Kir6.2 subunit) or *ABCC8* (encoding the SUR1 subunit) are causal for monogenic diabetes and result in the K_{ATP} channels being held in a permanently open conformation, thereby preventing membrane depolarisation (Gloyn, Pearson et al. 2004). Variants in *GCK*, which have

been identified in association with MODY, PNDM and gestational diabetes, mediate risk through impairment of GCK's function as a glucose sensor in the β cell (Vionnet, Stoffel et al. 1992, Matschinsky, Glaser et al. 1998, Shaat, Karlsson et al. 2006, Osbak, Colclough et al. 2009). The K_{ATP} channels of the β cell can also be modulated pharmacologically: sulphonylureas, such as tolbutamide are often used in research to mediate closure of the K_{ATP} channels and thus stimulate insulin secretion (Ashcroft 1996).

The amplifying pathway, which works to enhance insulin secretion, works mostly through elevated Ca^{2+} and downstream of glucose metabolism (Gembal, Gilon et al. 1992, Sato, Aizawa et al. 1992). Many of these mechanisms involve secondary messengers, such as cyclic AMP (cAMP), which is generated through the hydrolysis of ATP by adenylate cyclase. The effects of cAMP are mediated through protein kinase A (PKA) (Sharp 1979, Jones and Persaud 1998). Forskolin, which requires high glucose (16.7mM) to promote insulin secretion, works via the activation of adenylate cyclase, a subsequent rise in cAMP and Ca^{2+} influx (Seamon and Daly 1981, Wiedenkeller and Sharp 1983). This mediates cAMP-dependent activation of epac2 and PKA followed by Ca^{2+}-dependent exocytosis (Hermansen 1985, Henquin and Nenquin 2014).

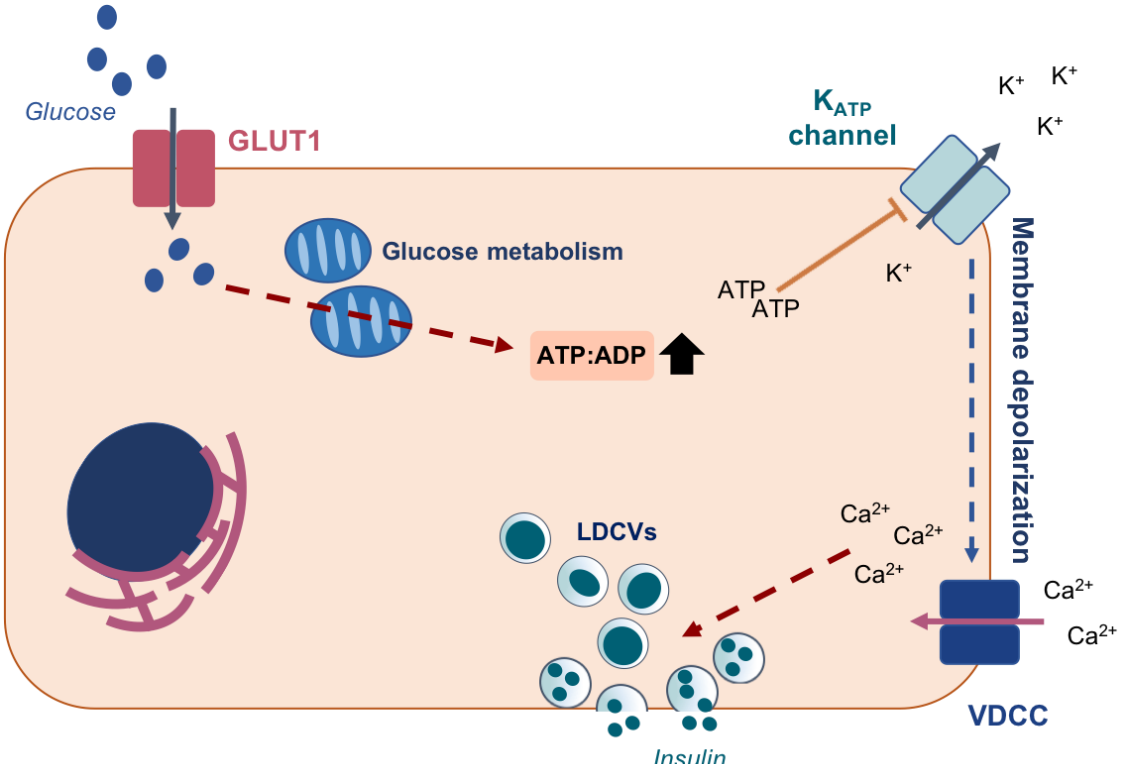

FIG. 1.5.3.1 Canonical pathway for glucose stimulated insulin secretion (GSIS). The triggering pathway begins with rapid entry of glucose through the GLUT1 transporter, which undergoes metabolism to generate ATP. This increase in ATP:ADP ratio drives closure of K_{ATP} channels, which results in membrane depolarisation and opening of voltage dependent Ca^{2+} channels (VDCC). The influx of Ca^{2+} drives the exocytosis of docked large dense core vesicles (LDCVs) and the release of insulin into the extracellular space.

1.7 AIMS OF THESIS

Throughout my introduction, I have described the main processes by which glucose homeostasis is governed, and the mechanisms by which dysfunction in the β cell can lead to impaired glycaemia and diabetes. Advances in human genetic research have informed our understanding of naturally occurring variants in the prevalence of T2D and its risk factors. This has been largely driven by the advent of large scale sequencing technologies and the identification coding variants in risk loci to distinguish effector

transcripts and allow for plausible links to be made between identified variants and underlying pathophysiological mechanisms.

My thesis has focused on the T2D susceptibility gene *PAM;* both in the context of its function in the pancreatic β cell and in the characterisation of coding variants identified either in a patient with neonatal diabetes or in association with T2D risk.

- In Chapter 3, I aim to characterise the main physiological roles for PAM in the pancreatic β cell, and in particular insulin secretion in a human clonal β cell line. Understanding the β cell-specific role for PAM will inform on the role of PAM in physiology and how disruption may proceed to cause disease.
- In chapter 4, I functionally characterise a novel, *de novo*, nonsynonymous variant R36S-PAM, identified in a patient with (monogenic) early onset diabetes mellitus. Exploration of the impact on PAM protein function and β cell function will help to identify the causal mechanisms underlying clinical disease in this proband.
- In chapter 5, I decipher the impact of rare coding variants that contribute to a gene-level signal for T2D risk in *PAM*. Characterisation of the impact of these variants on PAM function will explore the relationship between PAM protein perturbation and clinical effect and may identify critical regions of PAM required for β cell function

CHAPTER 2

Materials and Methods

2.1 GENERATION OF PAM EXPRESSION CONSTRUCTS

2.1.1 Procurement of the PAM vector constructs

Several constructs containing membrane-integral and luminal isoforms of PAM cDNA were provided by Dr. Anne Raimondo (OCDEM), all constructs were in either pCMV6-Entry or pIRES-puro2 expression plasmids. These PAM-containing constructs were generated by the following methods. A pCMV6-Entry expression plasmid containing human *PAM* was purchased from Origene (Maryland, USA), encompassing the membrane-integral isoform c of *PAM* (NM_138821.2), which lacks the intradomain linker region (exon A) between the two catalytic domains, encoded by exon 14. The untethered isoform encompassing 1-710 amino acids was generated by PCR by Dr. A. Raimondo (OCDEM) and sub-cloning of this fragment into a pCMV6-entry vector.

The pCMV6-Entry plasmids contain the kanamycin resistance gene (Kan^R) and pIRES-puro2 plasmids contain ampicillin resistance (Amp^R) for the purpose of bacterial selection. Bacterial stock plates for each plasmid vector were made as per section 2.2. A pCMV6-Entry empty vector was cloned for use as an negative control in experiments alongside the *PAM* vectors.

2.1.2 Agarose DNA Gel electrophoresis

To visualise plasmid samples for measure of plasmid integrity or check whether site directed mutagenesis PCR (section 2.4) had been successful, samples were run on an 0.8% agarose gel. The agarose gel was prepared by dissolving agarose (Sigma Aldrich) in Tris-acetate- ethylenediaminetetraacetic acid (EDTA) (TAE) buffer (Sigma Aldrich). Heat from a microwave was used to dissolve the agarose. The solution was then

cooled under a tap to ~50°C prior to the addition of ethidium bromide (Sigma Aldrich) at 0.002% of the final solution. Once the gel had set, the gel running tank was filled with 1X TAE buffer. DNA samples were mixed with 6X DNA gel loading dye (New England BioLabs) and loaded into the wells of the gel alongside a 1 kilobase (kb) plus ladder (New England BioLabs). Gels were run at 120V for 1 hour and viewed on a Gel Doc 2000 transilluminator system using the Image Lab 5.0 software (Bio-Rad).

2.2 TRANSFORMATION OF COMPETENT ESCHERICHIA COLI (E. coli)

Plasmid DNA was introduced into competent DH5α *E. coli* cells by transformation. Two to five microlitres of ligation reaction or up to 100 ng of plasmid DNA was added to 50 µl of DH5α competent cells (Life Technologies) in 1.5 ml tubes on ice. Reactions were mixed by gentle pipetting followed by incubation on ice for 30 min. Following incubation, reactions were subjected to heat shock at 42°C for 45s and then a further 2 min incubation on ice. Following this, 350ul pre-warmed super optical broth (SOC) media (Invitrogen) was added to each 1.5ml tube and mixed briefly by pipetting. Cultures were incubated with the SOC media for 60 mins at 37C in an orbital shaking incubator set to at 225 rpm (Stuart Orbital Incubator S150, Jencons Scientific Ltd).

For bacteria transformed with plasmid DNA, 50µl of the reaction was spread evenly across LB agar plates with the relevant antibiotic – 25µg/ml of kanamycin (Sigma Aldrich) or 50µg/ml of ampicillin (Sigma Aldrich). For bacteria transformed with a ligation reaction, the reaction was spun down in a microcentrifuge, and the pellet was resuspended in 100 µl of SOC and plated. Plates were incubated for 18 hours at 37°C.

Only cells that have taken up the ligated vector such as with Kan^R will grow on LB/Kan plates.

2.3 DNA EXTRACTION OF TRANSFORMED COMPETENT CELLS

2.3.1 DNA extraction by miniprep

Bacterial colonies were picked from LB agar plate and used to inoculate 10ml LB broth (Sigma Aldrich) supplemented with either 100 μg/ml ampicillin or 25 μg/ml kanamycin. Cultures were incubated overnight at 37°C for 16-18hrs in an orbital shaking incubator at 250 rpm (Stuart Orbital Incubator S150, Jencons Scientific Ltd). After incubation, glycerol stocks were initially created per culture, by taking 500 μl bacterial culture and adding an equal volume of 50% glycerol (in water). These glycerol stocks were then snap frozen in dry ice and stored at -80°C, and used in future for further plasmid preparation. The remaining culture was pelleted at 3900g for 15 mins (Heraeus Multifuge 3SR, Thermo Scientific). The supernatant was discarded and the DNA from the pellet extracted using the PureYield Plasmid Miniprep system (Promega), according to manufacturers' instructions. The DNA was eluted in 30μl of RNase-free water and stored at -20°C.

2.3.2 DNA extraction by midiprep

The glycerol stock for the plasmid DNA to be extracted was scraped with a plastic pipette tip and used to inoculate 100 ml LB medium with antibiotic (Sigma Aldrich) in 500 ml flasks and placed in the shaking incubator 250rpm (Stuart Orbital Incubator S150, Jencons Scientific Ltd) for 16 hours overnight at 37°C. After incubation, the culture was pelleted at 3,500 rpm for 10 min (Heraeus Multifuge 3SR, Thermo

Scientific). DNA from the pellet was extracted using the PureYield Plasmid Midiprep system (Promega), according to manufacturers' instructions. The DNA was eluted in at least 400 µl of RNase-free water and stored at -20°C.

2.3.3 DNA quantification

DNA was quantified using the NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), which measures the concentration of DNA based on sample absorbance at a wavelength of 260nm. Sample purity was assessed by measuring the 260nm/280nm absorbance ratio which provided an indication of any contamination by protein or organic compounds. DNA deemed as having good purity had a 260/280 ratio of ~1.8-1.9. Only DNA midiprep samples with concentrations above 100ng/µl and that were of ideal purity were subsequently used for transfections.

2.3.1 Sanger Sequencing

Samples were sent together with specific sequencing primers indicated in the relevant chapters to either Source Bioscience (Nottingham, UK) or Eurofins Genomics (Germany). For plasmids mutated by SDM, sequencing was performed to ensure that only the desired nucleotide changes were introduced.

2.4 GENERATION OF VARIANT CONSTRUCTS

2.4.1 Site directed mutagenesis

Site directed mutagenesis (SDM) was performed on the pCMV6-Entry-h*PAM*v3 (membrane integral) and pCMV6-Entry-h*PAM*v3 (luminal) constructs in parallel (see below for variant-specific details). Primers were designed using either the Quikchange

Primer Design (Agilent Technologies) or Primer X software. Both softwares are set up to design primers with a minimum GC content of 40% and that terminate in one or more C or G bases. Specific primers used for the generation of each variant discussed have been described in the relevant chapters 4 and 5 and were all purchased from Eurofins Genomics (Germany). The R36S variant was introduced into the pCMV6-Entry-h*PAM*v3 constructs using the Stratagene QuikChange II Kit (Agilent Technologies, California, USA) in collaboration with Dr. Anne Raimondo (Gloyn group, OCDEM) and confirmed by Sanger sequencing as per section 2.3.1 (Source BioScience, Nottingham, UK and Eurofins, Luxembourg). For R36S, both the membrane-integral and luminal PAM-containing constructs were modified. Variants shortlisted from the 52K Exome Sequencing dataset were introduced into pCMV6-Entry-hPAMv3 constructs by site directed mutagenesis in collaboration with Dr. Fernando Abaitua (Gloyn group, WTCHG) using Pfu Turbo and Dpn1 reagents from Promega and then confirmed by Sanger sequencing (Eurofins Genomics, Germany) as per section 2.3.1. For variants shortlisted from the 52K exome dataset, only the membrane-integral PAM-containing construct was modified.

SDM reactions were performed using a protocol adapted from the Quikchange II Site-Directed Mutagenesis (Agilent Technologies) kit manufacturer's guidelines and reagents purchased from Promega. The composition of an SDM reaction has been summarised in Table 2.4.1.1 and the programme used to run the reaction on in Table 2.4.1.2. Each reaction was set up with 50ng template DNA (e.g. pCMV6-h*PAM*v3 construct) and reactions were performed on a tetrad thermocycler (DNA Engine, MJ Research).

Component	Volume (μl)
Template DNA (20ng/μl)	7.2
Reaction buffer (10x)	5
Primer F (10μM)	1.36
Primer R (10μM)	1.37
dNTPs (10mM)	1
dH_2O	33.07
Pfu Turbo	1*

Table 2.4.1.1 Composition of reaction for one SDM reaction (50μl). Primers were purchased from Eurofins Genomics. All other reagents were purchased from Promega. *Reagent should be added last.

Stage	Cycles	Temperature (°C)	Time (min)
Denaturation	1	95	1
Primer annealing	18	95	0.5
		55	1
Extension		68	8[a]

Table 2.4.1.2 Programme for SDM PCR on thermocycler. [a] extension time is optimised for pCMV6-Entry-h*PAM*v3 constructs (~7kb).

Stage	Temperature (°C)	Time (min)
Dpn1 digestion	37	60
Cooling	4	10

Table 2.4.1.3 Programme for *Dpn1* digestion on thermocycler. Prior to the addition of *Dpn1*, 10μl of SDM PCR reaction was set aside to run as a control in the gel. *Dpn1* was purchased from Promega.

Following the SDM PCR reaction, tubes were placed on ice for 5min and then subsequently run for a Dpn1 digestion (1μl Dpn1 per 50μl SDM reaction). Template DNA which has been produced in bacteria is methylated and therefore a target for Dpn1. Reaction details are summarised in Table 2.4.1.3. Following Dpn1 cleavage, aliquots of samples were stored at 4°C overnight, frozen at -20°C or run immediately on a gel as per section 2.1.2. Only <10μl of reaction mix was run in a gel to see whether there was laddered bands to indicate successful Dpn1 cleavage and presence of a residual band to indicate successful SDM PCR amplification.

Following visualisation of bands by agarose gel electrophoresis, mutagenized constructs were transformed into *E. coli* DH5α bacteria (as described in section 2.2).

2.4.2 Insertion of mutagenized insert into pIRES-puro2 vector

Mutated luminal PAM constructs were sub-cloned into pIRES-puro2 vectors to allow for the generation of puromycin-resistant stable cell lines in HEK293 cells. Mutagenized inserts were generated by digestion of the mutagenized pCMV6-R36S hPAMv3 (luminal) construct by BglII/Xma1 (Promega, Wisconsin, USA), blunted by T4 polymerase (Promega), and sub-cloned into Hpa1-digested pIRES-puro2 plasmids to generate puromycin-resistant constructs. Constructs in both pCMV6- and pIRES-puro2-PAM backbones were verified by Sanger sequencing and checked to ensure no other mutations were introduced (Source Bioscience and Eurofins Genomics).

2.5 GENERATION OF ENDOC-βH1 CRISPR PAM KO CELL LINE

The PAM KO EndoC-βH1 cells were used as a physiological relevant model for functional characterisation, as they lack expression of the predominant *PAM* transcripts. Functional validation of this cell line ahead of its use in the characterisation of PAM has been performed in collaboration with Antje Grotz (section 4.3.1)

These PAM KO cells were generated and initially cultured by Antje Grotz (DPhil Student, OCDEM). Two single guide RNAs (sgRNAs) were selected from the TKO CRISPR v3 library and cloned into plentiCRISPRv2 constructs (#52961, Addgene, USA). The plentiCRISPRv2 plasmid was digested using FastDigest BsmBI (Fermentas, MA, USA) for 30mins and purified using gel-based extraction. sgRNA oligos were annealed and phosphorylated using T2 PNK (NEB, US) for 30min at 37°C and ligated using Quick Ligase (NEB, MA, USA) for 1h at RT. This ligated construct was transformed into Stbl3 competent cells (Thermo Fisher, MA, USA) and successful sgRNA integration was confirmed using Sanger Sequencing.

For lentiviral production, Lenti-X HEK293T cells were co-transfected with lentiviral packaging vectors in P/S-free media using JetPrime (Polyplus transfection, France) transfection reagents as per the manufacturer's instructions. The transfection mix consisted of pMD2.G (6.85µg), psPAX2 (10.3µg), the respective cloned plentiCRISPRv2 (12.85µg). Media was replaced after 16h into fresh complete culture media. Supernatant containing viral particles was collected 72h after transfection, spun down for 5min at 2000rpm and filtered through a 0.45µm filter. Supernatant was

ultracentrifuged for 2hrs at 4ºC and 29000rpm in a swinging-bucket rotor. The virus pellet was resuspended in 1.5% BSA in PBS, aliquoted and stored at -80°C.

EndoC-βH1 cells were transduced by lentirviral particles at an multiplicity of infection (MOI) of 8 with constructs containing two sgRNAs targeting *PAM* (exon 2 and exon 12) and the Cas9 enzyme (Table. 2.5.1). Following transduction, cells were subject to a puromycin selection for 7days (3µg/µl, Sigma Aldrich), thereby enriching for transduced cells only. Media was regularly changed to remove dead cells. Cells were maintained in culture in a manner similar to (unedited) EndoC-βH1 cells (section 2.6). To ensure the propagation of transduced cells only, media was supplemented with puromycin (4µg/µl) evert 4-6 passages.

sgRNA	Target exon	Sequence (5' to 3')
sgRNA1	2	GAACTAGCAGGCTAGGGACG
sgRNA2	12	GTTCAGAACCATACCACCAG

Table 2.5.1 Single guide RNA (sgRNA) sequences used in the generation of *PAM* KO EndoC-βH1 cells by CRISPR-Cas9 genome editing. Generation of cell line was performed by Antje Grotz. Single guide RNA (sgRNA) sequences were obtained from the TKO Library v3.

2.6 GENERAL MAINTENANCE OF CELL LINES

The functional characterisation of PAM as described in Chapter 3, 4 and 5 make use of HEK293 and EndoC-βH1 cells as relevant models. HEK293 cells, a human embryonic kidney line, were included as a non-neuroendocrine control, as well as for its lack of endogenous *PAM* expression and high transfection efficiency. EndoC-βH1 cells, the most robust glucose-sensitive human pancreatic β cell line to date, provides a more physiologically-relevant model for functional characterisation, has endogenous *PAM* expression, but does carry technical difficulties such as low transfection efficiency and an inability to survive clonal expansion. CRISPR PAM KO EndoC-βH1 cells were generated by Antje Grotz (DPhil student, Gloyn group) by CRISPR-Cas9 gene editing (section 2.5). All tissue culture activity was carried out in biological safety cabinets in the tissue culture facilities at OCDEM. All cell lines were tested negative for mycoplasma contamination using the MycoAlert Assay kit (Lonza). Cells were maintained at 37°C and 5% CO_2.

For routine cell maintenance, cells were passaged by the following protocol: cells were washed with 1X phosphate buffered saline (PBS) (Sigma Aldrich) and then subsequently lifted from plasticware via addition of 1x trypsin-EDTA solution (Life Technologies) (1ml TE used in a T75, and then scaled accordingly for other formats) and neutralisation by culture media. Cells were then collected and spun down at 1200rpm for 5min at RT (Beckman GS-15R, Beckman Coulter). Cells were then resuspended in culture media and counted using the Cellometer Auto T4 cell counter (Nexcelom Bioscience). Specific culture procedures are summarised in Table 2.6.1.

2.7 TRANSIENT TRANSFECTION

Cells were seeded approximately 24hr prior to transfection in penicillin/streptomycin-free media (for HEK293 cells) or normal media (EndoC-βH1), seeding densities summarized in Table 2.7.1. Forward transfection was performed for all experiments except for gene knockdown by siRNA as detailed in section 2.18. Transfection complexes were made up in Opti-MEM reduced serum media (Thermo Fisher) and then incubated with FuGene at the ratio of 1 DNA: 3 FuGene for 15min at RT. Transfection complexes were then added dropwise to cells and incubated at 37°C/5% CO_2 until harvest (unless stated otherwise).

2.8 STABLE TRANSFECTION

2.8.1 Generation of stable cell lines

For production of PAM stable cell line, HEK293 cells were transfected using the Fugene 6 kit (Promega) as per the manufacturer's instructions. Cells were seeded one day prior to transfection at $4x10^5$ cells/well in a 6 well plate in DMEM 6429 supplemented with 10% FCS. Cells were left to incubate for 72h prior to the addition of selective (puromycin) media. Initial puromycin selection involved growing cells in DMEM supplemented with 10% FCS, 1% P/S as above and supplemented with 2µg/ml puromycin for around 7-10 days. After establishment of stable cell line, the cells were routinely passaged and maintained in selective media containing 1µg/ml puromycin.

Cell type	Splitting procedure	Normal culture conditions
HEK293	Every 2-3 days Split 1:10	Dulbecco's Modified Eagle's Medium (DMEM) (D6429, Sigma Aldrich), 10% (v/v) foetal bovine serum (FBS) (10500- 064, Gibco by Life Technologies), 100 U/ml penicillin and 100 µg/ml streptomycin (15140122, Gibco by Life Technologies)
EndoC-βH1	Once weekly, seeded at $3.6x10^6$ cells per T75 and $9x10^6$ cells per T175. Trypsin neutralised with 7-9ml PBS (Sigma Aldrich) supplemented with foetal bovine serum (FBS) (10500-064, Gibco by Life Technologies).	DMEM (31885, Gibco by Life Technologies), Bovine Serum Albumin (BSA) fraction V (10775835001, Roche), 100 U/ml penicillin and 100 µg/ml streptomycin, 2 mM L-glutamine, 50 µM 2-mercaptoethanol, 10 mM. All plasticware was coated in advance (>2hr) with coating medium consisting of high glucose DMEM (41965, Gibco by Life Technologies), 100 U/ml penicillin and 100 µg/ml streptomycin with 2 µg/m fibronectin (Sigma Aldrich) and 1% (v/v) Extracellular Matrix (ECM) (Sigma Aldrich),
EndoC-βH1 CRISPR PAM KO	Same as EndoC-βH1	Same as for EndoC-βH1 Puromycin selection pressure was added every 4-6 weeks of culture: addition of normal media supplemented with 4µg/ml puromycin (Sigma Aldrich)

Table 2.6.1 Culturing procedures for HEK293, EndoC-βH1 and EndoC-βH1 CRISPR PAM KO cells. Common reagents between cell types were ordered from the same company unless stated otherwise. Generation of the PAM KO EndoC line is summarized in section 2.5.

Cell type	Seeding density	Harvest time (hr)	Specific conditions
HEK293	$4x10^5$/well	48	Seed in P/S-free media
EndoC-βH1*	$5x10^5$/well	72	-

Table 2.7.1 Conditions for transient transfection of *PAM* in a 6-well plate. Unless otherwise stated, cells were transfected by forward transfected using FuGene 6 (Promega) at a DNA: FuGene ratio of 1:3 using 1μg pcDNA/well. *refers to both EndoC-βH1 and EndoC-βH1 CRISPR PAM KO cells. Reagent quantities and seeding densities may be scaled proportionally for different plate formats.

2.8.2 Recombinant PAM protein production

HEK293-PAM stable cell lines were seeded into triple flasks (Nunc. Inc.) and allowed to reach confluence, before being serum-starved with 90ml DMEM 1145 (Sigma-Aldrich) supplemented with 1mM sodium pyruvate, for 3-5 days at 37°C/5% CO_2. PAM-containing supernatant was harvested by taking the media directly from the triple flasks and sterile filtering through 0.45μm filters into sterile storage bottles (Corning). Harvested supernatants were pH corrected to 5.5 with 1M hydrochloric acid and concentrated directly using Centricon-Plus 70 30kDa MW cut-off centrifugal Filter Device (Millipore) as per the manufacturer's instructions. Subsequently, concentrated supernatants were stored at -80°C.

2.9 PROTEIN EXPRESSION ANALYSIS

2.9.1 Cell harvest

Cells were harvested by trypsinisation using 1X trypsin-EDTA solution (Life Technologies) followed by neutralization by either normal media (HEK293) or PBS supplemented with FCS (EndoC-βH1). Cell suspensions were then spun at 1200 rpm for 5 min at room temperature (Beckman GS-15R, Beckman Coulter), then pellets were

suspended in 1ml PBS, transferred to a 1.5ml tube and then spun again to pellet cells. Cells for protein extraction were either kept on ice or frozen at -20°C for short-term storage.

2.9.2 Protein extraction

Unless otherwise stated, samples were lysed in RIPA buffer (Table 2.9.2.1) by resuspension of pellet in 30-100µl RIPA, briefly vortexed and then placed on a rotator (Blood Cell mixer, Voss of Maldon) at 4°C for 30 mins. Following this, samples were spun at 13K rpm for 30mins at 4°C (Heraeus Fresco 21 Centrifuge, Thermo Fisher). Following this, the pellet was discarded and the supernatant retained and quantified by DC assay.

Reagent	Final concentration
Tris pH7.4	50mM
NaCl	150mM
Triton X-100	1%
Sodium dodecyl sulfate (SDS)	0.1%
Sodium deoxycholate	0.5%
Distilled water	-

Table 2.9.2.1 Composition of RIPA lysis buffer used to generate whole cell extracts from pelleted cells. The above was used to make a stock, which was stored at 4°C for up to 3 months. Protease inhibitor cocktail was added at 1:100 to working stocks of RIPA and this was stored at 4°C for up to a week. Unless stated otherwise, all reagents were purchased from Sigma Aldrich.

2.9.3 Western blotting analysis

Protein lysates were separated by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) on a Bio-Rad Criterion TGX stain-free 4-12% pre-cast gel (mini: 15-well; midi: 26-well). The 1X Tris/Glycine/SDS (TGS) gel running buffer consists

of 25 mM Tris, 192 mM glycine and 0.1% SDS pH 8.3 (Bio-Rad Laboratories). Loading samples were prepared by adding 12µl protein to 5µl 4X Laemmli loading buffer (Bio-Rad Laboratories), supplemented with 1µl β-mercaptoethanol (Sigma Aldrich). Samples were heated at 80°C for 15 mins to allow for denaturation of the protein. A constant amount of protein (10µg unless stated otherwise) was loaded per lane for each gel run, together with a molecular weight marker (Precision Plus All Blue standard, Bio-Rad Laboratories). Mini gels were run at 300V for 15mins and midi gels were run at 250V for 30mins.

Subsequently, the gel was activated and imaged using the ChemiDoc MP Imaging System and Image Lab 5 software (Bio-Rad Laboratories). Proteins were transferred to a 0.2µm polyvinylidene difluoride (PVDF) membrane using the Trans-Blot Turbo transfer system (Bio-Rad Laboratories). Following transfer, the membrane was imaged to confirm protein transfer. Membranes were blocked in either 5% non-fat milk powder in TBST or 3% BSA in TBST for 1h before primary antibody incubation. Recipe for TBST has been shown in Table 2.9.4.1. Primary antibodies used are indicated in chapter-specific methods. Unless otherwise indicated, blots were incubated with primary antibodies for 1hr at RT, diluted in either 5% milk/TBST or 3% BSA/TBST. Following incubation, blots were washed in TBST for 3x5min and then incubated with secondary antibody for 1hr at RT. Secondary antibodies (summarized in table 2.9.4.2) were made up in either 5% milk/TBST or 3% BSA/TBST at 1:2500.

Blots were washed for 3x5min in TBST and then developed using ECL reagent (Clarity Western ECL substrate, Bio-Rad) where equivalent volumes of peroxidase and

luminol/enhancer solution are added to the blots, which is catalysed by the HRP conjugated to the secondary antibody resulting in light emission and incubated for 1-5min at RT prior to imaging using the ChemiDoc (Bio-Rad). All blots were washed for 15min in TBST prior to being re-probed using tubulin (sc-365791, Santa Cruz) as a loading control.

Reagent	Final concentration
Tris-HCl pH7.4	1mM
NaCl	5mM
Tween-20 (Thermo Fisher)	0.05%
Distilled water	-

Table 2.9.4.1 TBST wash buffer composition for western blot. TBST was stored at room temperature and made up as needed. All reagents are from Sigma Aldrich unless stated otherwise.

Antibody	Source	Dilution
Anti-mouse HRP	Thermo Fisher, 31450	1:2500
Anti-rabbit HRP	Thermo Fisher, 31460	1:2500

Table 2.9.4.2 Secondary antibodies used for western blot. All antibodies were reconstituted in 50% glycerol in distilled water and stored at -20°C. Antibodies were diluted in 3% BSA/TBST unless stated otherwise.

2.9.4 Densitometric analysis by Image Lab

Quantification of western blot bands was performed using the lane densitometry function of Image Lab 5.0 software (Bio-Rad, California). In brief, identically sized lanes were fitted onto the blot, bands assigned and then intensity values per band exported

to Excel (Microsoft Corporation, Washington, USA). For further details on how band intensities were analysed across experiments, see chapter-specific methods.

2.10 PROTEIN DEGRADATION ASSAYS

To investigate the role of protein degradation in the stability and abundance of variant PAM isoforms in EndoC-βH1 cells, cells were incubated with MG132 (474792, Sigma Aldrich), which is an inhibitor of the 26S subunit of the proteosome and was used at 10μM. DMSO (Sigma Aldrich) was added in equal volume to control wells.

2.11 CELLULAR LOCALISATION STUDIES

To visualise the subcellular localisation of proteins and their relevant variants within the cell, I performed immunofluorescent staining on fixed HEK293 or EndoC-βH1 cells. Cells were seeded in 4-well chamber slides (Falcon, 354114), transfected or treated and then fixed 24-72hr after seeding. Cells were fixed in 4% paraformaldehyde (Sigma Aldrich) following removal of cell culture media and a PBS wash (Sigma Aldrich). Fixed cells were then either stored short-term at 4°C or taken immediately for staining.

2.11.1 Immunofluorescence staining

Fixed cells were permeabilised in 0.1% Triton X-100 (Sigma Aldrich) for 30min at RT and then blocked in 5% swine serum (Dako) in Tris buffer (Sigma Aldrich) for 1hr at RT. Following this, primary antibodies were made up in Tris buffer (details within each chapter) and added to wells at 200μl/well and incubated at RT for 1hr or overnight at 4°C.

Slides were then washed 3x5min in Tris buffer and then incubated in secondary antibody (table 2.11.1.1), made up in Tris buffer. These steps were carried out in the dark to prevent bleaching of the light-sensitive antibodies. For slides co-stained with 2 or more (primary) antibodies, staining occurred in a sequential process to minimise the chance of cross-reactivity between antibodies. Slides were incubated with secondary antibody for 30min at RT in a foil-covered box. Slides were then washed 3x5min in Tris buffer and then incubated in Red dot2, a far-red nuclear stain (Biotium, 40061) at 1:200 in Tris for 15min at RT. Slides were then washed 2x5min in Tris and then mounted in Vectashield mounting media (Vector Labs, H-1200) and a glass coverslip. Slides were then sealed with commercially available nail varnish and left to curate in the dark overnight at 4°C before imaging.

2.11.2 Confocal microscopy

Slides were visualised on a Bio-Rad Radiance 2100 confocal microscope with a 60X 1.0 N.A. water immersion objective, unless otherwise stated. Images for the 3 colour channels (green, red, far-red) were acquired consecutively using the LaserSharp 2000 (Zeiss) software. Images were collected with different laser settings that were optimized for each sample and channel and therefore fluorescent intensities have not been compared across samples. Image files were exported using the LSM Image Browser 4.2 (Zeiss) and arranged with Adobe illustrator CC (Adobe systems).

Antibody	Source	Dilution
Anti-mouse Alexa Fluor 488	Invitrogen, A11029	1:500
Anti-rabbit Alexa Fluor 488	Invitrogen, A27034	1:200
Anti-mouse TRITC	Sigma Aldrich, T7782	1:200
Anti-rabbit biotin	Vector Laboratories, BA-1000	1:50
Streptavidin-488	Thermo Fisher, S11223	1:200

Table 2.11.1.1 Secondary antibodies used in immunofluorescence staining. Unless stated otherwise, antibodies were diluted in Tris buffer (Sigma Aldrich) and added at 200µl/well in 4-well chamber slides (Falcon, 354114).

2.12 CELLULAR ULTRASTRUCTURE STUDIES

To characterise the cellular ultrastructure, immunogold labelling and transmission electron microscopy was performed under the guidance of Dr. Anne Clark (OCDEM). First, cells were fixed in 2.5% paraformaldehyde/0.5% glutaraldehyde in 0.1M phosphate buffer at pH 7.2 for a minimum of 2h, washed in 0.1M acetate buffer and stained in 2% uranyl acetate 1h. Cells were then washed again in acetate buffer, dehydrated in graded methanol at -20°C and incubated in a 1:1 mixture of methanol and London Resin Gold (LRG) embedding medium (Agar Scientific, UK) at -20°C overnight to prepare for embedding. Cells were embedded in pure LRG medium at -20°C for 4h and fresh benzil (Agar Scientific, UK) in LRG at 0.1% w/v which acts as an activator was subsequently added in a clear air-free tube and cured under UV light at -20°C. LRG is an acrylic-based resin that provides better antigen preservation though poorer ultrastructural preservation as compared to epoxy resins. After polymerisation, the LRG blocks were cut into ultrathin sections of 70nm thickness using an ultramicrotome (Reichert-Jung Ultracut, AT) and placed onto 200 mesh Ni^{2+} grids (Agar

Scientific, UK) by Dr Anne Clark at the Department of Physiology, Anatomy and Genetics (DPAG) or OCDEM of the University of Oxford.

2.12.1 Immunogold labelling

Grids were washed in distilled water and blocked in PBS with egg albumin (PBS/EA) for 30min at room temperature. They were then incubated in primary antibody (see chapter-specific methods) diluted in PBS/EA for 1h, washed in PBS/EA, then incubated in secondary antibody (Table 2.12.1.1) for 1h. Grids were subsequently washed in PBS/EA and then distilled water. Finally, grids were contrasted with 2% uranyl acetate and lead citrate, washed with distilled water and dried for storage. Sections were examined in a Jeol 1010 transmission electron microscope with an accelerating voltage of 80 kV (Jeol, JP) fitted with a digital camera (Gatan, CA, USA) at DPAG by Dr Anne Clark.

Antibody	Source	Dilution
Anti-rabbit biotin	Vector Laboratories, BA-1000	1:50
Streptavidin gold 15nm	Biocell, EMSTP15	1:20
Anti-guinea pig gold 10nm	Biocell, EMGAG10	1:20
Anti-rabbit gold 15nm	Biocell, EMGAR15	1:20

Table 2.12.1.1 Secondary antibodies used for immunogold labelling for transmission electron microscopy. Antibodies were diluted in egg albumin in PBS (PBS/EA) unless stated otherwise.

2.13 *IN VITRO* AMIDATION ASSAY

This was used to measure the catalytic capacity of PAM-containing lysates in a colourmetric, cell-free assay that takes into account the PAM content in that particular sample lysate. This allows for quantification of amidation activity (per molecule of PAM) and comparison between wild-type and variant PAM isoforms generated via stable transfection of HEK293 cells (see section 2.8).

2.13.1 Quantification of variant PAM

Concentrated supernatants containing PAM (variant/WT) were quantified to achieve equimolar ratios in the amidation assays. 1µl aliquots (in triplicate) of concentrated supernatant were separated on 4-20% Criterion TGX stain-free gels (BioRad, California, USA) and then transferred to PVDF membrane using Trans-blot Turbo Transfer Pack (BioRad), according to the manufacturer's instructions. The membrane was blocked in 5% milk/PBST for 1 hr at room temperature and then incubated overnight at 4°C in 4A6 anti-Myc tag antibody (Merck-Millipore, Massachusetts, USA). Blots were developed using the Clarity Western ECL substrate (BioRad) as per the protocol provided and imaged on a ChemiDoc MP (BioRad). Lanes were drawn on using Image Lab 5.0 and band intensity quantified. Relative amount of variant PAM protein/µl was then calculated and normalised to that of WT to give a molar scale ratio, which could be fed into the amidation assay set up.

2.13.2 Kinetic amidation assay

The assay was developed in-house by Dr. A. Raimondo (a former post-doc in the group), and is based on the detection of chemically converted glyoxylate product of

the amidation reaction catalysed by PAM as a measure of its activity. The α-amidation reaction catalysed by PAM is summarised in FIG. 2.13.2.1; all reactants and products are in 1:1 stoichiometry with each other, allowing the activity to be inferred via the generation of glyoxylate by-product.

An assay buffer master mix (summarised in Table 2.13.2.2) was aliquoted into a glass vial for each sample of recombinant PAM (wild-type, empty vector, variant PAM). Equimolar amounts of PAM (relative to WT PAM used in that assay, and calculated by western blot) were added to each vial and incubated in a 37°C water bath. Vials were positioned such that the water outside encapsulated all the liquid in the vial and lids were removed as molecular O_2 is required for amidation by PAM. Triplicate 78µl aliquots of each vial were taken every 30 mins and added to a conical 96-well plate (4titude) containing 40mM EDTA pH 8.0 (Sigma Aldrich). After the final time point (t=210 mins), addition of 0.1% phenylhydrazine HCl (Sigma-Aldrich) resulted in the conversion of the glyoxylate product to glyoxylate phenylhydrazone. Plates were then incubated at room temperature for 15 mins and then saturated with concentrated (37%) hydrochloric acid (Sigma Aldrich). The contents of this conical plate were then aspirated and added to a flat-bottomed 96-well plate (Corning) containing 0.2% $K_3Fe(CN)_6$ using an i-Pipette Automated Pipettor (Apricot Designs, California, USA). Absorbance was measured at 515 nm using a VersaMax microplate reader (Molecular Devices, California, USA) to detect 1,5-diphenylformazan. Values were normalised to the maximal activity of wild-type (WT) (of read taken at 210min) to give a % difference in PAM activity per variant over time.

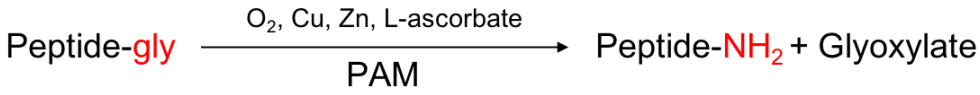

FIG. 2.13.2.1 Summary of α-amidation reaction catalysed by PAM. In the *in vitro* assay, hippuric acid is used as a non-peptide substrate and is present in excess. Given that PAM content is normalised between samples and there is a 1:1 stoichiometry between reaction components, PAM activity is the limiting factor and can be ascertained.

Component	Stock concentration	Final concentration
MES pH 5.5	1M	100mM
Copper sulfate ($CuSO_4$)	200μM	2μM
L-ascorbic acid	1M	10mM
Hippuric acid	0.5M	20mM
Bovine liver catalase	34mg/mL	0.1mg/mL
Distilled water	-	-

Table 2.13.2.2 Components of assay mix for the *in vitro* amidation assay. All components were purchased from Sigma Aldrich.

2.14 RADIOLABELLED SERUM AMIDATION ASSAY

To assess for the amidating profile in individuals, serum was isolated from carriers of alleles for investigation (see chapter-specific introduction for details on which alleles were characterised) and run in a radiolabelled amidation assay. The serum amidating profile is measured by calculating the decay of the radiolabelled isotope substrate: I125 Ac-tyr-val-gly, which is converted into an amidated product by PAM enzyme present in the serum samples. Reactions are made up in individual test tubes (per serum sample) containing, I125 radiolabelled Ac-tyr-val-gly and PAM co-factors (Table 2.14.1) and then topped with 4μl patient serum. Each serum sample is run in triplicate.

The reaction mix is then incubated at 37.5°C for 1hr, after which the reaction is stopped by placing the tubes in ice and chelating the reaction with 0.5mM EDTA pH 8.0. Samples are then counted on a Gamma counter then 700µl water-saturated ethyl acetate added to each sample. The amidated product is extracted using ethyl acetate, since it is more soluble than the glycine-extended precursor. Following phase separation, 350µl of the upper phase was placed into a separate tube and samples counted on Gamma counter for 5min. The serum amidating profile per sample was calculated using the following equation:

$$\text{Amidating profile} = \frac{(2\times amidated\ dpm)-(2\times blank\ dpm)}{20{,}000\ (or\ average\ total\ counts)} \times \frac{5000(substrate\ in\ pmol)}{1\ (time\ in\ hrs)} \times \frac{1}{4\ (vol\ serum\ uL)}$$

Since PAM content in the serum samples are not quantified for this assay, an equivalent volume of serum (4µl) is always added to each reaction tube. Therefore the assay detects the serum amidating profile (pmol/µl/hr of amidated product) rather than amidating activity per molecule of PAM.

Reagent	Stock concentration	Final concentration
Ascorbate	1M	0.5mM
Copper sulfate ($CuSO_4$)	200µM	4µM
NaMES pH 5.5	1M	150mM
Ac-Tyr-Val-Gly*	500µM	0.5µM
Bovine liver catalase	20mg/mL	0.1mg/mL
Distilled water	-	To 50ul

Table 2.14.1 Reaction master mix for serum amidation assay. Components are added to individual test tubes (per serum sample) on ice and then supplemented with 4µl serum. All reagents are purchased from Sigma Aldrich, except for the radiolabelled Ac-Tyr-Val-Gly* isotope which was purchased from Perkin Elmer.

2.15 GENE EXPRESSION ANALYSIS

2.15.1 RNA Extraction

Cells were lysed in TriReagent (1ml/well for a 6-well plate and scaled accordingly for other formats) and then extracted for RNA using the Direct-Zol RNA Miniprep kit (Zymo research, R2051) as per manufacturer's instructions. In brief, following lysis by TriReagent (Thermo Fisher), RNA is precipitated by the addition of 99% ethanol (Sigma Aldrich) and then purified by running through a column. Finally, RNA is eluted in 30μl nuclease-free water and quantified on a nanodrop 2000 spectrophotometer (Thermo Fisher Scientific).

2.15.2 cDNA synthesis

cDNA was synthesised using the GoScript™ Reverse transcription kit (A5001, Promega Corp., Wisconsin, USA) according to manufacturer's guidelines, from RNA samples obtained from cells samples collected in Trizol (section 2.15.1). Reactions were carried out in 96-well PCR plates (4titude Ltd) in the tetrad thermal cycler (DNA Engine, MJ Research). Input RNA was matched across samples within an experiment and were primed a mixture of oligo(dT) and random hexamers. cDNA synthesis reactions were stored at 4°C for short-term or at -20°C for long-term storage.

2.15.3 Real time Taqman® quantitative PCR

Quantitative PCR (qPCR) was performed on the Applied Biosystems 7900HT Fast Real-Time PCR system on the SDS v2.3 software using TaqMan® gene expression master mix (Applied Biosystems) and commercially-available TaqMan® gene expression assays (Applied Biosystems). Details of the assays used are provided in individual results

chapters. All samples were assayed in triplicate on 384-well plates (4titude Ltd) for the gene(s) of interest and at least two housekeeping genes. Each cDNA sample was analysed in triplicate wells and raw Ct values were analysed by the $2^{-\Delta\Delta Ct}$ method, which calculates fold change in gene expression between control and sample cells.

2.16 *IN SILICO* MUTATION PREDICTION ANALYSIS

To ascertain the likely impact of variants characterised in chapter 5, their impact on protein function was assessed by the *in silic*o tools: SIFT, PolyPhen2 and Condel. These tools were accessed via the following online platforms on the 5th February 2018:

- Sorting Intolerant from Tolerant (SIFT): https://sift.bii.a-star.edu.sg
- Polymorphism Phenotyping v2 (PolyPhen2); http://genetics.bwh.harvard.edu/pph2/
- Consensus deleterious score of non-synonymous single nucleotide variants (CONDEL): https://bbglab.irbbarcelona.org/fannsdb/help/condel.html

PolyPhen2 predicts the consequence of non-synonymous SNPs on protein structure and function by assessing impact on the 3D structural and functional features of the human protein. SIFT uses sequence alignment of related proteins to establish which amino acid substitutions at a given position are likely to be tolerated. These predictions are based on the position of the substitution and the type of amino acid change below this cut-off is predicted to be deleterious. Condel collates information from multiple in silico tools (SIFT, MutationAssessor, FatHMM), thereby providing a weighted score on whether a given mutation is deleterious to protein function. These

tools provide a prediction for the impact that a coding variant may have on protein function.

2.17 GENE KNOCKDOWN BY siRNA TRANSFECTION

Gene silencing was carried out on EndoC-βH1 cells using small interfering RNA (siRNA) and lipid-based reverse transfection. siRNA complexes were prepared by combining ON- TARGETplus siRNA (Dharmacon, GE Healthcare) and Lipofectamine RNAiMAX (Life Technologies) in Opti-MEM reduced serum medium (31985, Gibco, Life Technologies) and incubated for 15-20 min at RT. Transfection complexes were aliquoted into wells in 12-well cell culture plates (Corning) before resuspended cells in culture medium were added gently to the wells. Media was changed on 24 hours after gene knockdown and cells harvested for RNA or protein 72hr post-knockdown.

2.18 INSULIN SECRETION ASSAY

2.19.1 Glucose stimulation

For secretion experiments in EndoC-βH1 cells, glucose-free complete culture medium was made up in DMEM (11966, Gibco by Life Technologies) supplemented as stated in section 2.6.2 and with the appropriate amounts of glucose (Sigma Aldrich). Cells were placed in low glucose (2.8mM) media overnight either 48-hours after transfection or 24hrs after treatment. The following day, cells were starved in 0mM glucose for 30 mins and then placed in fresh pre-warmed condition media at the following conditions: 2.8mM Glucose, 10mM Glucose, 16.7mM Glucose, 16.7mM Glucose with 10µM Forskolin (Sigma Aldrich) and 100µM tolbutamide (Sigma Aldrich). Each condition was carried out with either 3 or 6 technical replicates within each experiment. After 1h of static incubation at 37°C, 80% of the supernatants were

carefully removed from the wells into 96-well plates (4titude) and spun down at 1000rpm for 5 min at 4°C (Allegra 64R centrifuge, Beckman Coulter). Then a further aliquot of supernatant (60µl) was removed into a fresh 96-well plate (4titude), parafilmed and frozen at -20°C. Remaining supernatant on the cells was discarded and cellular insulin content was harvested by the addition of ice-cold acid ethanol (Sigma Aldrich) at 100µl/well. Plates were then parafilmed and frozen at -20°C until analysis.

2.19.2 AlphaLISA assay

The concentration of insulin in supernatants and contents were measured using the human insulin AlphaLISA detection kit (AL204C, Perkin Elmer) according to manufacturer's guidelines. The immunoassay detects both proinsulin and insulin. Insulin supernatant samples were diluted 1:10 and insulin content samples were diluted 1:100 in 1X immunoassay buffer. Diluted samples were aliquoted into white 96-well 1⁄2-area microplates (Perkin Elmer) for the assay. Plates were read on the Enspire multimode plate reader (Perkin Elmer) using the Alpha Screen protocol. Insulin concentrations in the samples were calculated based on a 10-point standard curve generated using the insulin analyte (AL204S, Perkin Elmer) with every run. For details pertaining to how data was normalised, see chapter-specific methods.

2.19 STATISTICAL ANALYSES

Unless otherwise stated, analysis of data comparing one parameter between control and tested samples were first normalised to control sample(s) within each experiment, and then transformed by taking the $\log_2$ (hereafter referred to as log) of the normalised values. For comparison of one variable between 2 groups (e.g. control vs. sample) transformed values were analysed by using two-tailed unpaired Student's t-

test with Welch's correction. For comparison of 2 or more variables between 3 or more test groups, transformed values were analysed by one-way or two-way analysis of variance (ANOVA). This allowed for the direct comparison of multiple (>2) test groups (e.g. WT, S539W and R36S) within one single test. In chapter 3, comparison between PAM protein abundance between low and high glucose was analysed by the two-tailed unpaired Student's t-test, as data were not normalised to control. Plotting of graphs and statistical analyses were carried out on GraphPad Prism 8.0. A P-value <0.05 was considered statistically significant.

CHAPTER 3

Characterisation of endogenous PAM in the pancreatic β cell

Results from this chapter have contributed to the following manuscript:

S.K Thomsen*, A. Raimondo*, B. Hastoy*, S. Sengupta et al (2018)
"Type 2 diabetes risk alleles in *PAM* impact insulin release from human pancreatic β-cells."
Nature Genetics 50(8): 1122-1131

The full manuscript can be found in the Appendix

Unless stated otherwise, only the data I contributed to the manuscript has been included

3.1 INTRODUCTION

3.1.1 The role of PAM in glucose homeostasis

T2D-associated alleles in *PAM* were first identified to be associated with reduced insulinogenic index (Huyghe, Jackson et al. 2013, Steinthorsdottir, Thorleifsson et al. 2014, Fuchsberger, Flannick et al. 2016, Mahajan, Wessel et al. 2018). Functional *in vitro* characterisation of the two independently associated coding variants, D563G and S539W, revealed that both cause loss-of-function (LoF), which affected either amidation activity or protein stability respectively (Thomsen, Raimondo et al. 2018). This suggests that LoF alleles in *PAM* negatively affect β cell function and more specifically insulin secretion.

To fully understand the functional consequence for variants in *PAM* in the β cell, a comprehensive understanding of the normal, tissue-specific role for PAM is required. Previous efforts have identified roles for PAM in islets largely through the use of rodent models (Scharfmann, Leduque et al. 1988, Czyzyk, Ning et al. 2005). Heterozygous global knockout *Pam*$^{-/-}$ mice develop glucose intolerance when aged compared to their matched littermate controls, suggesting that PAM has a role in glucose homeostasis (Czyzyk, Ning et al. 2005).

Carboxyamidation is mediated by the sequential activity of the 2 catalytic domains present in the membrane-integral isoform of PAM, which results in the conversion of a C-terminally glycine-extended precursor to an amidated peptide (Eipper and Mains 1988). Amidation has been proposed to be key in maximising biological potency, activity and stability of the biological substrate (Chufan, De et al. 2009).

3.1.2 Elucidating tissue-specific functions for PAM in the pancreatic β cell

PAM is widely expressed in neuroendocrine tissues, but the nature of its function appears to differ between cell types (El Meskini, Mains et al. 2000). It therefore follows that PAM may have tissue-specific processing allowing for these distinctions in function (Eipper, Green et al. 1992, El Meskini, Mains et al. 2000). The biophysical parameters for carboxyamidation are well-established; characterisation has been performed largely in the anterior pituitary, where the predominant peptide proopiomelanocortin (POMC) is amidated by PAM (Kumar, Mains et al. 2016). However, insulin is not amidated by PAM. Therefore, mechanisms impacting insulin secretion are likely to be mediated through carboxyamidation of intermediate peptides or by non-catalytic functions of PAM in the β cell. The function of PAM in human pancreatic β cells and how it functions to regulate insulin secretion is currently undescribed.

Gene silencing of *PAM* in EndoC-βH1 cells, identified a role for PAM in insulin secretion and content (FIG. 3.1.2.1). This resulted in a reduction in the number of immunogold-labelled insulin particles in granules and a trend towards increased vesicle density (FIG. 3.1.2.2) (Thomsen, Raimondo et al. 2018). Proopiomelanocortin (POMC), a known PAM substrate, is synthesised and secreted from granules in the anterior pituitary (Kumar, Mains et al. 2016).

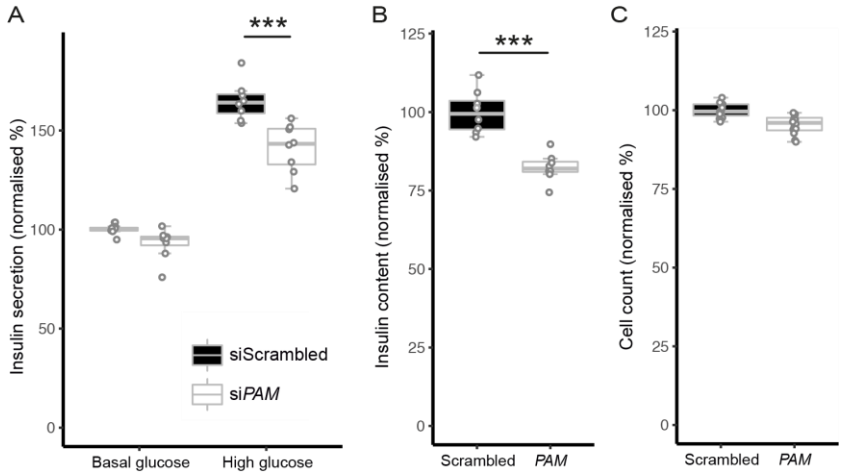

FIG. 3.1.2.1 *PAM* gene silencing results in reduced insulin secretion driven by a defect in insulin content in EndoC-βH1 cells. (**A**) Insulin secretion following PAM knock down (KD) is reduced at high glucose (10mM) compared to scrambled (Scr) controls. (**B**) Insulin content is reduced following *PAM* KD, when adjusted for cell count (**C**), which is unaffected. Adapted from (Thomsen, Raimondo et al. 2018). *** denotes P<0.001.

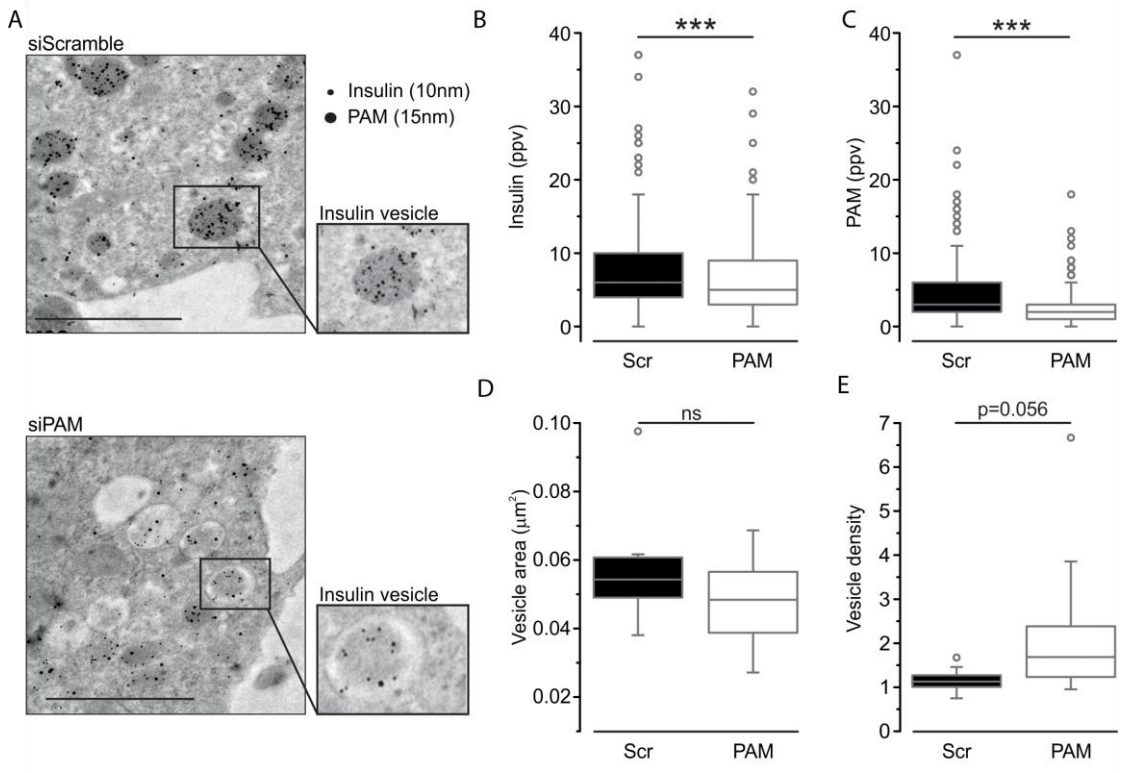

FIG. 3.1.2.2. *PAM* gene silencing in EndoC-βH1 cells results in reduced immunogold-labelled insulin particles per vesicle. (**A**) Electron micrographs following transfection with either scrambled or siPAM oligos. Quantification of the number of immunogold labelled particles for insulin (**B**) and PAM (**C**) per vesicle; the vesicle area (**D**) and vesicle density (**E**). Adapted from (Thomsen, Raimondo et al. 2018). *** denotes P<0.001.

3.1.3 Experimental aims

In this study, I aim to characterise the tissue-specific mechanisms by which PAM functions in insulin secretion and content in the human pancreatic β cell line, EndoC-βH1 cells. To achieve this, I will be examining the following experimental avenues:

- Elucidating the isoform-specific mRNA expression of *PAM* in human islets using long read RNA sequencing
- Characterisation of the expressed and cleaved PAM protein isoforms present in human β cells thereby examining putative tissue-specific processing
- Identifying the localisation of PAM in human beta cells using immunofluorescence and electron microscopy
- Examination of proinsulin expression and biosynthesis following PAM depletion
- Identification of chromogranin A as a likely substrate and mediator of PAM function
- Exploration of how chromogranin A function may be disrupted in the human β cell by depleted PAM function.

3.2 METHODS

3.2.1 Long read RNA sequencing

RNA extraction was performed on two independent samples of primary islets (ISL-021 and ISL-094) isolated at the Oxford Centre for Diabetes, Endocrinology and Metabolism Islet Isolation Facility by Dr. Carla Burrows (OCDEM). RNA was quantified using the Bioanalyser (Agilent, California, USA) to ensure that RNA was of sufficient quality (RNA integrity number, RIN score of >8) and yield to send for sequencing. Samples were then sent for long read sequencing at Pacific Bioscience (hereafter referred to as

PacBio, California, USA). Analysis of raw reads was performed using the PacBio IsoSeq pipeline, and were then mapped to human genome to identify transcripts. Resulting transcripts were then used as a reference for quantification with islet RNA-Seq data. Analysis was performed by Dr. Anne Ndungu (Wellcome Centre for Human Genetics, University of Oxford).

3.2.2 Transfection of EndoC-βH1 cells

EndoC-βH1 cells were cultured, split once weekly and seeded into pre-coated plasticware as described in section 2.6. For gene silencing, EndoC-βH1 cells were reverse transfected with 20nM SMARTpool ON-TARGET plus siRNA complexes for *PAM* or non-targeting control oligos (Table 3.2.1.1) (Dharmacon, Colorado, USA), made up in reduced serum OptiMEM (Thermo Fisher Scientific) using Lipofectamine RNAi reagent (Thermo Fisher, 13778030). Cells were incubated with siRNA oligos for 24hrs prior to replacement of culture media, and then harvested 48hr post-media change. I achieved a robust knock-down of *PAM* at both mRNA (-91%, n=3, P=0.017) and protein level (-35.7%, n=3, P=0.09) (FIG. 3.2.1.2).

Oligo target	Sequence (5' to 3')
PAM	CGUAAGGGCUACAGUCGAA
	GCAAAUAGGACUCGGACCA
	UCUUCAAGCCAGUGCGCAA
	GGUAGUAAGUGGAUACAGA
Non-targeting	UGGUUUACAUGUCGACUAA

FIG. 3.2.1.1 siRNA oligos used for PAM silencing experiments. All oligos were purchased from Dharmacon and were used at 20nM in reduced serum media.

For catalytic inhibition experiments, cells were treated with 500μM 4-phenyl 3-butenoic acid (4P3BA) or an equivalent volume of DMSO and incubated for 48hrs prior to harvest. Cells were either subject to lysis by RIPA buffer for protein extraction, or TriReagent for RNA extraction. RNA/cDNA was stored at -80°C and protein lysates at -20°C prior to qRT-PCR or western blotting respectively.

3.2.3 Immunofluorescence

EndoC-βH1 cells were seeded into 4-well chamber slides and fixed in 4% PFA approximately 48hrs after seeding. Fixed cells in chamber slides were then either sealed with parafilm and stored at 4°C or washed in PBS for immunostaining. For the latter, fixed cells were immunostained following the protocol detailed in section 2.11. Primary antibodies used are listed in Table 3.2.2.1, secondary antibodies listed in Table 2.11.1.1. Endogenous PAM (Cter PAM antibody) was amplified using biotin/streptavidin-488, which provides ~4x signal amplification. A minimal volume of 200μl/well was used for each of the incubations. Slides were visualised using a BioRad Radiance 2100 confocal microscope, images exported using the Zeiss LSM software and arranged using Adobe Illustrator (Adobe Systems)

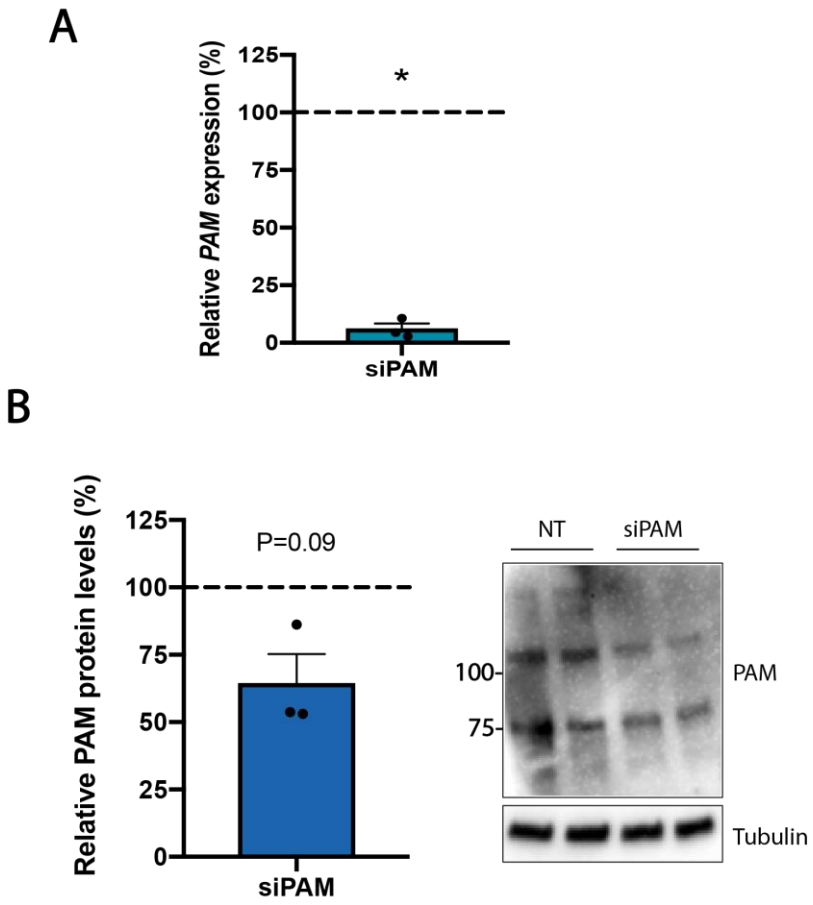

FIG. 3.2.1.2 Knock-down (KD) efficiency of *PAM* silencing in EndoC-βH1 cells. EndoC-βH1 cells were incubated with either non-targeting (NT) or siPAM oligos for 48hrs and then KD efficiency assessed using qRT-PCR (A) and western blotting (B) to assess *PAM* mRNA expression and PAM protein levels respectively. Data are normalised to NT-transfected cells (dashed line) and are representative of 3 independent experiments and are analysed by two-tailed unpaired Student's t-test with Welch's correction on transformed (log) data. * denotes P<0.05 and error bars are SEM.

3.2.4 Western blotting analysis

Cells were lysed using RIPA buffer, quantified using the DC assay (BioRad) and separated using SDS-PAGE (BioRad Criterion). Lysates were denatured by boiling at 80°C in 4x Laemmli buffer (BioRad) supplemented with β-mercaptoethanol; 10μg

protein was loaded per lane. Blots were blocked in 3% BSA/TBST unless stated otherwise. Antibodies used for Westerns are summarised in Table 3.2.2.1 and were incubated either overnight at 4°C or for 2hr at RT (see section 2.9 for more details). Visualisation of tubulin was always included as a loading control. Bands were quantified using lane densitometry (Image Lab 5.0, BioRad) and were adjusted for tubulin band intensities. Adjusted intensities were then normalised to the control within each experiment; each experiment was repeated at least 3 times and mean normalised band intensities were plotted in GraphPad Prism 8.0.

Antibody	Source	Dilution (WB)	Dilution (IF)	Dilution (EM)
Anti-CgA (rabbit polyclonal)	Abcam, ab15160	1:1000	1:100	1:50
Anti-CgA-gly (rabbit monoclonal)	Abcam, ab52983	1:1000	1:100	-
Anti-PAM (mouse monoclonal)	Santa Cruz, sc-514110	1:1000	1:100	1:20
Anti-PAM (rabbit monoclonal)	Abcam, ab109175	-	1:100	1:100
Anti-insulin (mouse monoclonal)	Santa Cruz, sc-377071	1:1000	-	-
Anti-C-peptide (mouse monoclonal)	ExBio, 11-247-C100	1:1000	1:200	1:50
Anti-insulin (Guinea pig polyclonal)	Dako, A0564	-	1:750	1:750
Anti-tubulin (mouse monoclonal)	Santa Cruz, sc-365791	1:2000	-	-

Table 3.2.2.1 Primary antibodies used for western blotting (WB), immunofluorescence (IF) and electron microscopy (EM). Antibodies were diluted in 3% BSA/TBST for western blot (WB), in Tris buffer for immunofluorescence (IF) and in PBS/egg albumin (PBS/EA) for electron microscopy (EM).

3.2.5 Transmission electron microscopy (TEM)

Transmission electron microscopy was performed under the guidance of Dr. Anne Clark at OCDEM and at the Department of Physiology, Anatomy and Genetics (DPAG). In brief, EndoC-βH1 cells for immunogold labelling for TEM were pelleted by trypsination (as per section 2.6) and processed as described in section 2.12. Immunogold labelling was performed using primary antibodies listed in table 3.2.3.1, with amplification using biotin/streptavidin conjugated gold particles for PAM (C-terminal) only.

3.2.6 Quantitative real-time PCR (qRT-PCR)

For quantification of gene expression, RNA was extracted, cDNA synthesised and qRT-PCR set up as described in section 2.15. For the gene expression analyses performed in this chapter, details on the Taqman probes used are summarised in Table 3.2.5.1. Data was analysed using the ddCT method, comparing gene expression of a test sample to a control sample in the same plate.

Taqman® probe	Assay ID	Specificity
*GAPDH**	Hs99999905_m1	Exon 3
*PPIA**	Hs99999904_m1	Exon 5
*TBP**	Hs00427620_m1	Spans exon 2-3
PAM	Hs00168596_m1	Spans exon 11-12
CHGA	Hs00900370_m1	Spans exon 2-3
INS	Hs00355773_m1	Spans exon 2-3

Table 3.2.5.1 Taqman probes used for assaying gene expression by qRT-PCR. All Taqman probes were purchased from Life Technologies. * indicates probes used as housekeeping genes in each qRT-PCR run.

3.2.7 Statistical analysis

Unless otherwise stated, differences between normalised data was analysed by first transforming values into log as the normalised data was not normally distributed. Statistical analyses were then performed using the two-tailed unpaired Student's t-test with the Welch's correction to compare one variable between two groups. Experiments were repeated at least 3 times and mean normalised values collated and plotted on GraphPad Prism 8.0. For ease, the data have been plotted as the normalised averages rather than log values. A $P<0.05$ has been deemed as statistically significant.

3.3 RESULTS

3.3.1 Isoform-specific *PAM* mRNA expression in human pancreatic islets

It is likely that both the mRNA expression of PAM isoforms and post-translational modifications are tissue-specific and that this may dictate the functional capacity for PAM in any specific cell type. Determining isoform-specific expression is a challenge, but can be key to understanding the full function of *PAM* in a tissue-specific manner. Short read sequencing as typically generated, only provides ~100bp reads and therefore cannot provide the resolution or coverage to identify isoform-specific expression with confidence. Long read sequencing, such as that provided by Pacific Bioscience (California, USA), provides greater transcript certainty and reads of greater than 500bp, thereby allowing identification of specific *PAM* mRNA isoforms (Rhoads and Au 2015). In human pancreatic islets, full-length *PAM* isoform (ENST00000438793) has the greatest abundance at 121 transcripts per kilobase million (TPM), accounting for 95% of the total *PAM* expression, with remaining isoforms identified listed at less than 3 TPM in abundance (Table 3.3.1.1).

Establishing the relative expression levels of specific isoforms can indicate the consequential protein isoforms that exist in the cell. The corresponding protein isoform of the predominant mRNA isoform ENST00000438793 translates to an integral protein isoform, which prior to any post-translational modifications, contains all the catalytically active and functional domains.

Transcript ID [a]	Length [b] (bp)	Protein isoform [c]	Size [d] (aa)	MW [e] (kDa)	TPM
ENST00000438793	5432	Integral	973	~120	121.120069
ENST00000304400	3156	Integral	974	~120	2.84602
ENST00000502472	540	No protein	-		1.274562
ENST00000509523	539	No protein	-		1.171969
ENST00000455264	2898	Luminal	905	100	0.362217
ENST00000510208	824	No protein	-		0.293786
ENST00000506127	510	No protein	-		0.161517

[a] ID as per Ensembl annotations is based on the Human GRCh38.p12 assembly for *PAM*
[b] Length of mRNA transcript prior to translation
[c] Listed as the type of corresponding protein isoform generated following translation and before any post-translational modifications or processing can occur. Membrane-integral PAM is denoted by the presence of C-ter transmembrane domain, which the luminal isoform lacks
[d] Size of corresponding protein isoform in terms of number of amino acids
[e] Putative molecular weight of corresponding protein isoform in kDa, as deduced by the assumption that 1 aa = 110 Da
Abbreviations listed as - bp: base pairs; aa: amino acids; MW: molecular weight; kDa: kilo Daltons; TPM: transcripts per kilobase million

Table 3.3.1.1 Isoform-specific expression of *PAM* in human pancreatic islets.
Pacific Bioscience (PacBio) sequencing was performed to identify the isoform-specific mRNA expression of *PAM* in human pancreatic islets. Table summarises the summary statistics for each of the isoforms identified including information pertaining to any corresponding protein isoforms after translation. Listed isoforms are ranked in order of expression in TPM, listed in the far-right column. Data collated as averages of multiple independent legacy samples.

3.3.2 Characterisation of PAM protein isoforms in the pancreatic β cell

Whilst the spliced (mRNA) isoforms will inform on the transcripts that will be expressed, function will be primarily mediated through the translated protein, which may undergo post-translational modifications (e.g. proteolytic cleavage). Characterisation of PAM processing in the anterior pituitary identified the need for processing enzymes present only in neuroendocrine cells and implied a link between tissue-specific cleavage and subsequent cellular function (Thiele, Marek et al. 1989). Knowledge of the extent to which PAM is cleaved post-translationally in the cell is therefore critical to understanding the functional demand for PAM in the β cell.

I postulated that newly translated PAM may be extensively modified and cleaved to generate smaller fragments in a tissue-specific manner and that the proportion and type of these cleaved isoforms may differ at low and high glucose states. In addition, untethered luminal PAM fragments (hereafter referred to as luminal) generated through endoproteolytic cleavage may be co-secreted with insulin following stimulation with glucose as per previous reports of PAM being co-secreted with hormones via regulated secretion (El Meskini, Galano et al. 2001, Bonnemaison, Bäck et al. 2014). To explore this, I examined the endogenous PAM species identified in both cell extracts and those secreted into the supernatant of EndoC-βH1 cells, incubated at different glucose conditions for 48 hours to represent the basal and stimulated stages of regulated secretion.

Lysis of EndoC-βH1 with RIPA buffer allows the examination of both newly translated protein and isoforms generated via endoproteolytic cleavage, thereby providing a

comprehensive overview of PAM species in the β cell. Under both basal (2.8mM) and suprastimulatory glucose conditions (16.7mM), three predominant PAM species were observed in the cell extract (FIG. 3.3.2.1), of which the largest (120kDa) corresponds in size to the translated protein of the predominant mRNA isoform identified in Table 3.3.1.1. The presence of multiple weaker bands laddered between these major bands may indicate further processing and modifications of PAM (e.g. glycosylation, phosphorylation) in the β cell, although these modifications were not further characterised.

Cumulative quantification of the three major bands at 120kDa, 75kDa and 50kDa respectively was performed to give an estimate of total PAM abundance. This showed no difference in the total PAM abundance (+80%, n=3, P=0.43) in the cell extract following incubation with 16.7mM glucose. I next calculated the relative proportion of each isoform (as a % of total PAM) to ascertain whether glucose-stimulation affected PAM processing. Processing in response to stimulation by secretagogues may point to specific functions for PAM that are only important under glucose-stimulated conditions. The relative proportions of each isoform were similar between basal and suprastimulatory glucose conditions, suggesting that this was not the case (FIG. 3.3.2.2). The most enriched isoform (termed Band 3 in FIG 3.3.2.1) is ~50kDa in size and constitutes 72% (at 2.8mM glucose) and 62% (at 16.7mM glucose) of total PAM respectively (FIG. 3.3.2.2). Deductions made from the size, position of the cleavage sites, and epitope of antibody suggest that this 50kDa isoform may be a membrane-integral monofunctional PAM species; although without further interrogation this cannot be confirmed.

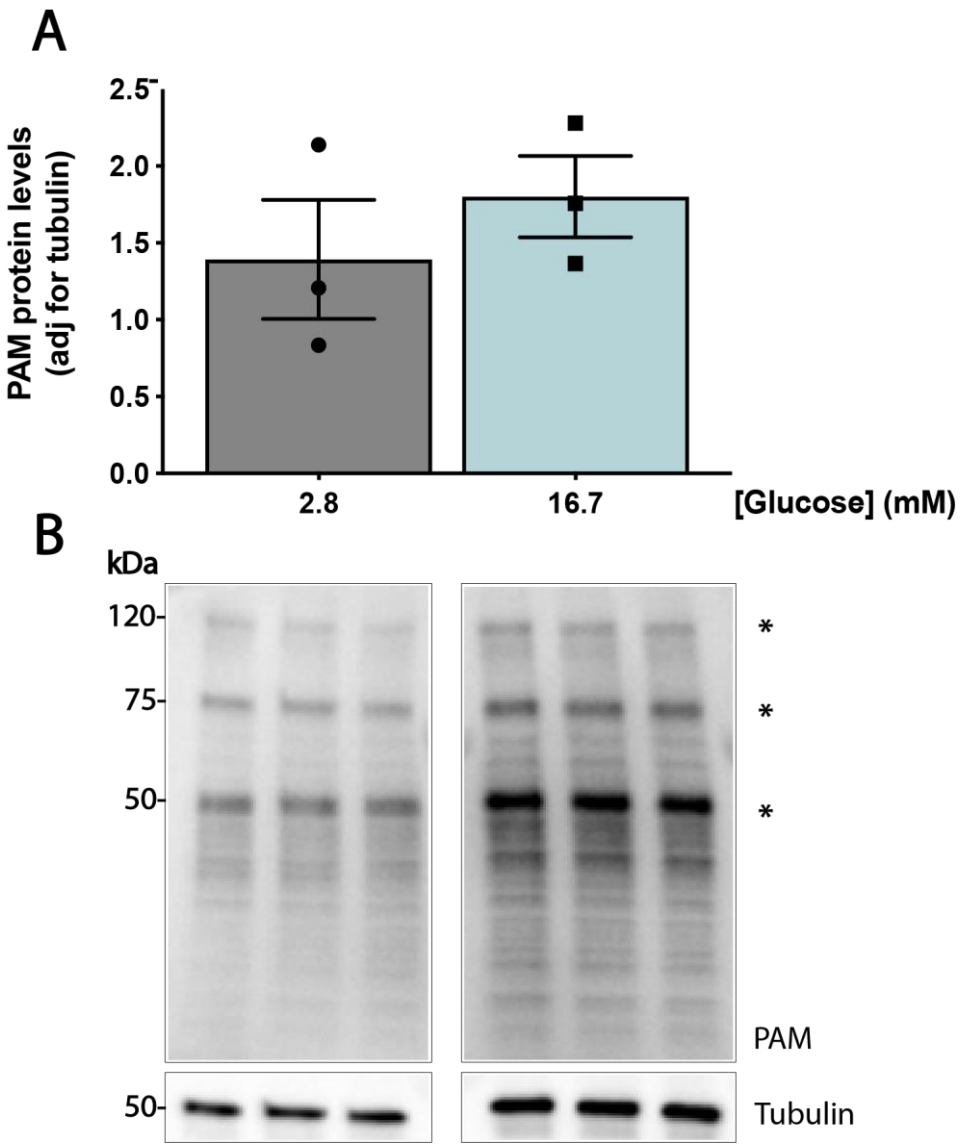

FIG. 3.3.2.1 Newly translated PAM is endoproteolytically cleaved to generate smaller isoforms. EndoC-βH1 cells were cultured at either 2.8mM or 16.7mM glucose conditions for 48hrs and then lysed to assess post-translational cleavage of PAM species in cell extract by western blotting (**B**). Quantification of total PAM (**A**) revealed no significant differences in the protein expression between glucose conditions, as assessed by two-tailed unpaired Student's t-test. Data are representative of 3 independent experiments, error bars are SEM. * indicates the bands deemed to be predominant PAM species in the β cell.

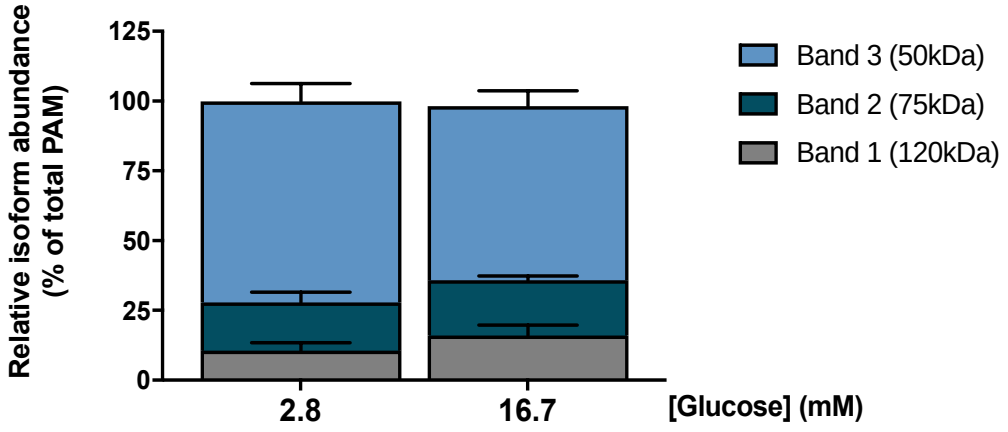

FIG. 3.3.2.2 Relative abundance of the 3 predominant PAM isoforms at basal and suprastimulatory glucose conditions in EndoC-βH1 cells. Band intensity for the predominant PAM species observed in FIG. 3.3.2.1(B) were quantified by lane densitometry and plotted as % of total PAM per glucose condition. No significant differences were observed as assessed by 2-way ANOVA of 3 independent experiments. Error bars indicate SEM.

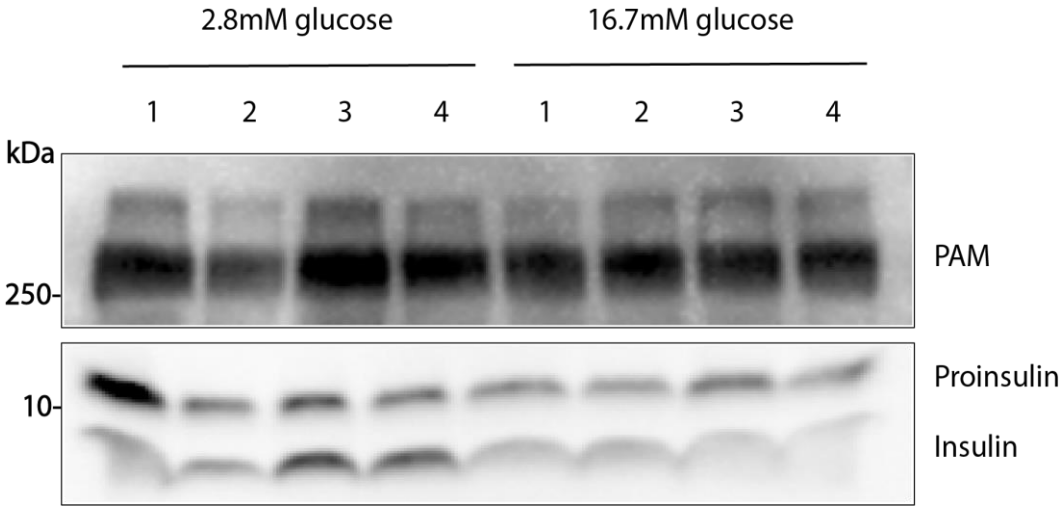

FIG. 3.3.2.3 PAM species secreted into the supernatant of EndoC-βH1 cells. Cells were cultured for 48hrs in BSA-free media containing either 2.8mM or 16.7mM glucose. Supernatant was collected, sterile filtered and concentrated 4-fold and then run on a western blot. Blot contains samples from 4 independent experiments as indicated by numbers above lanes. Insulin was included as a relevant marker for supernatant.

PAM could not be detected in the supernatant (data not shown) – although this is likely due to technical limitations in attempting to detect secreted PAM by Western blot. To address this, supernatant was concentrated ~4-fold in volume using Amicon Ultra-2 centrifugal filter units (MW cut off 30kDa, Merck-Millipore) and re-run using western blotting. A heavy band at ~250kDa was identified, but there was no obvious difference in relative amounts secreted between basal and high glucose (FIG. 3.3.2.3).

3.3.3 Subcellular localisation of PAM in human pancreatic β cells

I next utilised immunofluorescence to determine the subcellular localisation of PAM protein in human β cells. Identifying the subcellular compartments in which a protein resides can be particularly useful when trying to ascertain key cellular sites of function. Fixed EndoC-βH1 cells were co-stained for PAM and insulin using an antibody targeting the C-terminus of PAM. Co-localisation of PAM and insulin is observed near the nucleus with some punctate co-staining at the cell periphery (where the cytoplasm meets the plasma membrane) was observed (FIG. 3.3.3.1). This is in keeping with the previously reported co-localisation of PAM with the *trans*-Golgi network (TGN) and secretory granules in the anterior pituitary, gastrointestinal tract and pancreatic β cells (Martinez, Burrell et al. 1993, Martinez, Montuenga et al. 1993, Milgram, Kho et al. 1997).

To confirm this with greater resolution, immunogold labelling and transmission electron microscopy (TEM) was performed in EndoC-βH1 cells. This showed co-localisation of PAM with insulin in secretory granules, with very little labelling for PAM

in other subcellular compartments (FIG. 3.3.3.2). There appeared to be co-labelling in granules at varying stages of maturation. Other antibodies targeting epitopes at other positions in PAM (e.g. N-terminus and medial-PAM) were not successful in either immunofluorescence or immunogold labelling (data not shown).

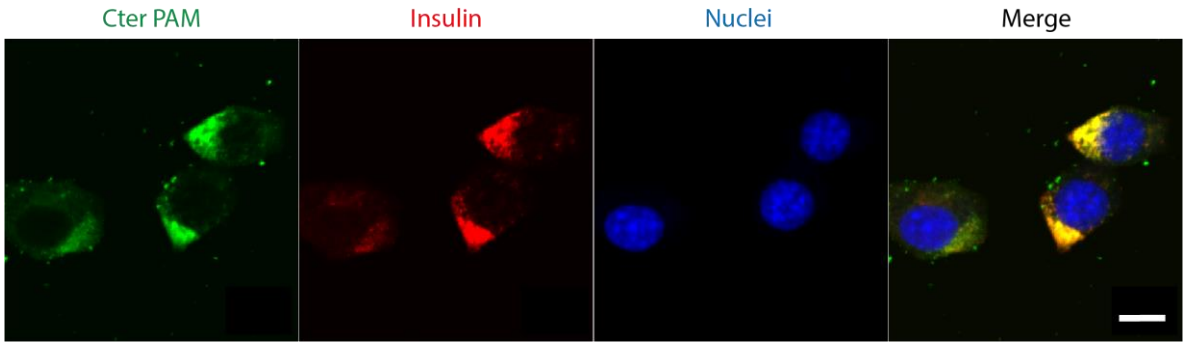

FIG. 3.3.3.1 PAM localises in punctate, granular-like staining at the periphery of the cytoplasm. Fixed EndoC-βH1 cells were double immunostained for PAM (Cter, green) and insulin (red) and nuclei stained using the far-red dye Red dot2. PAM staining was amplified using biotin/streptavidin-488 (x4 amplification). Cells were visualised using a BioRad Radiance 2100 confocal microscope. Scale bar is 50μm.

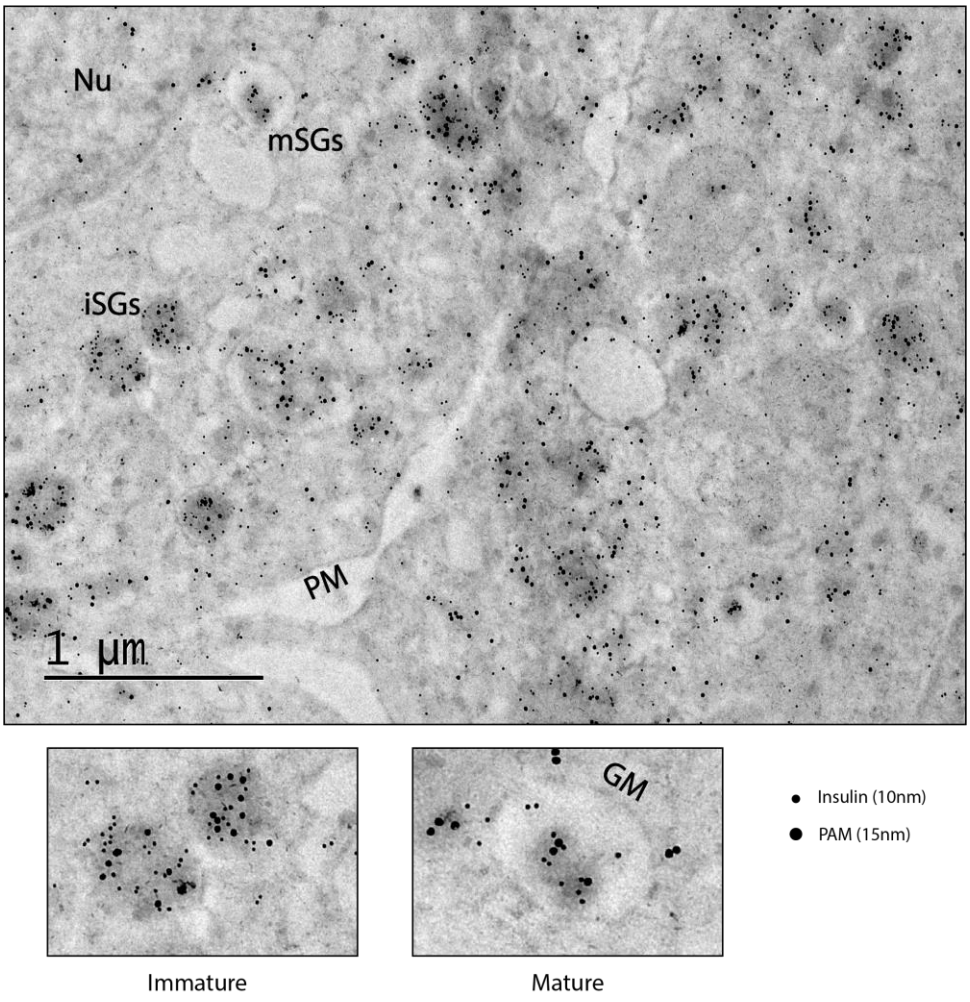

FIG. 3.3.3.2 Electron micrographs showing PAM labelling in insulin-containing secretory granules in EndoC-βH1 cells. Cells were immunogold labelled with both PAM (15nm gold) and insulin (10nm gold) and imaged using transmission electron microscopy. PAM co-labelling was observed in both immature and mature granules (lower panel) with little labelling in other organelles. iSGs: immature secretory granules; mSGs: mature secretory granules (denoted by presence of halo); Nu: nucleus; PM: plasma membrane; GM: granule membrane.

3.3.4 Reduced insulin content is mediated primarily through reduced proinsulin expression

The observation that silencing of *PAM* reduces insulin content (FIG. 3.1.2.1), may suggest a defect in the expression, processing or packaging and storage of (pro)insulin following loss of PAM function (Thomsen, Raimondo et al. 2018). To investigate a potential impact on expression, I performed qRT-PCR following *PAM* KD in EndoC-βH1 cells. I observed a trend towards reduced *INS* expression following *PAM* KD (-49.9%, P= 0.12) as compared to cells transfected with non-targeting oligos (FIG 3.3.4.1).

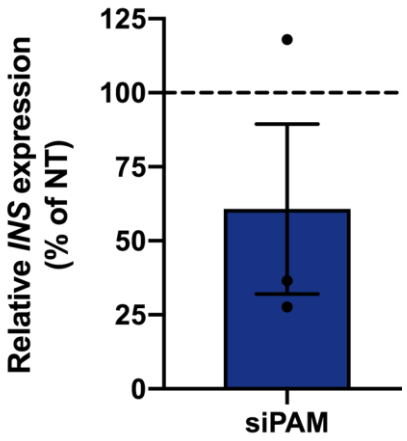

FIG. 3.3.4.1 *PAM* KD results in trends towards reduced *INS* expression in EndoC-βH1 cells. Cells were transfected with either non-targeting (NT) or siPAM oligos and incubated for 48hrs prior being assessed for gene expression by qRT-PCR. Data are representative of 3 independent experiments, are plotted as normalised averages to NT control (dotted line) and analysed by two-tailed unpaired Student's t-test with Welch's correction on transformed (log) data. Error bars are SEM.

I considered whether this effect would translate to an effect at protein level, and therefore ran western blots of cell lysates collected following *PAM* KD. Antibodies capable of distinguishing between proinsulin and insulin species (Table 3.2.3.1) based on their epitope and size on western blots were used to distinguish an impact on proinsulin expression from an effect on the conversion to mature insulin in EndoC-βH1 cells. I observed a reduction in the abundance of proinsulin (-51.4%, n=3, P=0.01) following *PAM* KD but no impact on insulin abundance (P=0.71) (FIG. 3.3.4.2). Irreversible catalytic inhibition of PAM by 4P3BA also showed a trend towards reduced proinsulin levels (-14.2%, n=4, p=0.09) and similarly no impact on insulin abundance (P=0.23) (FIG 3.3.4.3). Despite these impacts on proinsulin, both *PAM* KD and inhibition by 4P3BA resulted in no significant differences in the proinsulin: insulin ratio (FIG. 3.3.4.4).

These investigations imply that the reduced insulin content observed following loss of PAM function may be caused by reduced proinsulin biosynthesis (effects active at both mRNA and protein level) and that this may be driven, at least partially, by impaired amidating activity.

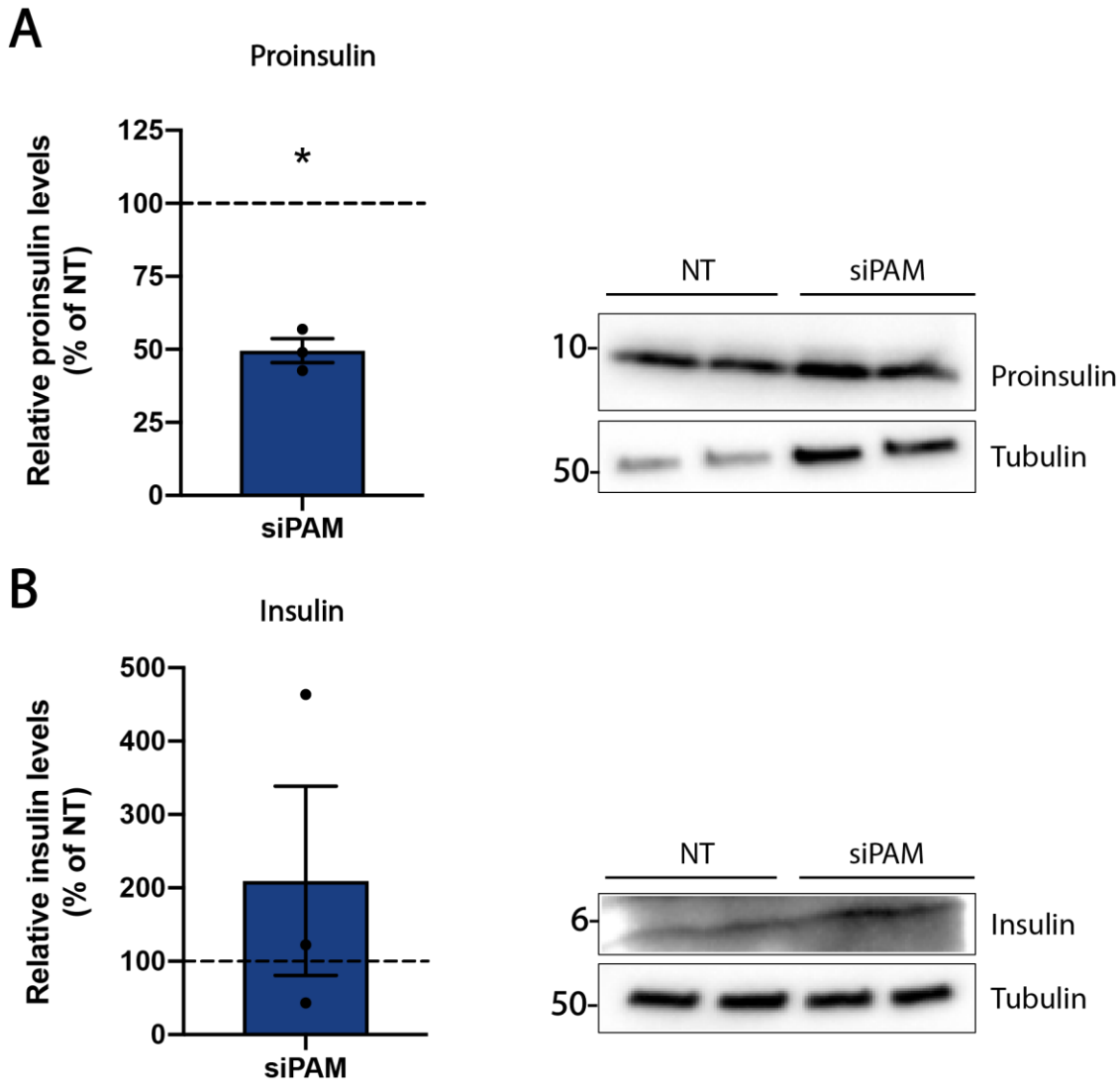

FIG. 3.3.4.2 *PAM* KD results in reduced proinsulin protein levels in EndoC-βH1 cells.
Cells were transfected with either non-targeting (NT) or siPAM oligos for 48hrs and then run on western blots for either proinsulin (A) or insulin (B). Bands were quantified using lane densitometry, normalised to NT-transfected data (dotted line) and analysed by two-tailed unpaired Student's t-test with Welch's correction on transformed (log) data. * denotes $P<0.05$.

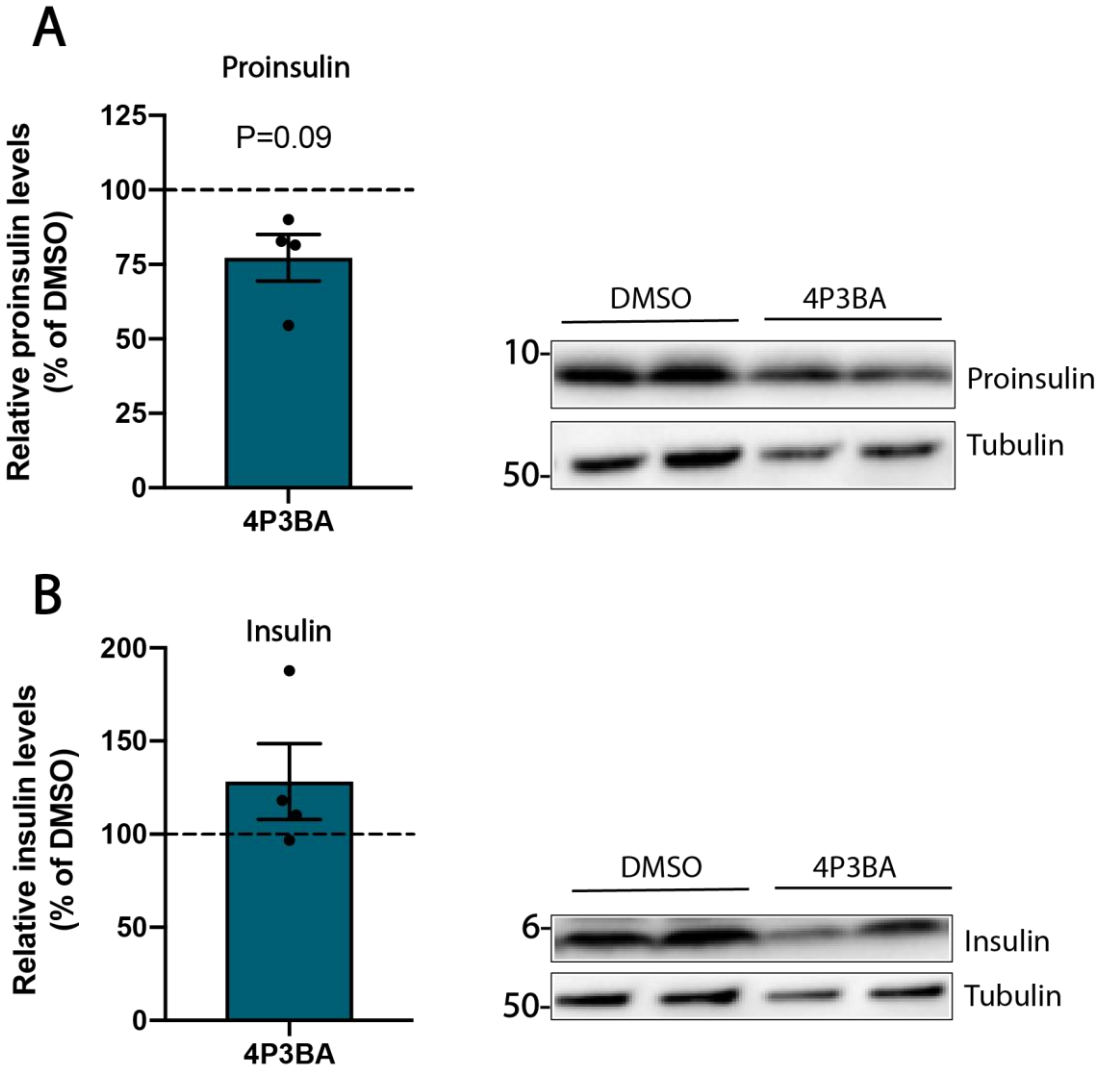

FIG. 3.3.4.3 Catalytic inhibition by 4P3BA results in reduced proinsulin protein levels in EndoC-βH1 cells. Cells were incubated with 500μM 4P3BA or DMSO and cell extracts examined by western blots for proinsulin (**A**) or insulin (**B**) abundance. Bands were quantified by lane densitometry, normalised to DMSO-treated cells (dotted line) and analysed by two-tailed unpaired Student's t-test with Welch's correction on transformed (log) data. Data are representative of 4 independent experiments and error bars are SEM.

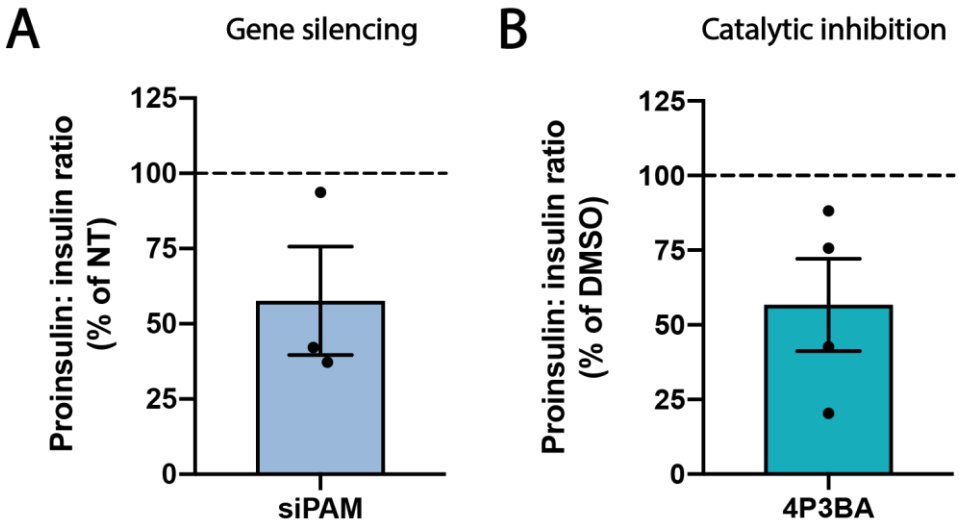

FIG. 3.3.4.4 Depletion of PAM function by either gene silencing (A) or catalytic inhibition (B) does not affect the proinsulin: insulin ratio in EndoC-βH1 cells. Relative abundance of proinsulin and mature insulin from FIG. 3.3.5.2 and FIG. 3.3.5.3 were collated to generated ratios of insulin processing (within experiments). Ratios were then normalised to control (NT for A, DMSO-treated for B) shown by the dotted line. Data are from 3-4 independent experiments and analysed by two-tailed unpaired Student's t-test with Welch's correction on transformed (log) data. Error bars are SEM.

3.3.5 Impact of PAM depletion on insulin biosynthesis and processing

I next questioned whether a reduction in proinsulin biosynthesis would result in a visible difference in the insulin content of large dense core vesicles (LDCVs). I first confirmed whether or not the proinsulin (C-peptide epitope) and insulin (B chain epitope) antibodies worked for immunofluorescence using PFA-fixed EndoC-βH1 cells cultured on slides. Cells were co-stained for PAM (C-terminal epitope), proinsulin (C-peptide epitope) and insulin (B chain epitope) and all three antibodies were successful for staining in EndoC-βH1 cells (FIG. 3.3.5.1). Proinsulin (α-C-peptide) largely co-localised to insulin-containing secretory granules, but there were also visible densely labelled regions of proinsulin only staining (white arrows, FIG. 3.3.5.1). This suggests

that the proinsulin antibody may detect both intact proinsulin and cleaved C-peptide fragments, which can be both found in mature granules.

In parallel, I performed immunogold labelling on EM sections using antibodies for either proinsulin (C-peptide epitope) or insulin (B-chain) on EndoC-βH1 cells transfected with either non-targeting or *siPAM* oligos. I observed good labelling of both antibodies to the secretory granules; there was a qualitative indication that C-peptide labelling was reduced in siPAM-transfected cells (FIG. 3.3.5.2).

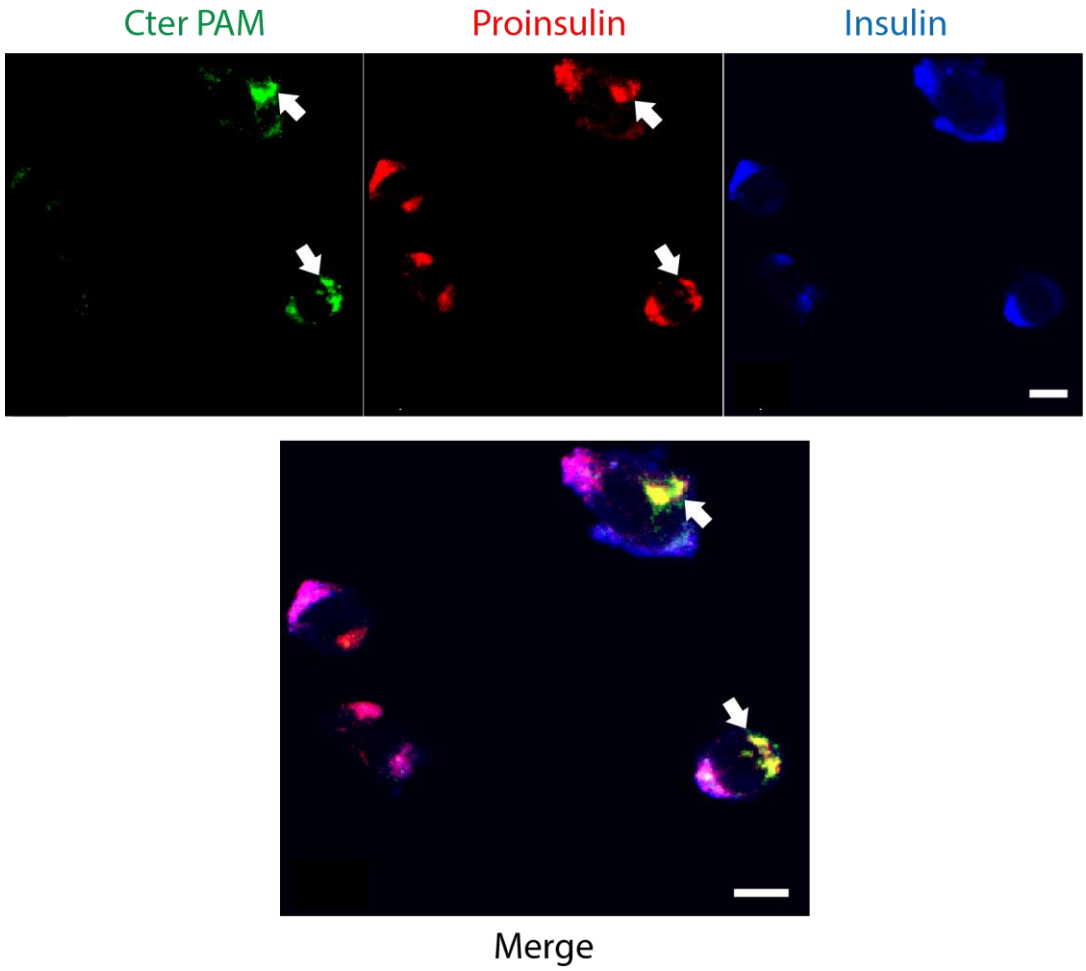

FIG. 3.3.5.1 Co-localisation of PAM and insulin species in EndoC-βH1 cells as assessed by immunofluorescence microscopy. Cells were triple immunostained for PAM (green), proinsulin (red) and insulin (blue) and imaged on a BioRad Radiance 2100 confocal microscope. White arrow indicates regions that are positive for proinsulin but negative for insulin. Scale bar indicates 10μm.

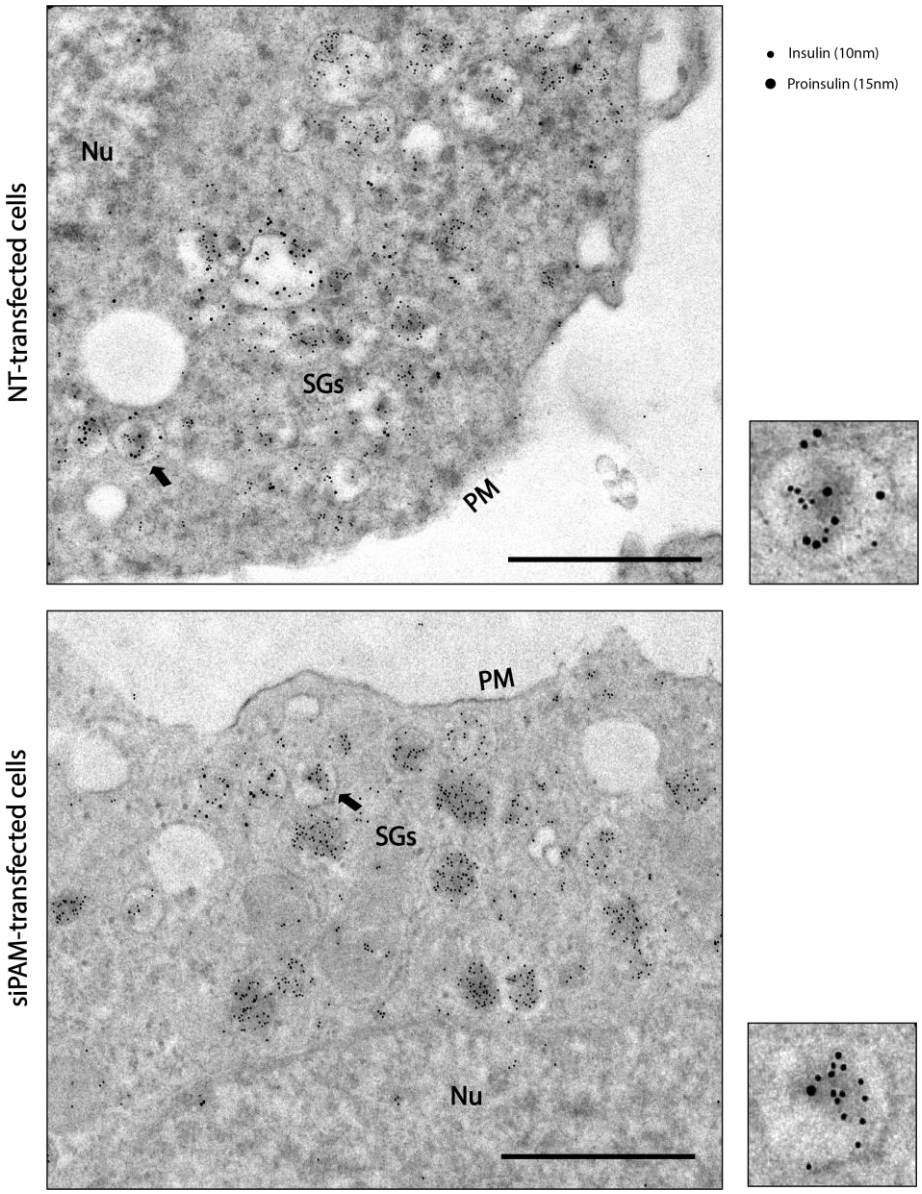

FIG. 3.3.5.2 Immunogold labelling of insulin species in NT-transfected or siPAM-transfected EndoC-βH1 cells visualised by transmission electron microscopy (TEM). Cells were co-labelled for C-peptide/proinsulin (15nm gold) and insulin (10nm gold) using immunogold labelling. Black arrow indicates position of granule enlarged in the panels on the right. NT: non-targeting; Nu: nucleus; SGs: secretory granules; PM: plasma membrane. Scale bar indicates 1μm.

I did not see an obvious difference in the maturity of labelled granules (mature granules are denoted by presence of a halo) containing either proinsulin or insulin between si*PAM*- and NT-transfected cells.

In conclusion, since distinguishing intact proinsulin from cleaved C-peptide fragments is not possible with immunostaining, I cannot conclude whether *PAM* KD has an effect on proportion of proinsulin and insulin present in packaged granules. As this is very preliminary work, quantification of gold particles for each antibody to elucidate the impact of PAM KD on granule insulin content was not pursued.

3.3.6 Chromogranin A is a likely mediator of PAM function

I have shown that PAM's amidating activity contributes to its function in maintaining insulin content through proinsulin biosynthesis, but the precise mediators of this in the β cell are undescribed. I hypothesised that there may be PAM substrates, which are co-expressed and co-localised in insulin-containing secretory granules that may be amidated and indirectly affect insulin secretion. One such candidate is chromogranin A (CgA), a critical component of the secretory granule, which is implicated in the biogenesis of secretory granules and terminates in a C-terminal glycine – making it a putative substrate for PAM catalysis (Kim, Tao-Cheng et al. 2001). Its derivatives, such as pancreastatin (PST) and serpinin have been shown to influence insulin secretion in the β cell (Ahren, Lindskog et al. 1988, Koshimizu, Cawley et al. 2011).

To test this, I first performed immunofluorescence on fixed EndoC-βH1 cells to confirm whether chromogranin A co-localised to insulin-containing secretory granules in

EndoC-βH1 cells, as we had observed with PAM. I was unable to perform simultaneous co-staining of both PAM and CgA since both of the primary antibodies are raised in rabbit but was able to co-label CgA with insulin. I observed co-staining of CgA to insulin-containing granules; typically in EndoC-βH1 cells the granule density is highest at the periphery of the cell in processes (FIG. 3.3.6.1). This is similar to co-localisation observed between insulin and PAM (FIG. 3.3.3.1).

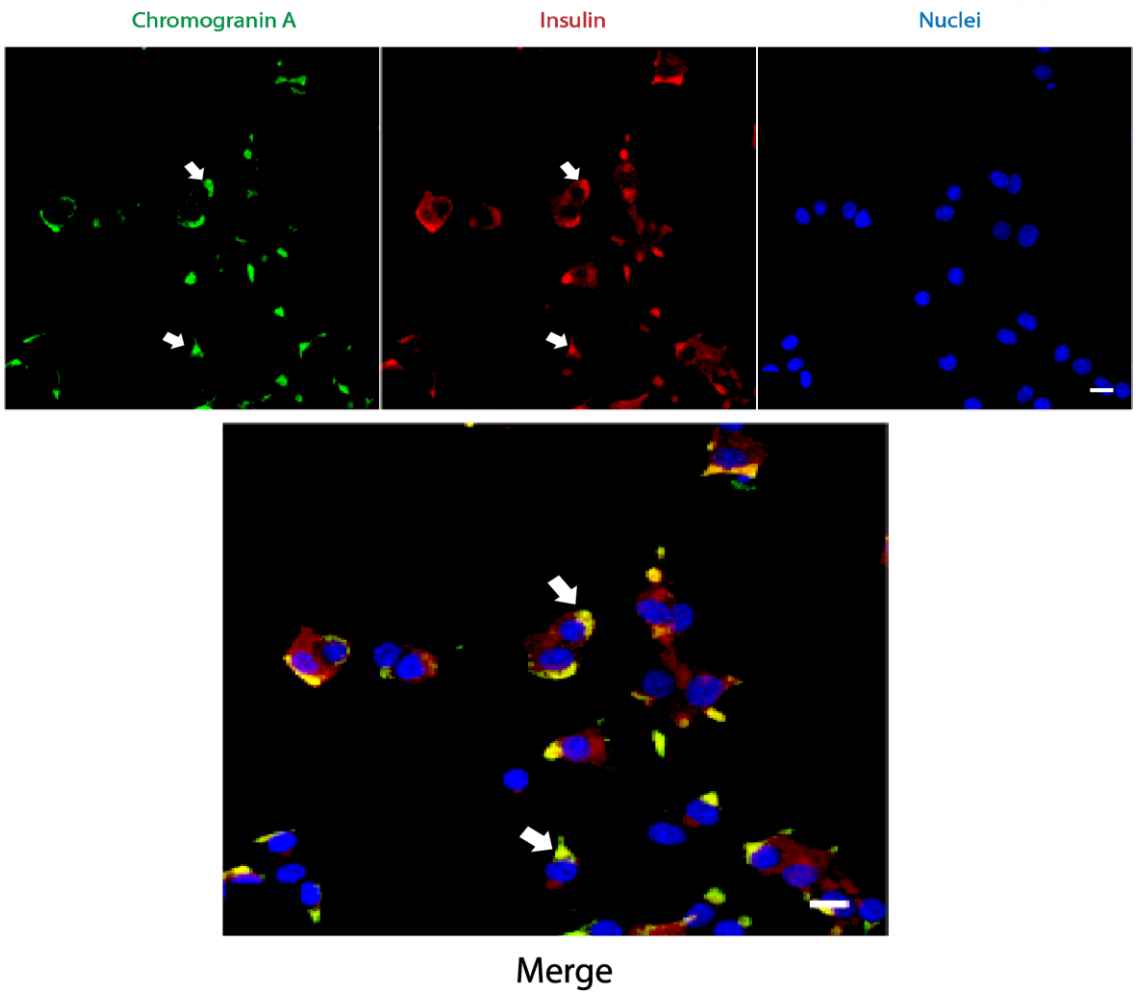

FIG. 3.3.6.1 Chromogranin A localises to insulin-containing granules at the periphery of EndoC-βH1 cells as assessed by immunofluorescence microscopy. Cells were double immunostained for chromogranin A (green), insulin (red) and red dot2 (nuclei, blue) and imaged on a BioRad Radiance 2100 confocal microscope. White arrows indicate regions of dense co-localisation of chromogranin A. Scale bars denote 10μm.

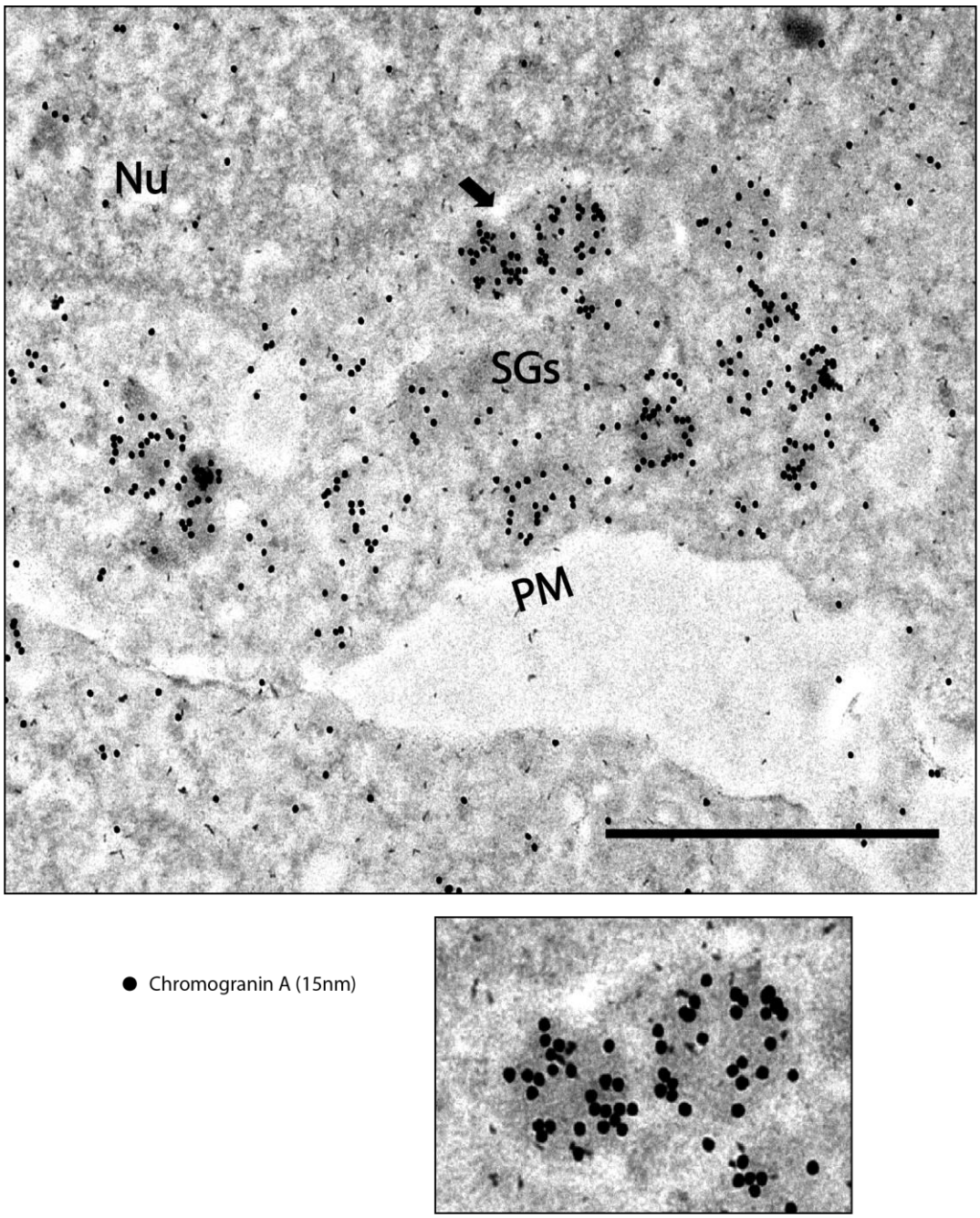

FIG. 3.3.6.2 Immunogold labelling of chromogranin A in EndoC-βH1 cells. Cells were immunogold labelled for chromogranin A (15nm gold) and imaged using transmission electron microscopy. Black arrow indicates position of granules shown in lower panel. SGs: secretory granules, PM: plasma membrane; Nu: nucleus. Scale bar denotes 1μm.

I was also able to observe this in electron micrographs of fixed EndoC-βH1 cells that had been immunogold labelled for chromogranin A (FIG. 3.3.6.2); this antibody had high background labelling but the density of labelling was highest in secretory granules.

Following this, I wanted to confirm whether or not CgA was indeed a substrate for PAM catalysis in the β cell. To do this, EndoC-βH1 cells were transfected with either non-targeting (NT) or si*PAM* and the supernatant and cells extracts examined for either total CgA or the glycine-extended precursor (to amidation) of CgA. The measure for total CgA was based on cumulative quantification of the major bands at ~50kDa.

I observed that silencing of *PAM* resulted in a trend towards the accumulation of chromogranin A precursor (CgA-gly). Abundance of CgA-gly provides an inverse measure of amidated CgA levels and confirms that chromogranin A is indeed amidated by PAM. This effect is visible in both cell extract (+336%, n=3, P=0.07) and supernatant (+63%, n=3, P=0.06) (FIG. 3.3.6.3).

When EndoC-βH1 cells were treated with an irreversible catalytic inhibitor (4P3BA) of PAM for 48hrs, the abundance of CgA precursor (CgA-gly) was elevated in the cell extract (+281%, n=3, P=0.02). However, there was no difference observed in the supernatant (P=0.3) (FIG. 3.3.6.4).

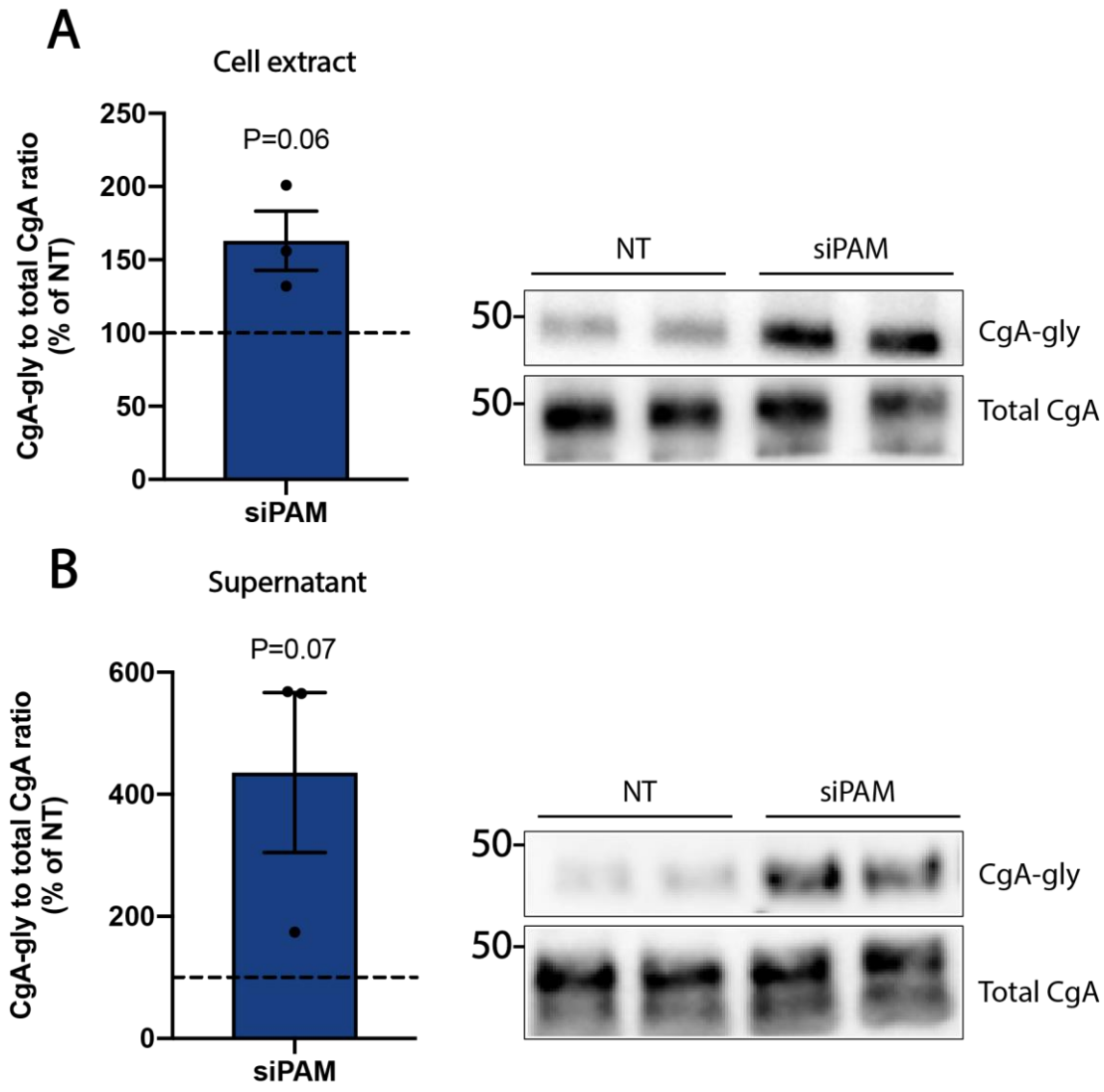

FIG. 3.3.6.3 Silencing of *PAM* results in accumulation of CgA-gly (precursor) in EndoC-βH1 cells. Cells were incubated with either non-targeting (NT) or siPAM oligos for 48hrs in BSA-free media and then cell extract and supernatant were run on western blots to measure protein abundance of CgA-gly (precursor) and total CgA. Blots were quantified by lane densitometry and plotted as CgA-gly: total CgA ratios normalised to NT-transfected cells (dotted line). Differences were analysed by two-tailed unpaired Student's t-test with Welch's correction on transformed (log) data and error bars are SEM.

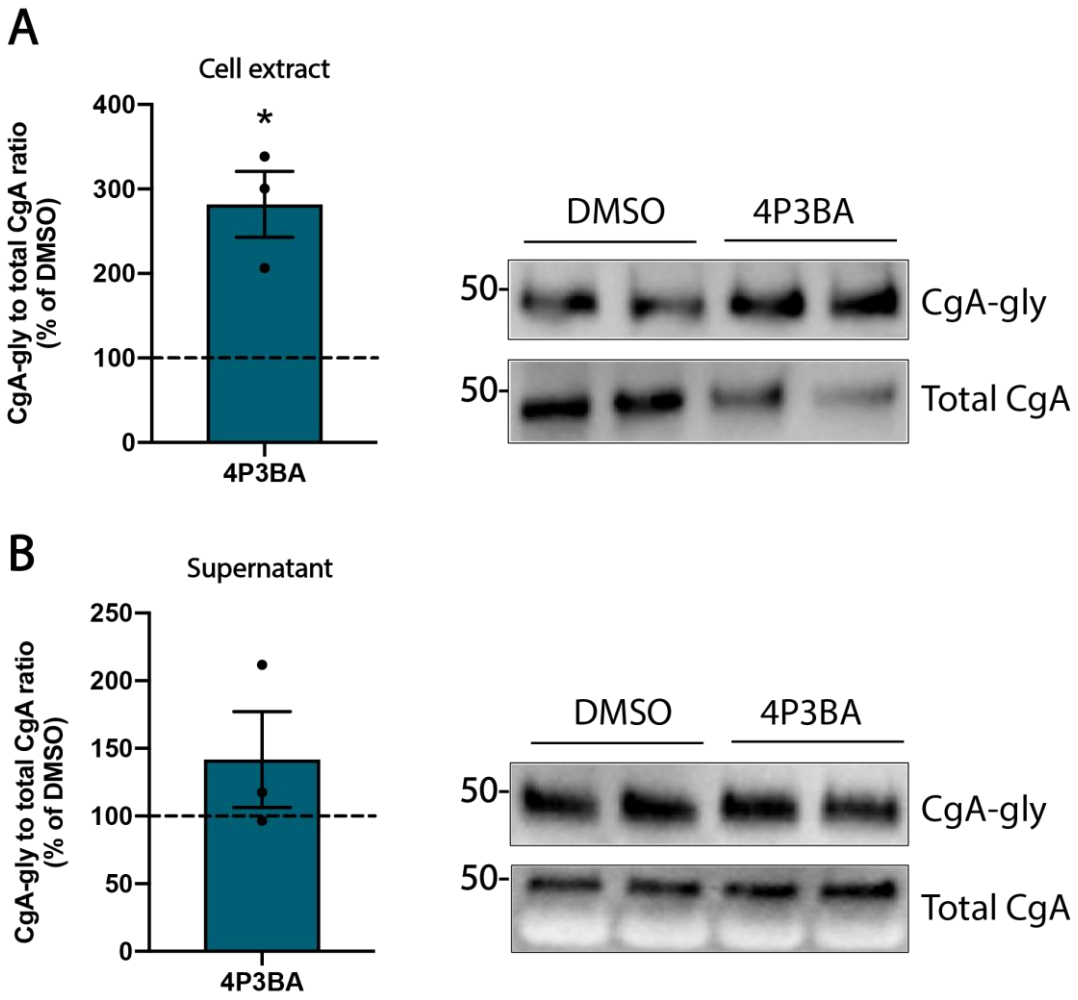

FIG. 3.3.6.4 Catalytic inhibition of PAM by 4P3BA results in accumulation of CgA-gly (precursor) in EndoC-βH1 cells. Cells were incubated with either 500μM 4P3BA or DMSO were examined for either total CgA or CgA-gly precursor in cell extract (**A**) and supernatant (**B**). Bands were quantified by lane densitometry and plotted as CgA-gly: total CgA ratios normalised to DMSO-treated cells (dotted line). Differences analysed by two-tailed unpaired Student's t-test with Welch's correction on transformed (log) data, error bars are SEM. * denotes P<0.05.

A minor observation from these studies was that whilst there was a clear accumulation of CgA-gly precursor molecules (relative to the total CgA protein levels) following gene silencing or catalytic inhibition, there also appeared to be a modest reduction in the total CgA levels observed (FIG. 3.3.6.4(A)). Examination by qRT-PCR demonstrated a reduction in *CHGA* mRNA expression (-29.5%, n=3, P=0.02) following *PAM* KD (FIG. 3.3.6.5).

This observation may indicate a feedback loop between PAM carboxyamidation activity, (pro)insulin expression and chromogranin A expression and processing, which is important for normal secretory pathway function, but this was not characterised further.

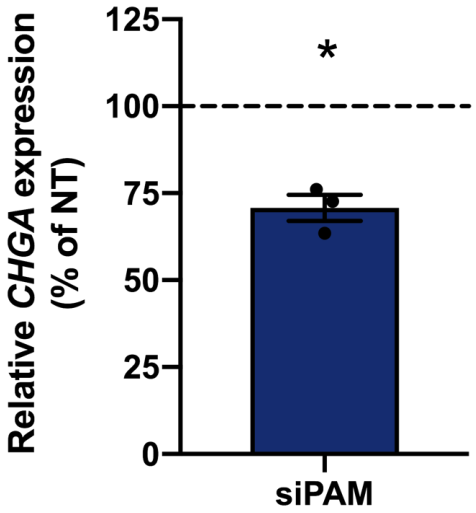

FIG. 3.3.6.5 *CHGA* mRNA expression is reduced following PAM KD in EndoC-βH1 cells. Cells were transfected with non-targeting (NT) or siPAM oligos for 48hrs and then collected for RNA. *CHGA* expression was assessed by qRT-PCR and is plotted as percentages of NT-transfected cells (dotted line). Data is representative of 3 independent experiments and analysed by two-tailed unpaired Student's t-test with Welch's correction on transformed (log) data. Error bars are SEM and * denotes P<0.05.

3.3.7 Impaired amidation of chromogranin A affects its localisation to the secretory granule

Given the implied role for chromogranin A in secretory granule biogenesis and the co-localisation with PAM in insulin-containing secretory granules, I hypothesized that impairment in the amidation of chromogranin A may result in altered trafficking to the secretory pathway of human pancreatic β cells. This may have a subsequent impact on protein interactions with other components of the developing granule and thus impede the correct formation of insulin-containing secretory granules at the level of the Golgi. Amidation has been proposed to improve electrostatic interactions and therefore may be important in favouring condensation of granule proteins at the TGN (Shalev, Mor et al. 2002, Sforca, Oyama et al. 2004). Previous reports of interactions of chromogranin A with other granule components (e.g. Secretogranin III and cholesterol) suggests that these interactions may be important in initiating granule biogenesis (Hosaka, Watanabe et al. 2002, Hosaka, Suda et al. 2004, Koshimizu, Kim et al. 2010). This, in combination with the reduced *INS* expression observed may contribute to the reduced insulin content observed in PAM-depleted β cells and in carriers of T2D-associated variants in *PAM.*

To investigate this, I performed immunofluorescence on CRISPR PAM KO EndoC-βH1 cells, which lack catalytically active PAM (details describing the generation of this cell line is in section 2.5 and functional validation provided in section 3.3.1). This provides a relevant model to examine whether impairment of PAM catalysis results in altered trafficking of chromogranin A. In the PAM KO cells, I observed a deviation in the staining pattern for chromogranin A, with regions of densely stained chromogranin A

that did not co-localise with insulin (white arrows, FIG. 3.3.7.1). This is different to the staining pattern for chromogranin A observed in control EndoC-βH1 cells previously, despite using the same antibody detecting total CgA, where there was complete co-staining of chromogranin A and insulin in processes located near the cell periphery (FIG. 3.3.6.1).

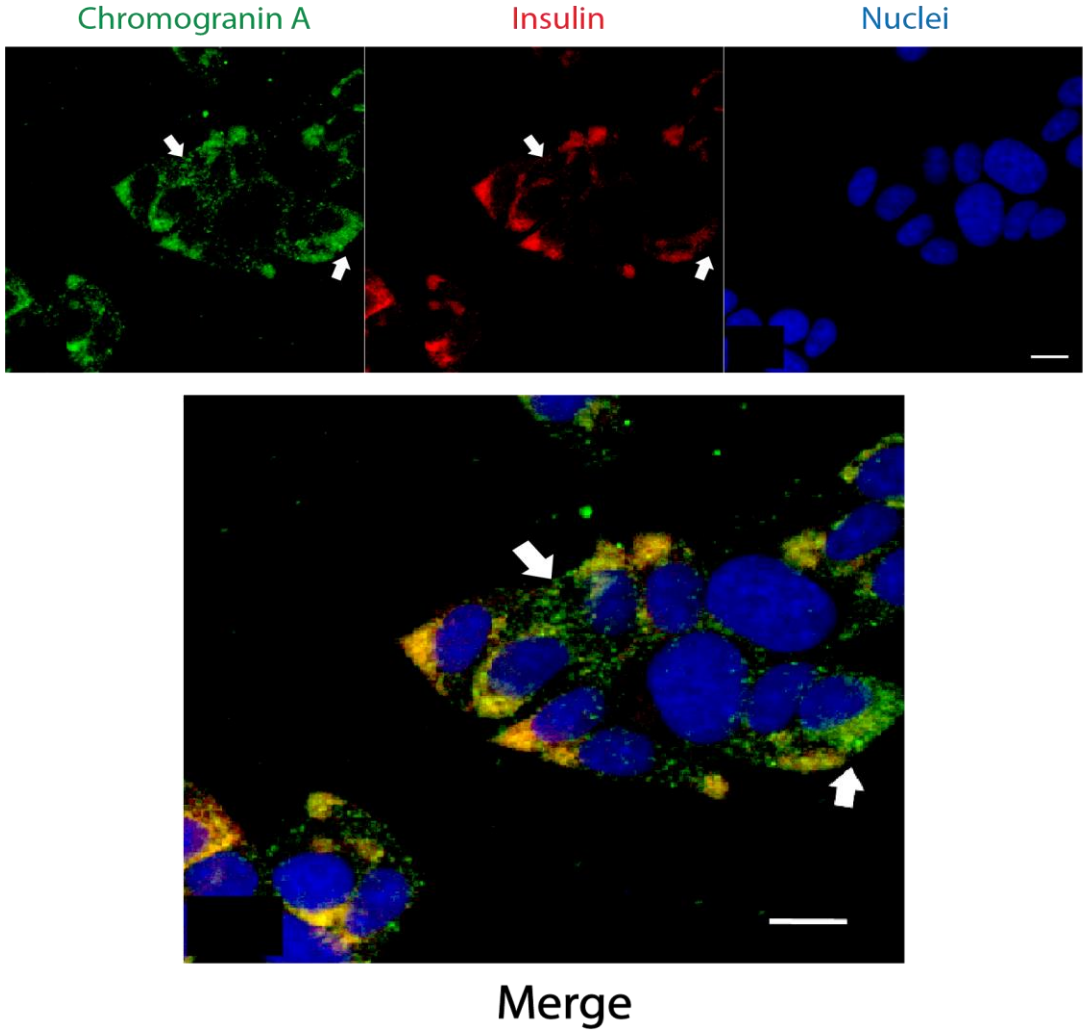

FIG. 3.3.7.1 Localisation of chromogranin A is altered in CRISPR PAM KO EndoC-βH1 cells. Cells were double immunostained with Chromogranin A (green), insulin (red) and nuclei stained using red dot2 (blue). Cells were imaged with a BioRad Radiance 2100 confocal microscope. White arrows indicate regions where there are regions of dense chromogranin A staining but no co-localisation with insulin. Scale bars indicate 10μm.

3.4 DISCUSSION

In this chapter, I have described the tissue-specific biosynthesis and processing of PAM, and the processes by which PAM function may play a role in insulin secretion. I used EndoC-βH1 cells as my model of human β cells, which were generated from human foetal pancreatic buds and so may lack some functional maturity (Ravassard, Hazhouz et al. 2011). It is therefore important to consider that my observations may differ from the mechanisms observed in human islets or in human carriers of T2D-associated variants in *PAM*.

Previous characterisation of PAM biosynthesis and function has been carried out largely in non-islet tissues, namely in the anterior pituitary (Eipper, Mains et al. 1983, Eipper, Milgram et al. 1993, Kumar, Mains et al. 2016). Using PacBio sequencing, I observed that the predominant mRNA isoform in human pancreatic islets encodes a bifunctional membrane integral isoform and as such contains all functional domains, including both catalytic domains (PHM and PAL), making it capable of performing complete carboxyamidation. In addition to splicing isoforms, post-translational endoproteolytic cleavage can generate additional complexity in the types of protein isoforms present in the cell (Eipper, Green et al. 1992). The types of protein isoforms present will dictate the functional capacity for PAM and mediate differences in cellular function between tissues (Thiele, Marek et al. 1989, Eipper, Green et al. 1992).

Exploration of the protein isoforms by western blotting identified 3 major PAM isoforms (120kDa, 75kDa, 50kDa) present in the human pancreatic β cell (FIG. 3.3.2.1), the largest of which corresponds to the newly translated intact PAM zymogen. The 2

remaining isoforms constitute post-translationally cleaved derivatives, which are likely to carry distinct functions important for β cell function. These are as yet unknown, but would provide useful information on the additional mechanisms in which PAM functions in the β cell, especially in hormone storage and secretion. Comparison of size and position of dibasic cleavage sites in PAM suggest that the 75kDa form is likely to comprise an untethered (luminal) isoform, which would be trafficked to the lumen of insulin secretory granule and likely be co-secreted with insulin. Unfortunately, I could not observe this isoform in the supernatant of EndoC-βH1 cells, but did see a higher molecular weight band at >250kDa (FIG.3.3.2.3); this could be a complex containing PAM bound to another granule peptide or a form of glycosylated PAM. Standard SDS-PAGE protocols are unlikely to be sufficient to solubilise PAM complexes. Using a similar approach, the 50kDa isoform is likely to be the membrane-integral monofunctional PAM isoform containing only PAL activity. These predicted isoforms are deduced from using one antibody, confirmation would require additional antibodies targeting epitopes in different regions of PAM or with mass spectrometry. There was no change in the relative abundance of these 3 isoforms (as a % of total PAM) following stimulation by high glucose, suggesting that PAM function is independent of relative isoform expression. Additional isoforms lacking the PAL domain may also exist or be generated through cleavage, but without additional antibodies these cannot be detected.

I observed that PAM co-localised with insulin-containing granules in processes at the cell periphery using both immunofluorescence and transmission electron microscopy using an antibody targeting the C-terminus of PAM. This is consistent with previous

reports of its localisation in secretory granules in other neuroendocrine cells (Martinez, Burrell et al. 1993, Martinez, Montuenga et al. 1993). These data suggest that PAM follows a similar trafficking pathway to insulin, starting at the ER where its N-terminal ER signal peptide is cleaved and post-translational processing such as N-linked glycosylation may occur (Milgram, Johnson et al. 1992). PAM would then be transported to the TGN, where it is packaged into insulin-containing secretory granules along with other key granule proteins such as chromogranin A. Immature secretory granules undergo acidification, which permits proteolytic cleavage of proinsulin by PC1/2 and PC2 and post-translational cleavage of PAM. This is likely mediated by γ-secretase and PC1/3 (given the abundance of dibasic motifs), thus generating the 2 smaller PAM isoforms (75kDa and 50kDa) observed in the cell extract of EndoC-βH1 cells.

Large dense core vesicles (LDCVs) would also be the site of carboxyamidation of granular components (e.g. chromogranin A and islet amyloid polypeptide, IAPP), for which a pH 5.5 is required. Understanding the full trafficking pattern of PAM isoforms requires additional antibodies. In addition, use of fusion proteins with a fluorophore at different positions throughout the PAM molecule would help to elucidate this through both fixed- and live-cell imaging techniques.

Having established the tissue-specific expression of *PAM* in human β cells, I next focused on the mechanisms surrounding the effects of PAM on insulin secretion and content in human pancreatic β cells. Insulin content was shown to be reduced following *PAM* knock-down (KD), exhibiting as a reduction in both total intracellular

insulin content and the number of insulin particles per granule (Thomsen, Raimondo et al. 2018). Given this, I first investigated the impact of *PAM* KD on the expression and abundance of insulin species in the β cell. Here, I identified a reduction in both proinsulin mRNA expression and protein abundance. Treatment of cells with 4P3BA, an irreversible inhibitor of amidating activity, displayed a similar reduction in proinsulin abundance, suggesting that the carboxyamidation process is at least partially responsible for maintaining normal insulin content. Further characterisation through high throughput methods such as RNA-seq may help to elucidate the intermediate processes between impaired PAM function and reduced proinsulin biosynthesis, which culminate in impaired insulin content (Alam, Caldwell et al. 1996, Alam, Steveson et al. 2001).

Given that insulin is not amidated by PAM, I hypothesised that an alternative mechanism by which PAM may function in insulin secretion, is through the amidation of auxiliary proteins that would be co-expressed with PAM. Chromogranin A (CgA) is implicated in the biogenesis of insulin-containing secretory granules and contains a C-terminal glycine, making it putative a substrate for PAM. Similar to PAM, I observed co-localisation of CgA with insulin in the granules. I observed evidence that CgA is amidated in the β cells, since there was an accumulation of CgA-gly (the conjugate precursor to amidated CgA) following *PAM* KD.

I postulated that impaired amidation of CgA may result in instability that could alter its trafficking and therefore result in a reduction in the packaging and storage of insulin-containing granules. In CRISPR PAM KO EndoC-βH1 cells, where catalytically-active

PAM is lacking, I observed a deviation in the localisation pattern of chromogranin A, which normally completely co-localises with insulin. I observed regions of punctate granular-like staining of CgA with no co-staining for insulin. Further characterisation is necessary to identify whether these are dysfunctional granules (e.g. granules lacking insulin cargo or with very little insulin) or alternative cellular compartments (e.g. lysosomes) to which chromogranin A may be trafficked following PAM depletion. Due to the high background of chromogranin A immunogold labelling in EM sections, I was unable to confirm whether impaired PAM function (amidating activity) resulted in aberrant trafficking of chromogranin A to LDCVs or lysosomes.

Additional mechanisms may be mediated through other proteins, which are amidated or interact with PAM, such as cleaved derivatives of chromogranin A. Whilst I have focused on characterising chromogranin A as an entity, this may not provide the best resolution in deciphering processes involved in granule biogenesis and insulin secretion. There have been reports of cleaved fragments (e.g. pancreastatin and serpinin) having roles on insulin secretion, but the biology in a PAM-dependent context is undescribed (Ahren, Lindskog et al. 1988, Koshimizu, Cawley et al. 2011, Wollam, Mahata et al. 2017).

The processes by which PAM functions in the β cell are complex and difficult to interpret. Much appears to be centred in the biogenesis and storage of insulin-containing secretory granules, and appears to be driven at least partially through its established function in the carboxyamidation of glycine-extended peptides. Future work should focus on deciphering the tissue-specific mechanisms by which PAM

depletion may impede this, which may lead to elevated T2D risk. Efforts should also be placed to understand the non-catalytic roles for PAM.

CHAPTER 4

Characterisation of R36S-PAM, a novel *de novo* candidate causal variant identified in a patient with early onset diabetes mellitus (NDM)

This work was presented as an oral abstract presentation at the ADA 79th Scientific Session in San Francisco on the 7th June 2019.

A manuscript is currently being prepared.

4.1 INTRODUCTION

4.1.1 Diabetes: a continuum of allelic frequency and clinical severity

Several genes implicated in monogenic diabetes (e.g. *HNF1A*) have been shown to harbour variants identified in association with T2D risk (Stanescu, Hughes et al. 2012, Estrada, Aukrust et al. 2014). This overlap in genetic aetiology, exemplifies the relationship between allelic frequency and severity of clinical presentation present in diabetes.

Monogenic diabetes, which encompasses both neonatal diabetes mellitus (NDM) and maturity onset diabetes of the young (MODY), is caused by mutations in a single gene mediating pathogenesis (Hattersley, Bruining et al. 2006). NDM has been defined by a clinical onset before 6 months of age, during which time the likelihood of T1D is rare (Aguilar-Bryan, Clement et al. 1998, Edghill, Dix et al. 2006, Polak and Cavé 2007, Hattersley, Bruining et al. 2009). Recently, cases for NDM have also been identified in individuals presenting at < 1 year of age, and as such early onset diabetes mellitus may, in some cases, be a more appropriate term for the condition (Rubio-Cabezas, Flanagan et al. 2012). Autosomal dominant cases are commonly caused by mutations in *KCNJ11* and *ABCC8*, which together contribute 40-60% of cases (Hattersley and Ashcroft 2005, Flanagan, Edghill et al. 2006).

Maturity onset of the young (MODY) is typically but not exclusively, diagnosed later in childhood (<25 years of age) and is inherited in an autosomal dominant manner (Ellard, Bellanne-Chantelot et al. 2008). The majority of cases for MODY can be

attributed to variants in *GCK* and *HNF1A,* which each constitute 30-60% of all MODY cases (Fajans, Bell et al. 2001, Stride, Ellard et al. 2005).

Currently, causal mechanisms underlying monogenic diabetes pathogenesis can be segregated into 3 main groups: impaired pancreatic and/or β cell development (e.g. *GATA6, PTF1A*), defects in β cell function (e.g. *KCNJ11*) and β cell destruction (e.g. *INS*, *EIF2AK3*), some of which can be non-T1D autoimmune (e.g. *STAT3*) (Gloyn, Pearson et al. 2004, Edghill, Flanagan et al. 2008, Allen, Flanagan et al. 2011, Flanagan, Haapaniemi et al. 2014, Houghton, Swift et al. 2016). Both MODY and early onset diabetes can present as either isolated diabetes or as part of a syndrome with extra-pancreatic clinical features (e.g. for *KCNJ11* and *HNF1B*) (Bingham, Bulman et al. 2001, Gloyn, Diatloff-Zito et al. 2006, Rubio-Cabezas and Ellard 2013).

Standard insulin therapy is invasive and does not combat the underlying pathophysiology (Stumvoll, Goldstein et al. 2005). Knowledge on the genetic aetiology and causal mechanisms underlying disease can inform on the clinical management for disease. Early onset diabetes mellitus caused by mutations in K_{ATP}-channels are treated with high dose sulfonylureas rather than insulin, thereby exemplifying the clinical benefit of understanding the genetic and cellular mechanisms underlying disease and the application of precision medicine (Pearson, Flechtner et al. 2006, Rafiq, Flanagan et al. 2008).

4.1.2 Utilising rare variants to uncover biological function

Common variants associated with polygenic traits, such as those in T2D risk, often carry modest or small effect sizes, thus making the elucidation of biological function for the associated loci difficult to uncover (Hemminki, Försti et al. 2008). Rare variants associated with monogenic disease identified in known T2D risk loci, can help to characterise the function of that gene, as rare alleles are generally more penetrant and are likely to severely impact protein function and/or cellular physiology (Bonnefond and Froguel 2015). Identification of rare variants in known T2D susceptibility loci can expose the causal mechanism supporting the monogenic disease-association in that locus and provide insights into the cellular mechanisms disrupted for T2D risk alleles (e.g. *KCNJ11*) (Gloyn, Weedon et al. 2003, Sandhu, Weedon et al. 2007, Wegner, Andersen et al. 2007).

4.1.3 Risk alleles in *PAM* provide causal link between impaired PAM function and disease risk

Functional characterisation of causal coding T2D-risk alleles in *PAM*, such as D563G and S539W, provide a direct causal link between impaired PAM function in the β cell and increased T2D risk. The low frequency T2D-risk allele, S539W (MAF=0.65%), resulted in altered stability of the untethered luminal isoform and ER-retention of the membrane-integral isoform. Despite these stark cellular phenotypes, S539W is only associated with an increase in T2D risk (OR=1.47 in Europeans) but not neonatal or perinatal diabetes (Huyghe, Jackson et al. 2013, Steinthorsdottir, Thorleifsson et al. 2014, Thomsen, Raimondo et al. 2018). In addition, *PAM* gene silencing (-79% at *PAM* mRNA level) did not eliminate insulin secretion or drastically impair β-cell function

(Thomsen, Raimondo et al. 2018). This suggests that impaired carboxyamidation alone is not sufficient to cause monogenic diabetes. Therefore, a rare allele identified in *PAM*, which is casual for monogenic diabetes, is likely to be working through alternative or additional mechanisms to those implicated for T2D risk.

4.1.4 Experimental aims

In this chapter, I aim to characterise the functional impact of a candidate causal *de novo* variant in *PAM* identified by collaborators in a patient with early onset diabetes mellitus. Characterisation of known T2D-associated alleles have established the link between PAM function and insulin secretion, but characterisation of this rare allele will provide a greater insight into PAM and the cellular dysfunction underlying diabetes risk. To achieve this, I examined the following experimental avenues for R36S-PAM in EndoC-βH1 cells:

- Catalytic amidating activity (*in vitro* and *ex vivo*)
- Amidation of chromogranin A
- The impact on PAM abundance and protein stability of both the untethered (luminal) and membrane-integral isoforms
- Cellular localisation
- Impact on the unfolded protein response (UPR) and consequential pathological ER stress
- β cell proliferation and survival
- Insulin secretion and content

I have included S539W-PAM as a T2D-associated reference allele to allow distinctions between T2D- and monogenic diabetes-associated cellular mechanisms to be made.

4.2 METHODS

For full methods see chapter 2, details pertaining to this chapter are summarised below. Unless stated otherwise, cellular characterisation was performed in CRISPR *PAM* KO EndoC-βH1 cells.

4.2.1 Generation of constructs

Constructs encompassing either luminal or membrane-integral R36S-PAM were generated through site directed mutagenesis of the WT-PAM constructs (section 2.4) in collaboration with Dr. Anne Raimondo (OCDEM, University of Oxford). Primers used for the generation of R36S-PAM are listed in Table 4.2.1.1. Constructs for S539W and D563G alleles were provided by Dr. Anne Raimondo. Sanger sequencing was performed to ensure that only the desired nucleotide changes were introduced.

Variant (p.)	Primer sequence (5' to 3')
R36S	GGTTTAAAGAAACTACCAG**T**CCATTTTCCAATGAATG

Table 4.2.1.1 Primer sequences to generate variant PAM constructs using site-directed mutagenesis. Primers were purchased from Eurofins Genomics. SDM was performed on a pCMV6 constructs encoding either human membrane-integral WT-PAM (NM_138821.2) or luminal WT-PAM. The latter construct had been generated through PCR cloning of the 1-710 amino acids by Dr Anne Raimondo previously. Nucleotide change is highlighted in red.

4.2.2 Culture and Transfection of CRISPR *PAM* KO EndoC-βH1 cells

The generation of this cell line is covered in section 2.5 and the details on how cells were cultured in section 2.6. To maintain selection pressure, cell media was routinely supplemented with puromycin (4µg/µl every 4-6 weeks, Sigma Aldrich). Functional validation is described in section 3.3.1. Cells were forward transfected 24 hours post-seeding, using FuGene 6 (Promega) (section 2.7). Unless otherwise stated, cells were incubated with transfection complexes for 72hr at 37°C/5% CO_2 prior to harvest by trypsination (as per section 2.9). Constructs used for transfection are summarised in Table 4.2.2.1. Cells were subject to lysis in RIPA buffer (Table 2.9.2.1) for protein extraction and lysates were stored at -80°C prior to western blotting (see section 2.9).

Vector	Insert	Genotype
pCMV6-Entry	GFP	-
pCMV6-Entry	Membrane-integral PAM	WT
pCMV6-Entry	Membrane-integral PAM	S539W
pCMV6-Entry	Membrane-integral PAM	R36S
pCMV6-Entry	Luminal PAM	WT
pCMV6-Entry	Luminal PAM	R36S

Table 4.2.2.1 Constructs used to transfect (CRISPR PAM KO) EndoC-βH1 cells. All pcDNA were quantified on the Nanodrop prior to transfection. All PAM constructs were C-terminally Myc- and FLAG-tagged.

4.2.3 Western blotting analysis

For details on quantification and running of SDS-PAGE and western blots see section 2.9. Blots were blocked and all antibodies were diluted in 3% BSA/TBST. Primary antibodies used to detect transfected PAM are provided in Table 4.2.3.1, ER stress/pro-apoptotic markers are in Table 4.2.3.2. Tubulin detection was used as a loading control. Secondary antibodies used are listed in Table 2.9.4.2. Bands were quantified using lane densitometry (Image Lab 5.0, BioRad) and adjusted for surrounding background signal. They were then adjusted for tubulin band intensities (as a loading control) and were normalised to the WT-transfected or DMSO-treated cells (as indicated in individual figure legends) within each experiment. Mean values across experiments were analysed in GraphPad Prism 8.0 (GraphPad Software, San Diego, California).

Target	Species	Specificity	Source	Dilution (WB)	Dilution (IF)
Myc tag	Mouse	Clone 4A6	Merck Millipore, 05-724	1:10,000	-
FLAG tag	Mouse	Clone M2	Sigma Aldrich, F1804	-	1:100
PAM (Nter)	Mouse	PAL domain	Santa Cruz, sc-514110	1:1000	-
Insulin	Guinea pig	B chain	Dako, A0564	-	1:750
TGN46	Rabbit	*Trans*-Golgi network	Invitrogen, PA5-23068	-	1:300
Tubulin	Mouse	Loading control	Santa Cruz, sc-365791	1:2000	-

Table 4.2.3.1 Primary antibodies used to detect exogenous PAM (upper panel) and for markers of cellular compartments (lower panels). Dilutions for western blot (WB) were made up in 3% bovine serum albumin in Tris-buffered saline with tween (BSA/TBST). For immunofluorescence (IF) dilutions were made up in Tris buffer. Cter: C terminus; Nter: N terminus.

Target	Source	Dilution
CHOP	Abcam, ab179823	1:1000
Total IRE1	Abcam, ab37073	1:1000
Phospho-IRE1	Abcam, ab48187	1:1000
PERK	Cell Signalling, 3192	1:1000
ASK1	Cell Signalling, 3762	1:1000
TRAF2	Cell Signalling, 4724	1:1000
Phospho-JNK	Cell Signalling, 9251	1:1000

Table 4.2.3.2 Antibodies used to detect markers of the unfolded protein response (UPR) and downstream pro-apoptotic signalling. Antibodies were used for western blotting analysis and all dilutions were made up in 3% BSA/TBST unless stated otherwise.

4.2.4 Protein degradation assays

MG132 (Sigma Aldrich), which is an inhibitor of the 26S subunit of the proteasome, was used to investigate the impact of proteasomal degradation on abundance of luminal R36S-PAM in EndoC-βH1 cells. Cells were transfected with luminal PAM as per section 2.7. Cells were then treated with either 10μM MG132 or DMSO for 2hr prior to harvest by trypsinisation and examination of PAM abundance using western blotting analysis (section 2.9).

4.2.5 Immunofluorescence

EndoC-βH1 cells were immunostained following fixation (section 2.11). Primary antibodies used are listed in table 4.2.3, secondary antibodies in table 2.11.1.2. Nuclei were stained with a far-red stain called Red Dot2 (Biotium). Slides were visualised using a BioRad Radiance 2100 confocal microscope, images exported using the Zeiss LSM software and arranged using Adobe Illustrator (Adobe Systems).

4.2.6 Transmission electron microscopy (TEM)

Transmission electron microscopy was performed under the guidance of Dr. Anne Clark at OCDEM and at the Department of Physiology, Anatomy and Genetics (DPAG). In brief, EndoC-βH1 cells for immunogold labelling for TEM were pelleted by trypsination (as per section 2.6) and processed as described in section 2.12. Immunogold labelling was performed using primary antibodies listed in table 3.2.3.1, with amplification using biotin/streptavidin conjugated gold particles for PAM (C-terminal) only.

4.2.7 Amidation assays

See chapter 2 for full details on the *in vitro* (section 2.13) and serum (2.14) amidation assays, including the generation of recombinant PAM for the former assay through stable HEK293 cells. For the serum amidation assay, serum samples were shipped over on ice from collaborators in Finland and stored at -80°C in Oxford.

4.2.8 Caspase-3 activity assay

Caspase-3 activity was probed using the EnzChek Caspase-3 assay (E13183, Invitrogen, California). It is based on the proteolytic cleavage of the aminomethylcoumarin (AMC)-derived Z-DEVD-AMC substrate and is specific to caspase-3/-7 activities. Cells were lysed in a buffer provided with the kit as per the manufacturer's instructions. As a guide, 50μl lysis buffer was used to lyse $5x10^5$ cells/sample, volume of lysis buffer was scaled up/down as per estimated number of cells in sample at harvest. Following this, samples were run on the assay as per kit guidelines and cleaved substrates extrapolated from the standard curve included with every run. Relative caspase-3

activity was calculated by normalisation of caspase-3 activity to that of WT-transfected cells within each experiment and plotted as mean normalised values (%) across experiments per variant on GraphPad Prism 8.0. Results (caspase-3 activity) quoted in text are given as percentages for R36S-PAM or S539W-PAM compared to WT-PAM transfected cells.

4.2.9 CyQUANT cell proliferation assay

Viable cell count was measured using the CyQUANT Direct Cell Proliferation Assay kit (C35012, Thermo Scientific) following manufacturer's guidelines. Cells were seeded into 96-well plates (Corning) and transfected with membrane-integral PAM constructs (pCMV6-hPAMv3 constructs: WT, S539W or R36S) or empty vector (EV: pCMV6-Entry). CyQUANT was performed after either 0, 3 or 6 days post-transfection. CyQUANT components were diluted in pre-warmed cell culture medium before use and cells were incubated in the detection reagent for 1hr at 37°C in the dark. Sample fluorescence was measured at 485nm/530nm (excitation/emission) on the Enspire multimode plate reader (Perkin Elmer). Signals were normalised to those from empty vector (EV)-transfected cells within each time point and plotted as mean normalised signals (%) over time on GraphPad Prism 8.0.

4.2.10 Insulin secretion assays

See chapter 2 for full details on how insulin secretion assays were performed (section 2.19). For the characterisation of variant PAM (exogenous), CRISPR PAM KO EndoC-βH1 cells were transfected with either pCMV6-GFP or membrane integral PAM (Table 4.2.2.1) 24hr post-seeding. Cells were incubated in 2.8mM glucose media 48hrs post-

transfection for overnight at 37°C/5% CO_2. Glucose-stimulated insulin secretion assays were then performed the following morning (72hr post-transfection) and cells were stimulated with media containing either 2.8mM glucose, 16.7mM glucose or 16.7mM + 100μM tolbutamide + 10μM Forskolin for 1hr at 37°C/5% CO_2. Supernatant and cellular extracts were then harvested and frozen at -20°C prior to detection using the AlphaLISA kit (Perkin Elmer). Raw insulin secretion reads (pg) were adjusted to intracellular content measures (pg) per well and normalised to GPF-transfected cells at basal (2.8mM) glucose and plotted as mean secretion (%) across the experiment.

4.2.11 Quantitative real-time PCR (qRT-PCR)

For quantification of gene expression, RNA was extracted, cDNA synthesised and qRT-PCR set up as described in section 2.15. For the gene expression analyses performed in this chapter, details on the Taqman probes used are summarised in Table 4.2.11.1. Data was analysed using the ddCT method, comparing gene expression of a test sample to a control sample in the same plate.

Taqman® probe	Assay ID	Specificity
*GAPDH**	Hs99999905_m1	Exon 3
*PPIA**	Hs99999904_m1	Exon 5
*TBP**	Hs00427620_m1	Spans exons 2-3
PAM_t1	Hs00168596_m1	Spans exons 11-12
PAM_t2	Hs01084050_m1	Spans exons 24-25

Table 4.2.11.1 Taqman probes used for assaying gene expression by qRT-PCR. All Taqman probes were purchased from Life Technologies. * indicates probes used as housekeeping genes in each qRT-PCR run.

4.2.12 Statistical analysis

Unless otherwise stated, data were normalised to control samples (usually either WT-PAM or empty vector (EV) transfected cells) within each experiment and collated as mean normalised values (%) across an experiment. For statistical analysis, mean data points were transformed into log and analysed using the two-tailed unpaired Student's t-test with the Welch's correction when comparing one variable or the ANOVA when there were multiple variables. For ease, data have been plotted as the normalised averages (%) rather than log values in GraphPad Prism 8.0.

4.3 RESULTS

4.3.1 Functional validation of the PAM KO EndoC-βH1 cell line

The PAM KO EndoC-βH1 cell line was generated by a fellow DPhil student, Antje Grotz, as described in section 2.5. I first performed functional validation on this cell line, since I would be using this cell line to characterise the impact of R36S-PAM on both PAM function and the consequential impact on β cell function. This was performed in collaboration with Antje Grotz, with relative contributions detailed in the individual figure legends.

Following generation of the cell line, western blots confirmed that there was a complete loss of PAM protein in these PAM KO cells as compared to controls cells (FIG. 4.3.1.1). Bands were detected using an antibody targeting the PAL domain (hereafter referred to as Nter PAM, Table 4.2.3.1). Following this, I performed immunofluorescence staining and immunogold labelling/TEM to confirm the loss of endogenous PAM. Here, I used an antibody targeting the C-terminus (referred to as Cter PAM, Table 3.2.3.1). Both methods confirmed PAM localisation to vesicles containing insulin. EM shows that PAM is detectable on both immature/mature granules as indicated by the presence of a halo in mature granules (FIG. 4.3.1.2). The localisation pattern observed was similar to that observed in unedited EndoC-βH1 cells (FIG. 3.3.3.1 (IF), FIG. 3.3.3.2 (EM)).

These data suggest that there may not be a complete *PAM* KO in these cells, a discrepancy that was not identified looking at the western blots alone. Presence of (Cter) PAM staining in both IF and EM, but absence of (Nter) PAM bands in westerns

suggests that this residual isoform lacks the PAL domain but contains the Cter domain of endogenous PAM. There is a short C-terminal transcript that has a transcriptional start site (TSS) after exon 12 and so was not edited following CRISPR-Cas9 genome editing. Therefore, this PAM peptide would only be detected with the Cter PAM antibody.

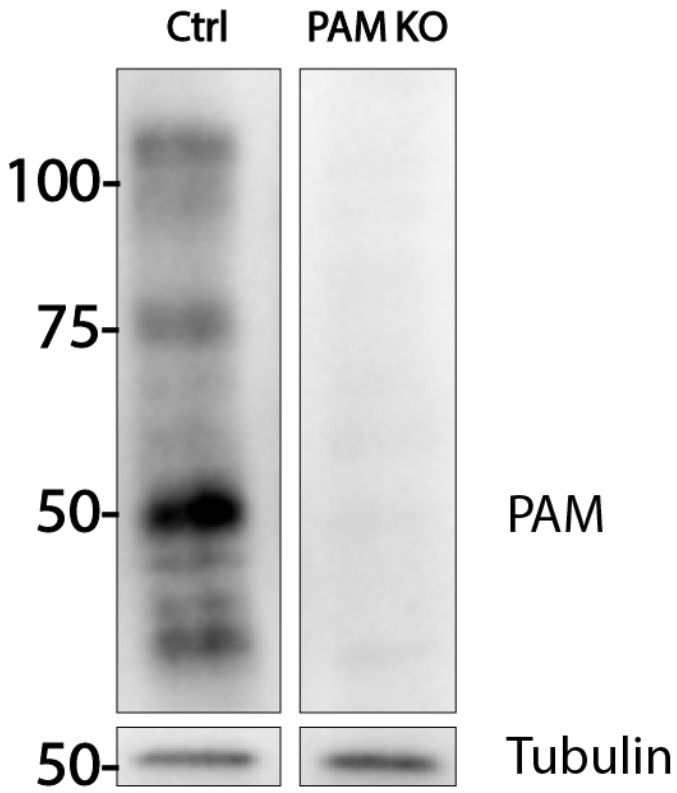

FIG. 4.3.1.1 Western blot analysis exhibits loss of endogenous PAM protein expression in PAM KO EndoC-βH1 cells. Blots representative of 3 independent experiments and were performed by Antje Grotz (DPhil Student, OCDEM). Ctrl (control) samples indicate unedited EndoC-βH1 cells.

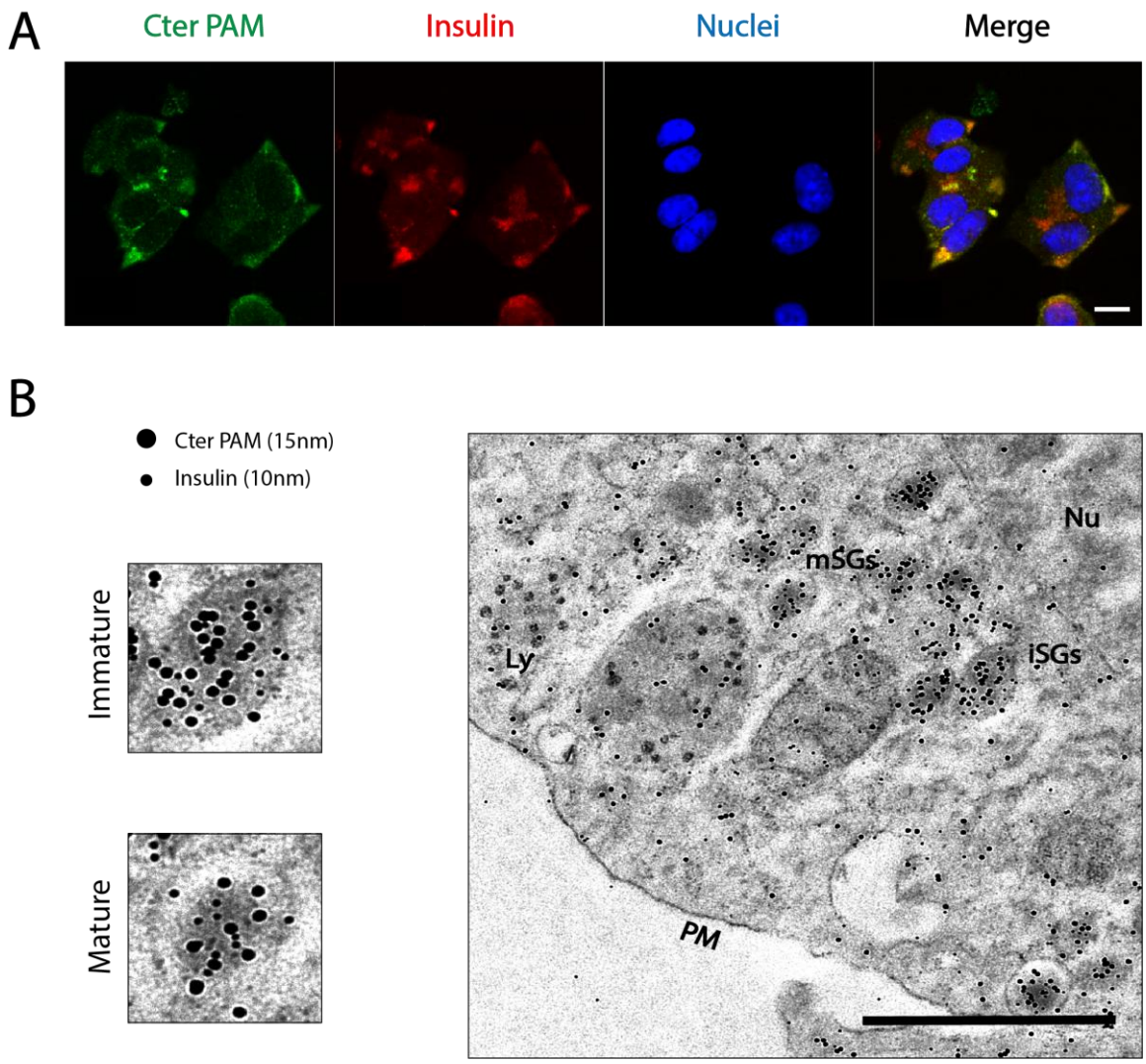

FIG. 4.3.1.2 Immunostaining of CRISPR PAM KO EndoC-βH1 cells displays (Cter) PAM staining. (**A**) Fixed PAM KO EndoC-βH1 cells were double immunostained for PAM (Cter, green) and insulin (red) and nuclei stained using the far-red dye Red dot2. PAM staining was amplified using biotin/streptavidin-488. Cells were visualised using a BioRad Radiance 2100 confocal microscope. Scale bar indicates 10µm. (**B**) Cells were immunogold labelled with both PAM (15nm gold) and insulin (10nm gold) and imaged using transmission electron microscopy. Scale bar indicates 1µm. PAM labelling was observed in both immature (iSGs) and mature granules (mSGs) as defined by the absence/presence of a halo. Cter: C terminal; PM: plasma membrane; Nu: nucleus; Ly: lysosome.

Gene expression analyses were also conducted to characterise the expression of *PAM* in these cells. The Taqman probe, *PAM*_t1, targets exons 11-12 and is therefore capable of identifying the predominant endogenous transcripts in EndoC-βH1 cells. An alternative probe, *PAM*_t2, which targets exons 24-25, will detect endogenous PAM transcripts in EndoC-βH1 cells but is also capable of detecting a shorter C-terminal transcript which may not be detected by *PAM_t1* in the PAM KO cells. There was reduced expression with both probes in PAM KO cells compared to unedited EndoC-βH1 cells (*PAM_t1*: -92%, P=0.01; *PAM_T2*: -73.2%, P=0.024; n=3) (FIG. 4.3.1.3). There was higher *PAM* expression detected by *PAM_t2* compared to *PAM_t1* (+18.8%, P=0.02), suggesting the detection of a residual C-terminal *PAM* transcript as per my hypothesis. Both western blots and qRT-PCR were performed by Antje Grotz (DPhil student, OCDEM).

The presence of a residual C-terminal isoform in the PAM KO EndoC-βH1 cells is important to consider when interpreting my data in the remainder of this chapter (and chapter 5). Data so far suggests that since this transcript is unlikely to encode either catalytic domain, the PAM KO EndoC-βH1 cells are likely to be catalytically-inactive. The relative contribution of this C-terminal fragment to β cell function remains undescribed and has not been characterised further. Although there is residual expression of a C-terminal transcript, for practicality this cell line will be referred to as PAM KO (EndoC-βH1) cells.

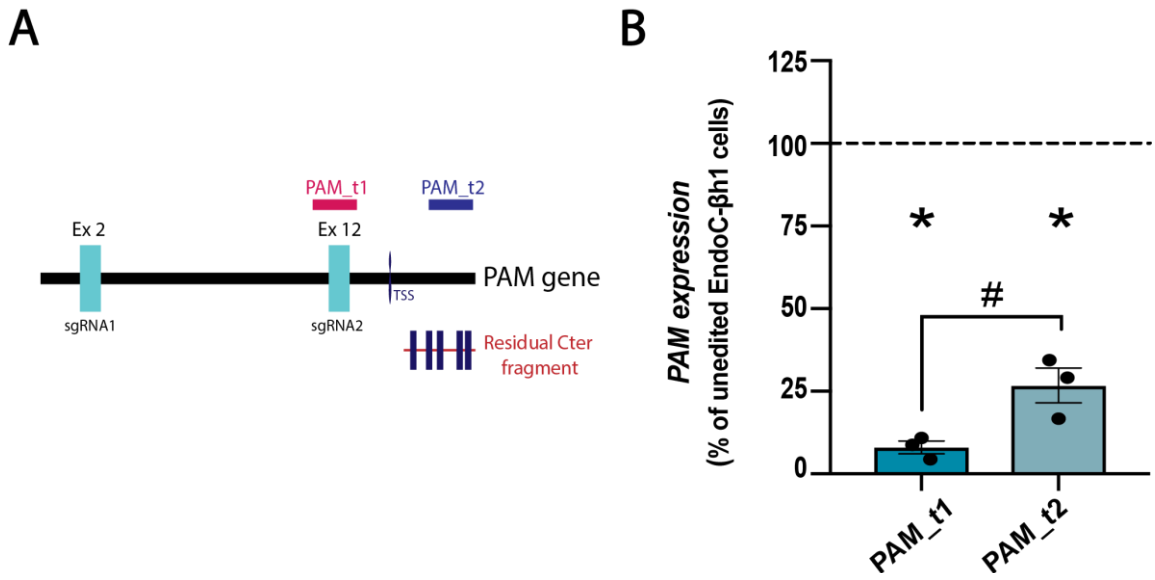

FIG. 4.3.1.3 Gene expression analyses confirm expression of short C-terminal PAM transcript in PAM KO EndoC-βH1 cells. (**A**) Schematic showing the position of the Taqman™ probes and single guide RNA (sgRNA) binding sites relative to *PAM*. The residual C-terminal (Cter) transcript predicted to be remaining in the PAM KO cells has a transcriptional start site (TSS) after the 2nd sgRNA binding site at exon (Ex) 12. (**B**) qRT-PCR was performed by Antje Grotz (DPhil Student, OCDEM) using *PAM_t1* (targeting exon 11-12) and *PAM_t2* (targeting exons 24-25) comparing to unedited EndoC-βH1 cells (dotted line). Data are representative of 3 independent experiments and error bars are SEM. Statistical analyses performed using two-tailed unpaired Student's t-test with Welch's correction on transformed (log) values. * compares either PAM_t1 or PAM_t2 expression in PAM KO cells to the unedited control EndoC-βH1 cells (dotted line). # compares PAM_t1 with PAM_t2 expression in PAM KO cells only. * and $^{\#}$ denote $P<0.05$.

4.3.2 Exome sequencing identified R36S-PAM as a candidate causal variant in proband with neonatal diabetes mellitus (NDM)

Exome sequencing was performed in 15 probands with presenting with early onset and/or neonatal diabetes mellitus and their unaffected parents (Lori Bonnycastle and Francis Collins, NIH, USA) to identify *de novo* causal variants driving disease pathogenesis. Exome sequencing targets the protein-coding portion of the genome and therefore is capable of identifying coding and splicing variants only. The *de novo* mutation rate in humans is 1.2e-08, which equates to ~ 38 mutations per offspring across the genome (Francioli, Polak et al. 2015). Shortlisting of *de novo* variants was performed through a 3-tier process, based on the following criteria: a) predicted to be damaging (nonsynonymous, premature stop codon, splice acceptor site disruption, frame shift), b) segregation with disease status, c) variant prevalence that fits with the mode of inheritance (autosomal dominant alleles <0.0001, recessive alleles <0.01).

This identified one proband with a novel, nonsynonymous, *de novo* variant in *PAM* (FIG. 4.3.2.1). *PAM* is a known T2D susceptibility gene, making R36S-PAM a good candidate causal allele for monogenic diabetes in this proband. Several other variants were also identified as candidate causal variants, one of which R1289-DOCK3 is previously reported in GnomAD and therefore unlikely to be causal (Table 4.3.2.2). *DOCK3* is predominantly expressed in the central nervous system (CNS) and functions as a guanine nucleotide exchange factor (GEF) mediating intracellular signalling (Marie, Hay et al. 2014). *CHD3* encodes a helicase involved in chromatin remodelling and is involved in brain development (Rodriguez-Castaneda, Lemma et al. 2018, Snijders Blok, Rousseau et al. 2018). *CDH2* encodes cadherin-2, which is closely related to N-

cadherin (Nieto-Estevez and Hsieh 2018). N-cadherin mediates cell-cell adhesion between pancreatic β cells and global knock-out mouse models have defects in β cell turnover (Johansson, Voss et al. 2010). However, CRISPR-Cas9 mediated knock-out of *CHD2* in Ins1 cells did not uncover any defects in insulin secretion as compared with unedited Ins1 control cells (personal communication, Lori Bonnycastle, NIH). This suggests that the Q448*-CDH2 variant is unlikely to be driving NDM pathogenesis in this proband. The variants identified in *CHD3* and *DOCK3* have not been functionally characterised.

The proband presented at 8 months of age with moderate diabetic ketoacidosis (pH 7.13), elevated plasma glucose (32.6mmol/L) and reduced plasma C-peptide (0.11nmol/L). There were no extra-pancreatic clinical features detected. The proband was born in term during the 40^{th} week of gestation and weighed 3.95 (+0.2 SD) with a length of 54 (+2.2 SD) and had an Apgar score of 8 (at 1min) and 9 (at 5min). The Apgar test (appearance, pulse, grimace, activity, respiration) is performed to check the health of a new-born baby; a score of 1-10 is given with 10 being the most healthy (Apgar 1953). For the proband, these observations were all within normal range. A normal birth weight is unusual for a diagnosis of NDM, and is consistent with normal foetal insulin secretion *in utero*. The proband was started on insulin at a dose of 1.3iu/kg/day and the latest dose is 0.6iu/kg/day. There was no evidence of autoantibodies to islet proteins which would indicate a Type 1 diabetes (T1D) phenotype, except for insulin positive autoantibodies following the initiation of insulin therapy.

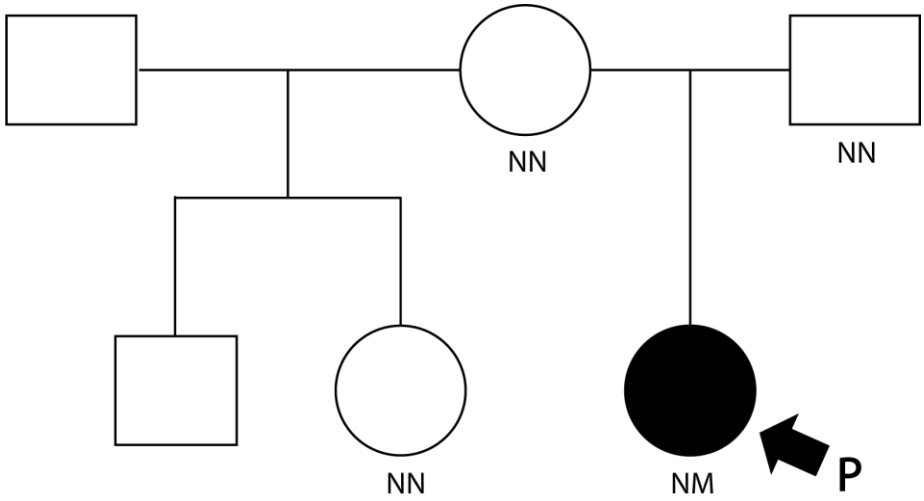

FIG. 4.3.2.1 Pedigree of affected proband with early onset diabetes mellitus. Exome sequencing was performed on a family trio where the proband (black filled circle, arrow) was diagnosed with early onset diabetes but the parents were unaffected (hollow), to identify a *de novo* causal variant underlying diabetes in the proband. NN: homozygous wild-type (non-carriers); NM: heterozygous for (likely) causal variant driving NDM pathogenesis.

Gene	Protein	Variant [a]	Frequency [b]
CDH2	Cadherin 2	Q448*	-
CHD3	Chromodomain helicase binding DNA binding protein 3	M1063L	-
DOCK3	Dedicator of cytokinesis 3	R1289Q	$4.03e^{-6}$ (non-Finnish Europeans)
PAM	PAM	R36S	-

Table 4.3.2.2 Exome sequencing and shortlisting by *in silico* analysis identified 4 heterozygous *de novo* candidate causal variants in one proband. [a] Variant position in protein sequence**;** [b] Frequency of minor allele as listed in GnomAD, accessed on 25th April 2019, at https://gnomad.broadinstitute.org. The variant of interest (in *PAM*) is highlighted in red.

R36S is positioned between the N-terminal ER signal peptide and the proregion in PAM, which makes it unlikely to have a direct impact on (catalytic) amidating activity (FIG. 4.3.1.3). There may, however, be an effect on the global folding and stability of PAM, and this may result in degradation or impaired trafficking through the cell (Liu, Hodish et al. 2007, Cunningham, He et al. 2017). The proregion in PAM has been shown to assist in the secretion of its soluble isoform and has a conserved role in the folding and assembly of the global structure (Mains, Milgram et al. 1995).

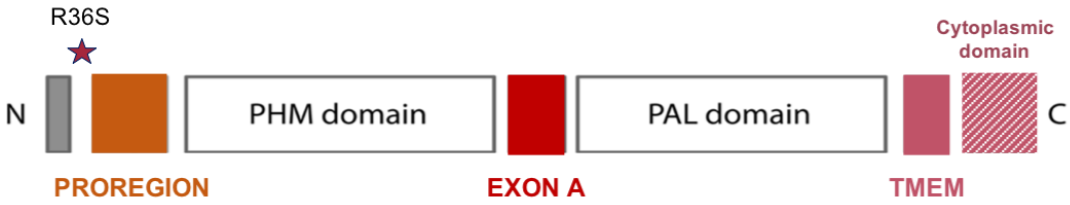

FIG. 4.3.2.3 R36S sits at the N-terminal of the proregion in PAM. Position of variant (star) outside catalytic domains, makes it unlikely to directly affect amidating activity and more likely to affect PAM protein folding, stability and/or trafficking. TMEM: transmembrane domain.

4.3.3 Impact of R36S-PAM on amidation activity

Although R36S is positioned outside the catalytic domains, an impact on the global protein folding of PAM may initiate deficiencies in its ability to amidate peptides. As such, serum samples from the proband and immediate family were collected and run through a radiolabelled amidation assay, which examines the total amidating profile in serum samples (using a fixed volume of serum per assay). This was performed by a fellow DPhil student in the lab, Mahesh Umapathysivam.

Serum amidating profiles of individuals with/without known T2D risk alleles were included as functional standards. These show a step-wise reduction in serum amidating profile between wild-type (WT) individuals without T2D-risk alleles (392pmol/µl/hr), heterozygous carriers of the D563G allele (314pmol/µl/hr), homozygous carriers of the D563G allele (240pmol/µl/hr) and heterozygous carriers of the S359W allele (188pmol/µl/hr). D563G-PAM has been shown previously to impair amidating activity and S539W-PAM results in reduced expression, thereby also affecting the serum amidating profile (Thomsen, Raimondo et al. 2018). Neither the proband, nor her immediate family were found to carry either of the known T2D-asociated variants (personal communication, Lori Bonnycastle, NIH). Despite this, I observed from the data that both the proband and her father displayed a reduced amidating serum profile, which was more than 1 standard deviation unit below that of WT individuals (FIG. 4.3.3.1).

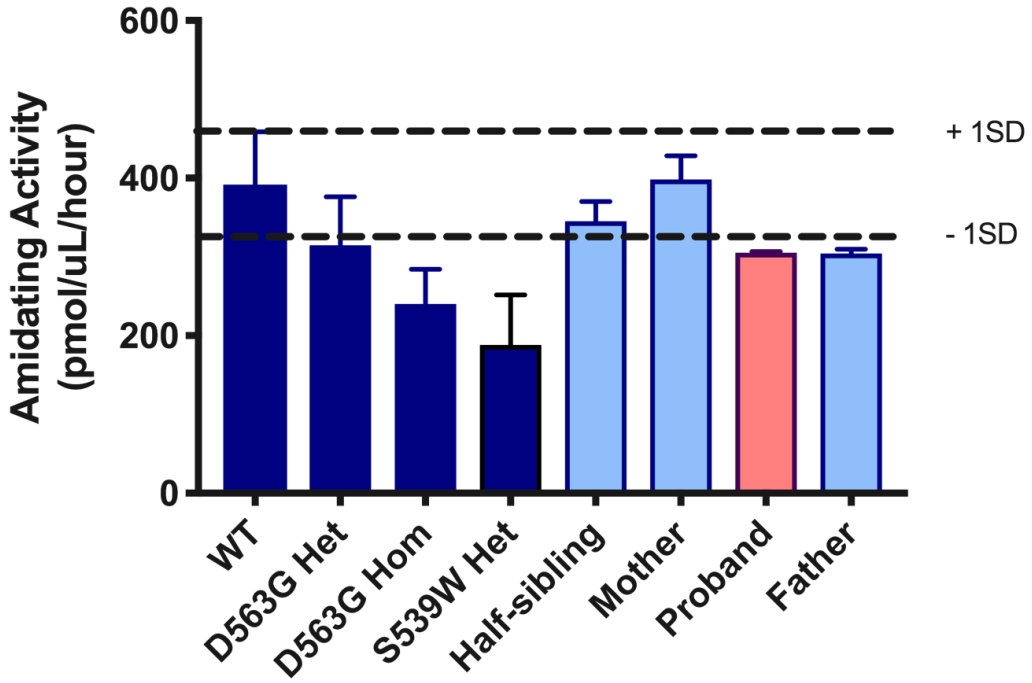

FIG. 4.3.3.1 Serum amidation profile of proband is nominally reduced. Serum samples were collected from the proband (pink), mother, father and half-sibling (light blues) and run in a radiolabelled amidation assay run by Mahesh Umapathysivam. Samples were run in technical triplicates and compared to serum standards (navy) collected from individuals with/without T2D risk alleles (WT: 26 individuals, D563G Hets: 25 individuals, D563G Homs: 21 individuals, S539W Hets: 26 individuals) in *PAM*. D563G is a common (MAF=5%) T2D risk allele with known defects in amidation activity; S539W is a low-frequency (MAF=~1%) T2D risk allele with known mislocalisation and expression defects, and therefore results in reduced serum amidating profile (Thomsen, Raimondo et al. 2018). N=1 for proband and family, error bars are standard deviation (SD) Het: heterozygous; Hom: homozygous; WT: wild-type.

Whilst the serum amidation assay provides a physiological and relevant indication of PAM activity, there are some caveats. Firstly, samples are not matched for PAM content, and so deviations from normal/wild-type serum amidating activity could be mediated by reduced activity or reduced PAM abundance, but the contributions of these effects cannot be quantified. The variability observed in (control) individuals without PAM variants (denoted as WT) could be attributed to variation between multiple individuals (n>1), whereas there is only an N=1 for the proband and immediate family.

To address these caveats, I performed an *in vitro* amidation assay in parallel, in which activity is normalised to the PAM content per sample and such gives a more accurate measure. Recombinant luminal (untethered) PAM generated through stable transfection of HEK293 cells was harvested in the supernatant, concentrated and quantified using western blots (FIG. 4.3.4.2). There were no differences in the amidating activity of R36S-PAM compared to WT-PAM (FIG. 4.3.3.2), suggesting that the reduced amidating activity observed in the proband serum is likely supported by a reduced abundance of PAM in the β cell. By contrast, D563G had been previously shown to impair amidating activity by ~25% of WT-PAM activity in this assay (Thomsen, Raimondo et al. 2018).

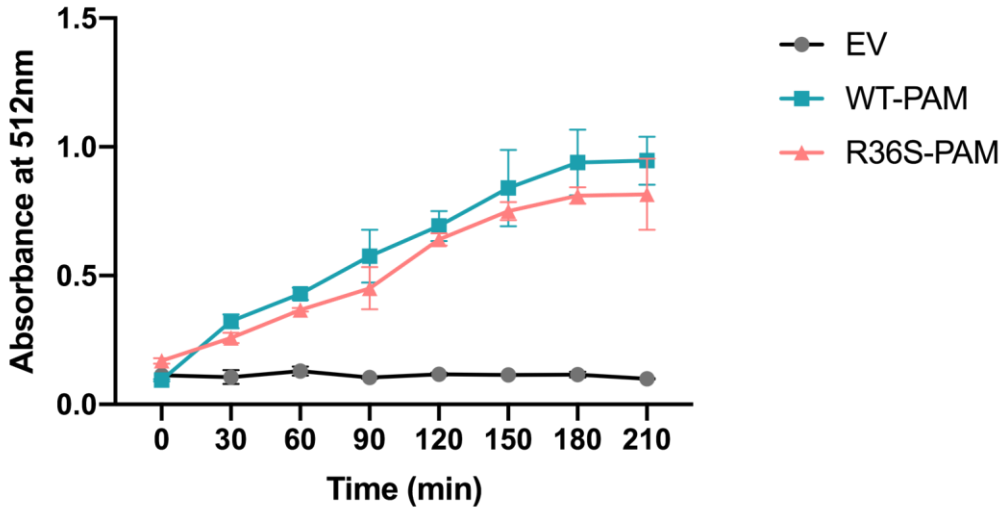

FIG. 4.3.3.2 R36S-PAM exhibits normal amidating activity as assessed by an *in vitro* amidation assay. Supernatant samples from HEK293 cells stably transfected with luminal PAM were quantified for PAM content within each sample to generate a molar scale ratio relative to WT-PAM samples. This was then used in the amidation assay to normalise loading amounts for PAM content. Assay detects the production of glyoxylate by-product, which is in a 1:1 stoichiometry of PAM and hippuric acid (substrate) amount. Data is from one experiment, consisting of 3 technical replicates per time point per variant PAM sample. Error bars are standard deviation. EV: empty vector; WT: wild-type.

4.3.4 The luminal R36S-PAM isoform is unstable and degraded by the proteosome

To determine the impact of R36S on untethered/luminal PAM abundance (hereafter referred to as luminal), I transiently transfected CRISPR PAM KO EndoC-βH1 cells with the construct for untethered luminal PAM (pCMV6-Luminal PAM, WT or R36S) and used western blot analysis to detect PAM species in the lysate. I observed reduced abundance of luminal R36S-PAM as compared to WT-PAM (-78%, n=3, P= 0.12), but this was not statistically significant (FIG. 4.3.4.1). I had also seen a similar reduction in

the abundance of luminal R36S-PAM in the supernatant of stable HEK293 cell lines (FIG. 4.3.4.2).

Since the reduction in R36S-PAM content seems a common feature in EndoC-βH1 and HEK293 cells, I hypothesized that this is due to altered protein stability rather than a β cell-specific effect. To assess the protein stability, EndoC-βH1 cells transfected with luminal PAM were then incubated with either a proteosomal inhibitor, MG132 (Sigma Aldrich, 10μM) or DMSO. Following 2hr incubation with MG132, luminal R36S-PAM was rescued in EndoC-βH1 (FIG. 4.3.4.3), suggesting that luminal R36S-PAM is degraded by the proteosome, thereby resulting in reduced abundance in the cell extract and supernatant.

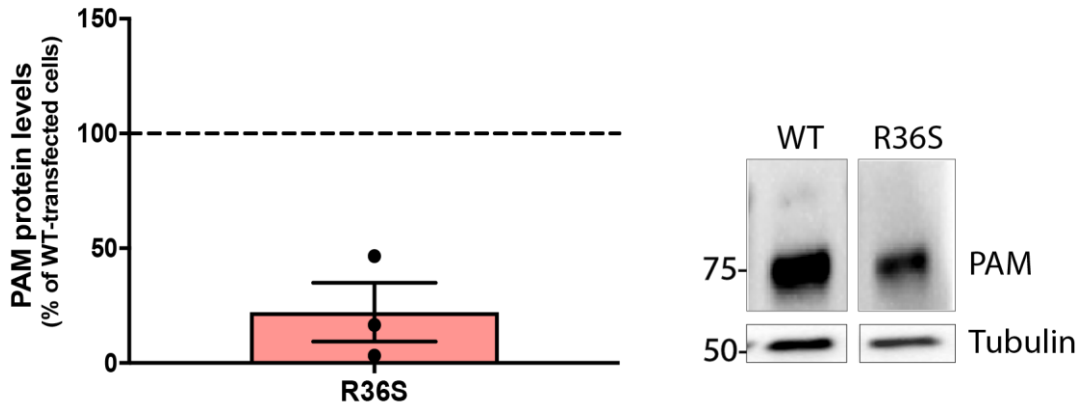

FIG. 4.3.4.1 R36S results in a trend towards reduced luminal R36S-PAM levels in CRISPR PAM KO EndoC-βH1. Cells were transfected with luminal PAM for 72hr and then PAM protein abundance assessed by western blot. Bands were quantified using lane densitometry, adjusted for tubulin band intensities and normalised to WT (dotted line) within each experiment. Data are analysed by unpaired Student's t-test with Welch's correction on transformed (log) values. Data are representative of 3 independent experiments and error bars are SEM.

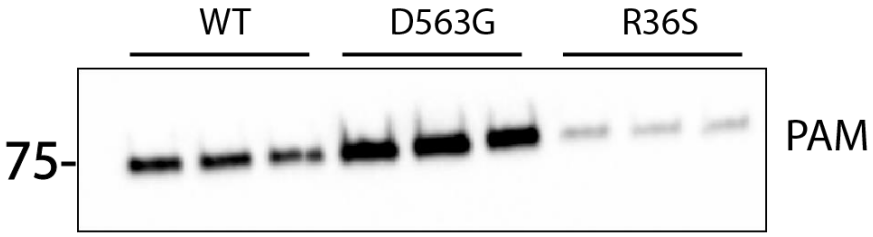

FIG. 4.3.4.2 R36S-PAM is depleted in the supernatant of stably transfected HEK293 cells. Supernatant collected from HEK293 cells stably transfected with luminal R36S-PAM were assessed by western blot to measure secreted PAM abundance. Blot is representative of 4 independent experiments containing 3 technical replicates per experiment (supernatants collected from sequential passages of stable HEK293 PAM cell lines). No loading control included as this is supernatant. WT: wild-type; D563G is a known T2D-associated allele with impaired amidating activity (Thomsen, Raimondo et al. 2018).

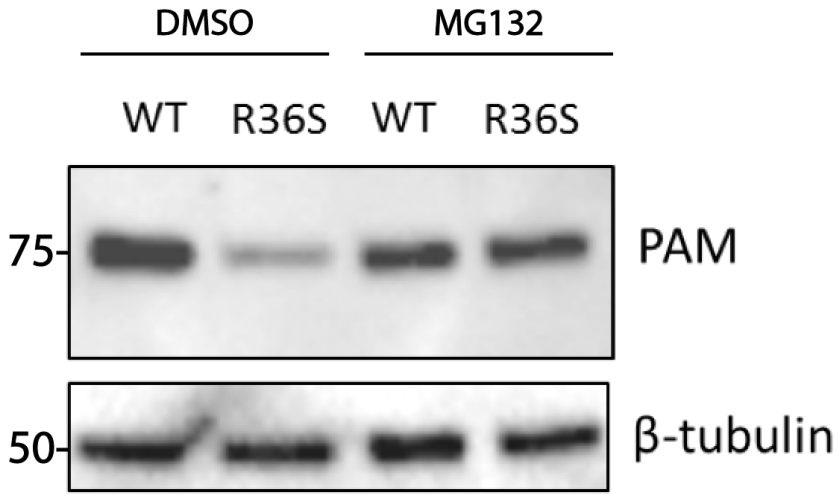

FIG. 4.3.4.3 Proteosomal inhibition rescues luminal R36S-PAM in EndoC-βH1 cells. Cells were transfected with luminal PAM and then incubated with 10μM MG132 (or DMSO control) for a further 2hrs. Lysates were examined for PAM abundance using western blot. Blot representative of two independent experiments.

4.3.5 Luminal R36S-PAM and the unfolded protein response (UPR)

The UPR is a dynamic, physiological mechanism that protects the cell from undue stress and dysfunction from misfolded proteins (Walter and Ron 2011). However, when there is a prolonged activation of the UPR, these mechanisms may switch from homeostatic to pathological resulting in cellular dysfunction and/or apoptosis (Sano and Reed 2013). Since luminal R36S-PAM is proteosomally degraded, I hypothesised that any activation of UPR markers would be minimal and promote the clearance of the misfolded protein in a physiological manner. Nevertheless, I examined the impact on the abundance of key UPR markers (CHOP, IRE1, PERK; see section 1.6.2.1) by western blot. These markers are indicative of different arms of the UPR, so unless there is generalised and severe ER stress where there may be trans-activation of multiple sensors, I would not expect to see activation of all markers. I did not observe differences in any of the markers, except for a modest trend towards increased phospho-IRE1 (p-IRE1) in R36S-PAM transfected cells compared to WT-PAM transfected cells (+30%, n=3, P=0.08) (FIG. 4.3.5.1). This, in combination with the evidence that membrane-integral PAM is the predominant isoform in the β cell (FIG. 3.3.1.1), suggests that the membrane-integral isoform is more relevant to characterising the pathophysiology resulting from the R36S substitution. As such, I have focused on the effects of membrane-integral R36S-PAM only for the remainder of this chapter.

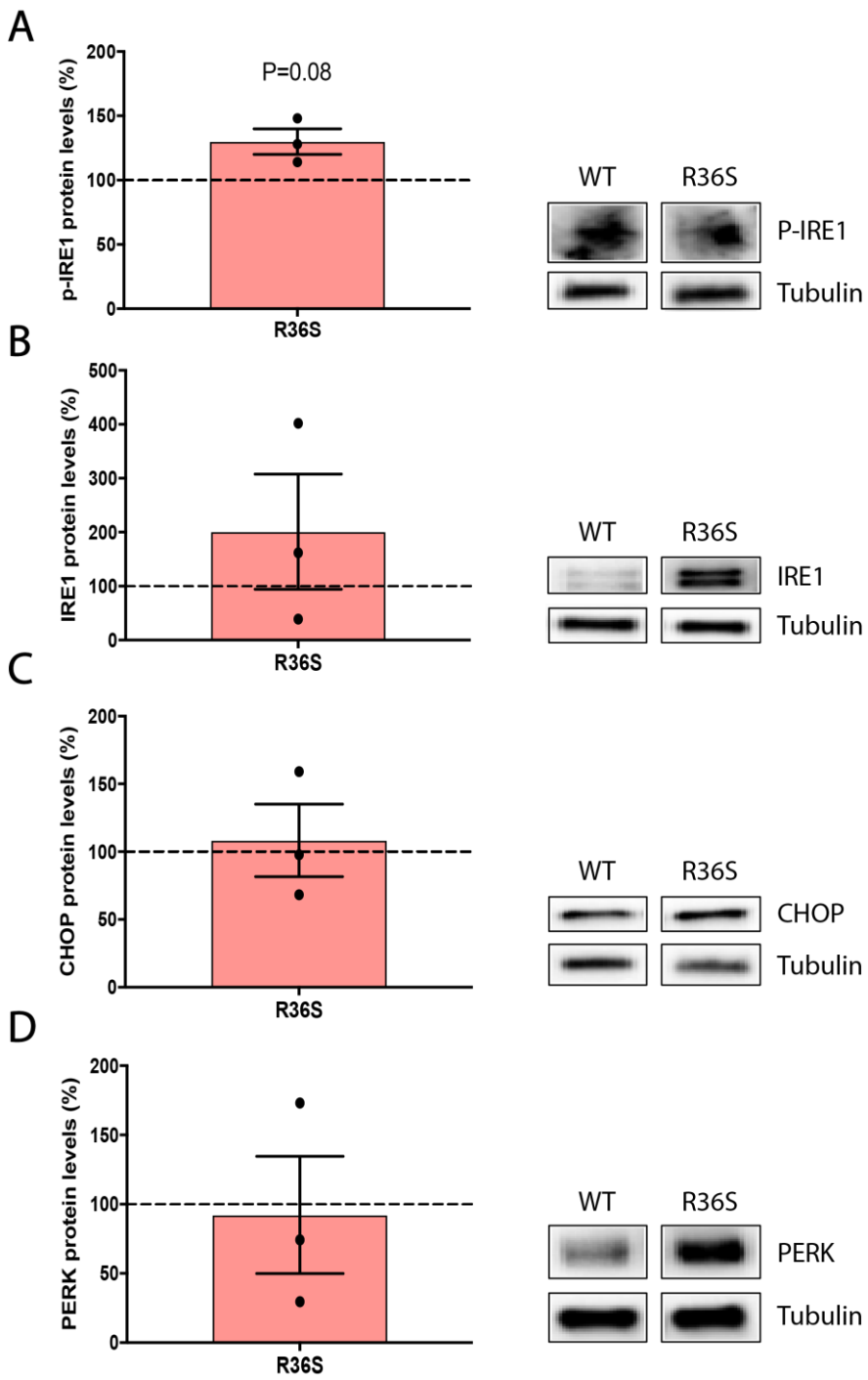

FIG. 4.3.5.1 Luminal R36S-PAM does not result in elevation of UPR markers. Cells transfected with luminal PAM were characterised for abundance of P-IRE1 (**A**), IRE1 (**B**), CHOP (**C**) and PERK (**D**) measured using western blot. Bands were quantified using lane densitometry and normalised to WT-transfected cells (dotted line). Data are representative of 3 independent experiments and analysed by unpaired Student's t-test with Welch's correction on the transformed (log) data. Error bars are SEM.

4.3.6 The membrane-integral isoform is retained within the ER

The predominant transcript for *PAM* in the β cell encodes a membrane-integral protein isoform (FIG. 3.3.1.1). Following the same approach, I next wanted to examine the impact of R36S on membrane-integral PAM in the β cell. I examined PAM abundance in the cells transfected with membrane-integral PAM using western blots. Here, I identified no difference in the abundance of membrane-integral R36S-PAM as compared to WT-PAM (+76.3%, n=3, P=0.56) (FIG. 4.3.6.1).

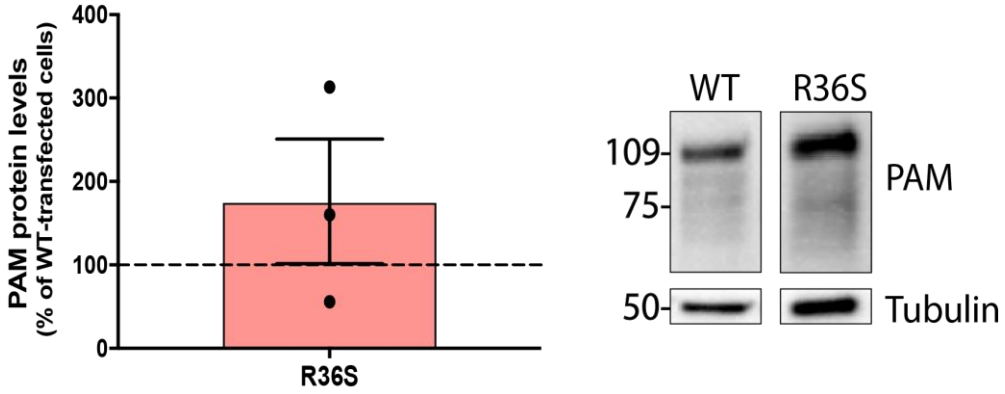

FIG. 4.3.6.1 Membrane-integral R36S-PAM displays similar protein abundance to WT-PAM. Cells transfected with membrane-integral PAM were examined for PAM abundance by western blot. Bands were quantified using lane densitometry, adjusted for tubulin band intensities and then normalised to WT (dotted line) within each experiment. Mean abundance was analysed using unpaired Student's t-test with Welch's correction on transformed (log) values. Data are representative of 3 independent experiments and error bars are SEM.

I probed the effect of the R36S substitution on the trafficking and localisation of membrane-integral PAM. To investigate this, I performed IF on fixed cells, which were co-stained for PAM using anti-Myc targeting a C-terminal tag and anti-insulin. I observed a staining pattern which may imply ER-retention of membrane-integral R36S-PAM, which was not observed with WT-PAM (FIG. 4.3.6.2). Unfortunately, staining for standard ER markers were not successful despite several attempts (Table 4.3.6.3).

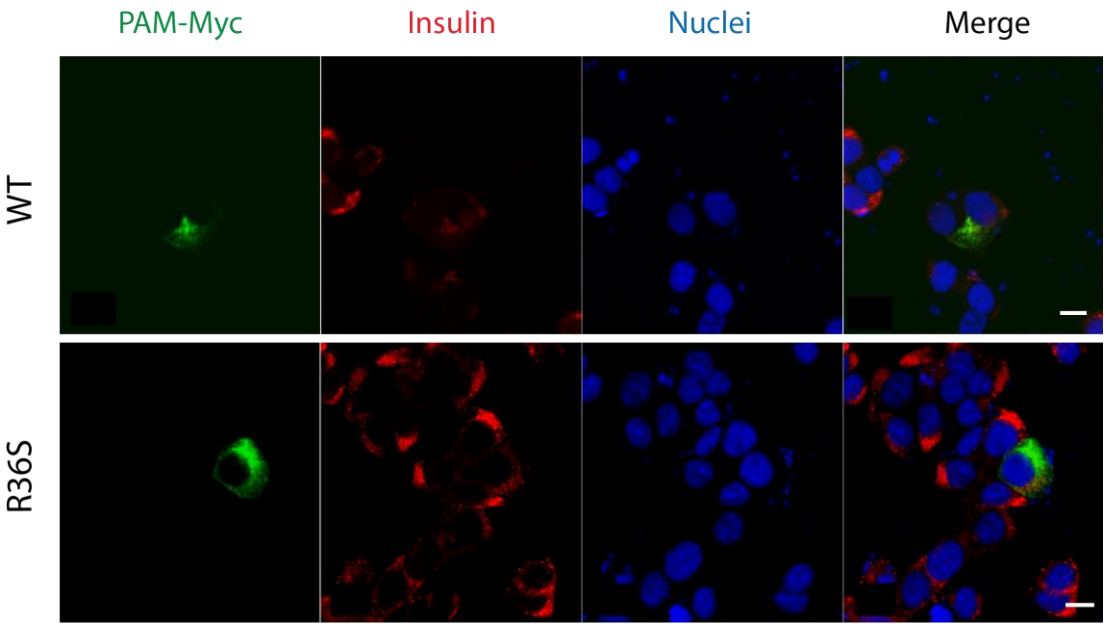

FIG. 4.3.6.2 R36S results in ER-retention of membrane-integral PAM in CRISPR PAM KO EndoC-βH1 cells. Cells were transfected with WT (**A**) or R36S (**B**) membrane-integral PAM, and co-stained for PAM-Myc (green), insulin (red) and nuclei (red dot2, blue) and imaged using a BioRad Radiance 2100 confocal microscope. Scale bar indicates 10μm.

Antibody	Species	Source
Anti-calnexin	Rabbit monoclonal	Santa Cruz, sc-11397
Anti-KDEL	Mouse monoclonal	Santa Cruz, sc-58774
Anti-PDI	Rabbit monoclonal	Cell Signalling, C81H6

Table 4.3.6.3 Antibodies used as markers of the ER for IF were unsuccessful for co-staining EndoC-βH1 cells. Immunostaining of the ER proved difficult when cells were also co-stained for PAM-Myc. Antibodies were diluted in Tris buffer. PDI: protein disulfide isomerase.

4.3.7 Consequences of impaired trafficking on the amidation of chromogranin A

Despite observing no impact of R36S-PAM on either amidation activity or serum amidating profile in the proband, I considered whether the impaired trafficking of membrane-integral PAM may impact its ability to amidate peptides within the β cell. In chapter 3, I confirmed that chromogranin A (CgA) was amidated by (endogenous) PAM in EndoC-βH1 cells (see section 3.3.7) and considered the importance of amidation in the correct trafficking and function of CgA. I therefore hypothesized that I could use the amidation of CgA as a proxy to investigate whether amidation of peptides related to insulin secretion may be affected in R36S-PAM transfected cells. I surveyed the abundance of CgA-gly (precursor) and total CgA by western blot. I observed trends towards increased total CgA abundance in R36S-PAM transfected cells compared to WT-transfected cells (+74.8%, n=5, P=0.07) (FIG. 4.3.7.1). This resulted in a trend towards a reduced proportion of amidated chromogranin A for R36S-PAM compared to WT-PAM (-20.4%, n=5, P=0.09).

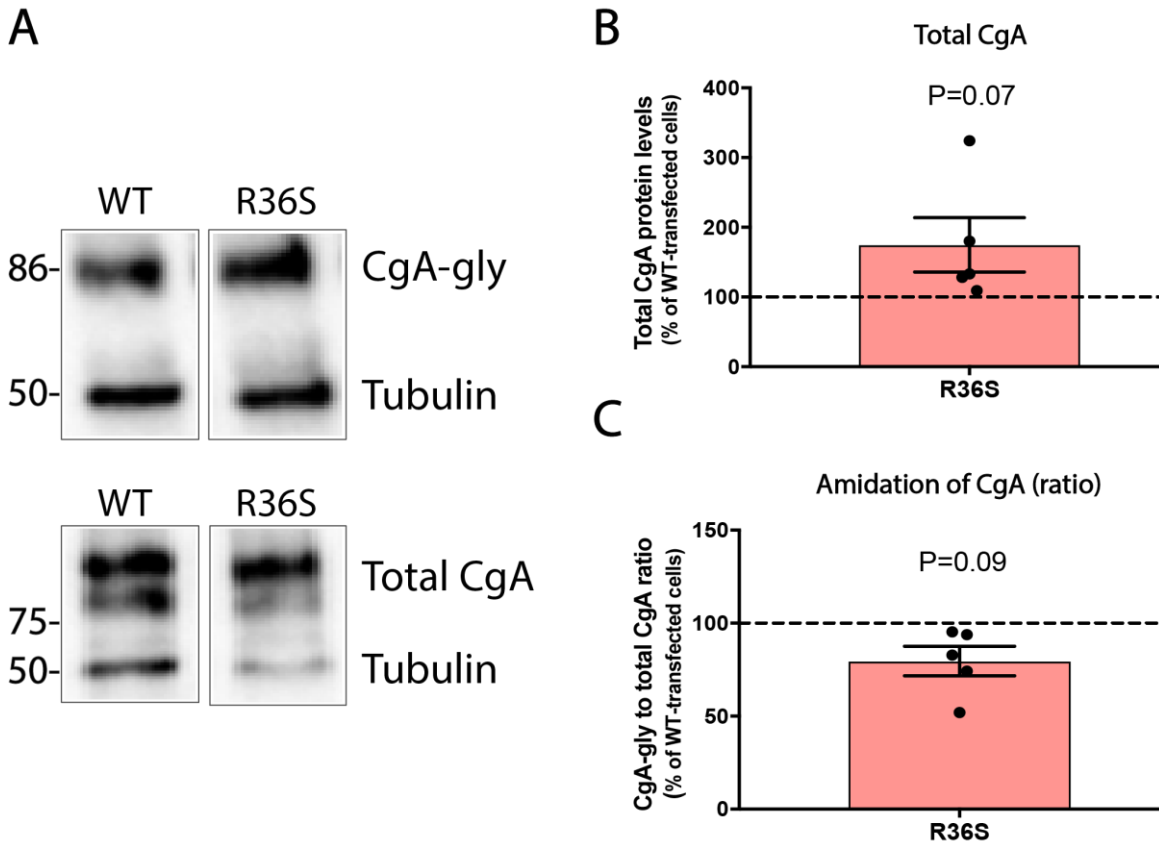

FIG. 4.3.7.1 Membrane-integral R36S-PAM mediates trend towards reduced proportions of amidated chromogranin A (CgA). Cells transfected with membrane-integral PAM were examined using western blot to detect CgA-gly and total CgA abundance (**A**). Quantification of bands demonstrated an increase in total CgA (**B**) and a reduced proportion of amidated CgA relative to total CgA (**C**). Analysis performed by unpaired Student's t-test with Welch's correction on transformed (log) data. Data are representative of 5 independent experiments and error bars are SEM.

4.3.8 Membrane-integral R36S-PAM results in activation of the unfolded protein response (UPR)

Given the observation that membrane-integral R36S-PAM is ER-retained (FIG. 4.3.6.2), I hypothesised that the R36S substitution could result in activation of the unfolded protein response (UPR) and if unresolved this could lead to ER stress and result in cellular dysfunction and apoptosis (Cnop, Foufelle et al. 2012, Sano and Reed 2013). This is of particular relevance given that I observe no degradation of membrane-integral R36S-PAM. This suggests that a restorative ER response, such as ER-associated degradation (ERAD), may not be active.

To investigate this, I examined the abundance of key markers of the UPR using western blot following transfections with membrane-integral PAM. The T2D-risk allele S539W, which is also ER-retained, was included as a positive control (Thomsen, Raimondo et al. 2018). I observed an increase in phospho-IRE1 and a trend towards increased CHOP abundance in R36S-transfected cells compared to WT-transfected cell (p-IRE1: +97%, n=3, P=0.04; CHOP: +15%, n=6, P=0.08) (FIG. 4.3.8.1).

These markers were not significantly elevated for S539W-transfected vs. WT-transfected cell (CHOP: +13%, n=3, P=0.3; p-IRE1: -16.3%, n=3, P=0.47) (FIG. 4.3.8.1). There were no elevations in the abundance of PERK or IRE1 for either variant compared to WT-transfected cell. The P-IRE1: IRE1 ratio, which gives a measure of IRE1 activation, showed a trend towards phospho-activation of IRE1 for R36S (+54%, n=3, P=0.1) suggesting that there is an overall activation of IRE1 relative to the total IRE1

present (FIG. 4.3.8.2). Although not significant, the difference is diluted by the low transfection efficiency in EndoC-βH1 cells. It is likely that the majority of the (total) IRE1 signal comes from cells that are not transfected.

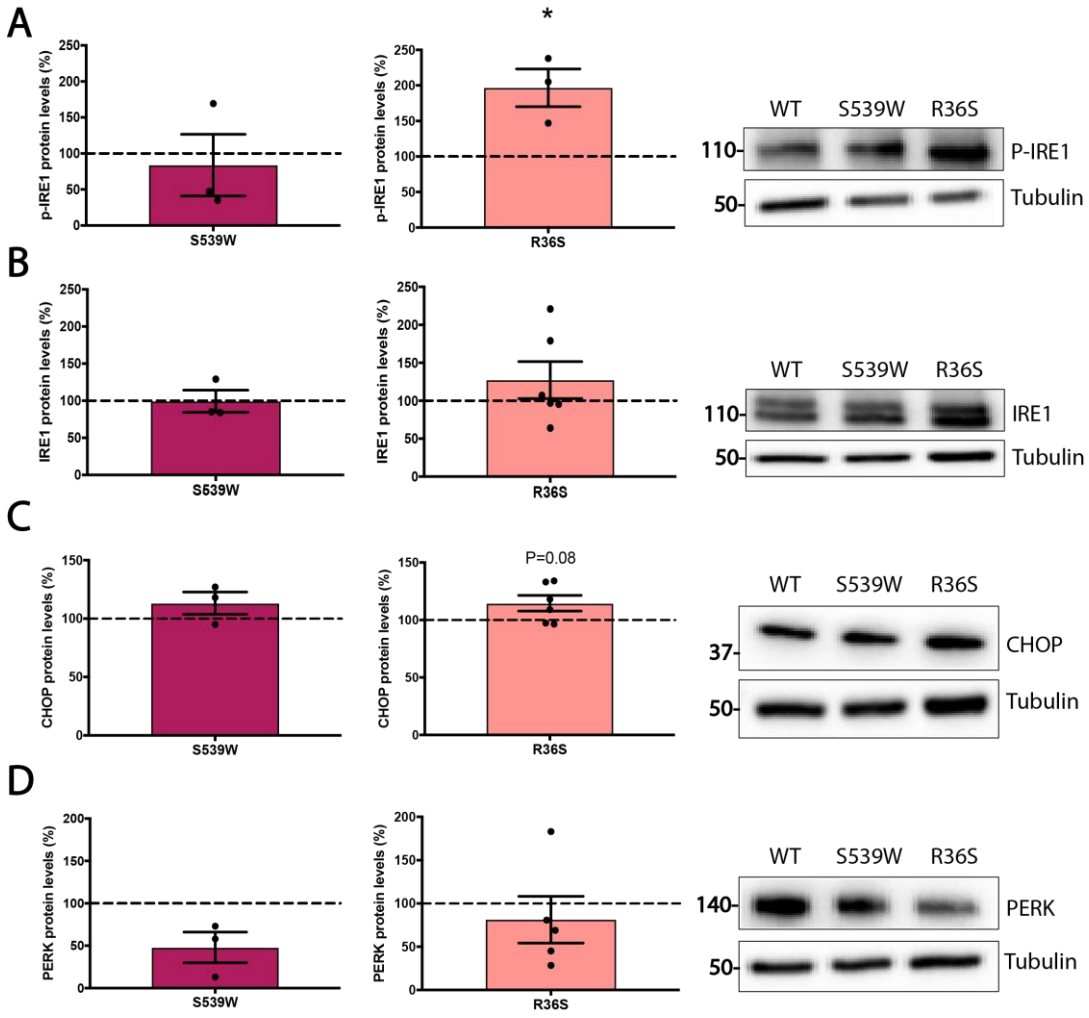

FIG. 4.3.8.1 Markers of the UPR are elevated for R36S- but not S539W-PAM transfected cells. CRISPR PAM KO EndoC-βH1 cells were transiently transfected with membrane-integral PAM and then assessed by western blotting analysis. Relative abundance of P-IRE1 (**A**), IRE1 (**B**), CHOP (**C**) and PERK (**D**) were quantified using lane densitometry and have been normalised to levels in WT-transfected cells (dotted line). Analysis performed using two-tailed unpaired Student's t-test with Welch's correction on transformed (log) fold changes comparing R36S or S539W to WT. Data are representative of 3-6 independent experiments and error bars are SEM. * denotes P<0.05. UPR: unfolded protein response P-IRE1: phospho-IRE1.

The lack of UPR activation observed for S539W provides the first indication of differences in how the β cell handles the 2 different variant PAM proteins (S539W and R36S) despite both exhibiting retention in the ER. Further characterisation of these deviations may help to understand the mechanisms that result in such different clinical presentations associated with each risk allele (S539W: T2D risk; R36S: early onset DM).

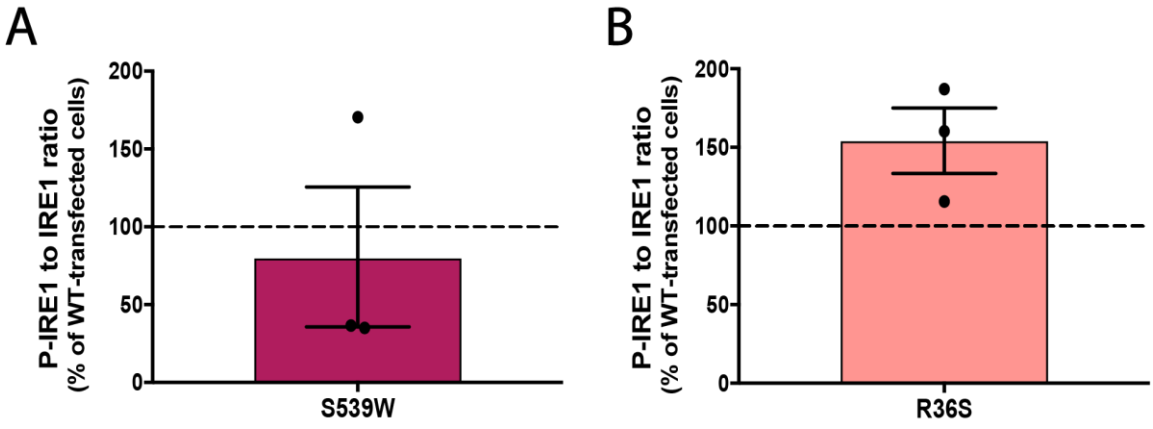

FIG. 4.3.8.2 R36S results in a trend towards elevated phospho-activation of IRE1. Ratios of P-IRE1 to total IRE1 levels from protein abundance measured by western blots in FIG. 4.3.6.1 for R36S (**A**) and S539W (**B**) were calculated and plotted as mean values normalised to WT (dotted line). Analysis was performed by unpaired Student's t-test with Welch's correction on the transformed (log) data. Data representative of 3 independent experiments, error bars are SEM.

4.3.9 R36S-mediated UPR activation initiates downstream pro-apoptotic signalling

Activation of P-IRE1 by membrane-integral R36S-PAM expression may result in pro-apoptotic signalling downstream and may impact cell fate. To investigate this, I examined the protein abundance of key effectors involved downstream of IRE1 activation (FIG. 4.3.9.1).

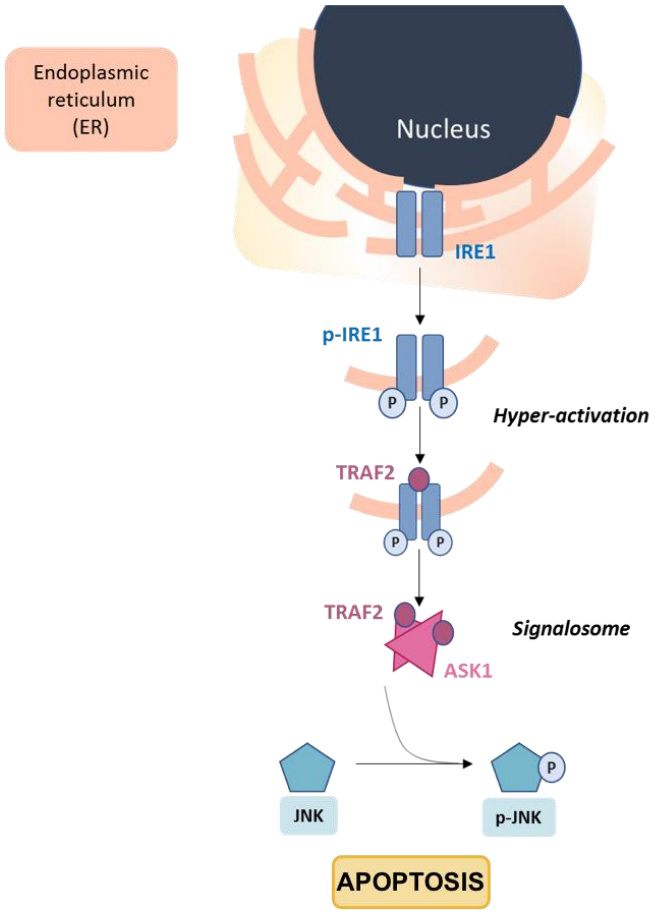

FIG. 4.3.9.1 Pro-apoptotic signalling cascade downstream of IRE1 activation can result in apoptosis. The main effectors illustrated are recruited following hyper-activation of IRE1 and result in the assembly of the ASK1/TRAF2 signalosome, which can lead to the phosphorylation of JNK. This can lead to activation of the apoptotic cascade. TRAF2: TNF-α

associated factor 2; ASK1: apoptosis signal-regulating kinase 1; JNK: c-Jun N-terminal kinase.

CRISPR PAM KO EndoC-βH1 cells were transiently transfected with membrane-integral PAM and the abundance of these effectors measured by western blot. ASK1 is a pro-apoptotic marker and TRAF2 is anti-apoptotic (Nishitoh, Matsuzawa et al. 2002, Hu, Han et al. 2006). For R36S-transfected cells, there was an elevation in ASK1 abundance and a reduction in TRAF2 abundance compared to WT-transfected cells (ASK1: +37%, n=4, P=0.01; TRAF2: -33.1%, n=4, P=0.01). This was not observed for S539W-compared to WT-PAM transfected cells (ASK1: +24%, P=0.07; TRAF2: -11%, P=0.48) (FIG. 4.3.9.2). Unfortunately, P-JNK is very lowly expressed in cells and therefore very difficult to detect (data not shown).

These data provide evidence that membrane-integral R36S-PAM, but not S539W-PAM, mediates pro-apoptotic signalling cascade through IRE1 activation, further clarifying mechanisms which are distinct between the two variant PAM proteins.

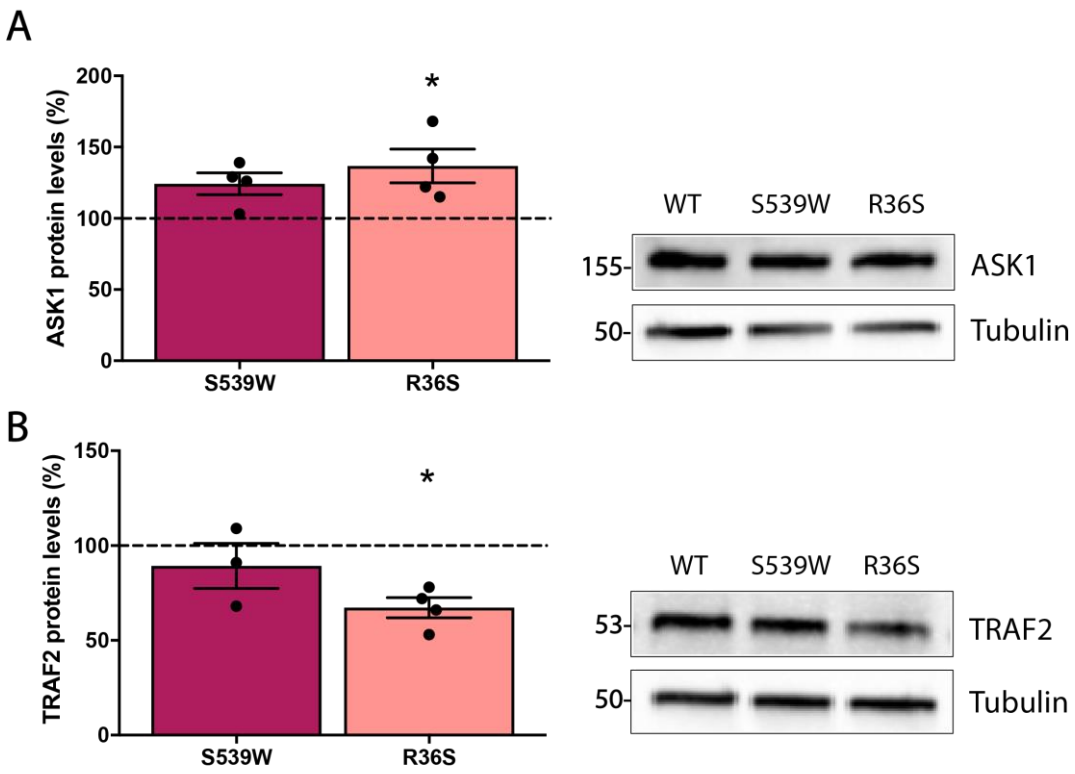

FIG. 4.3.9.2 R36S-PAM results in activation of a pro-apoptotic signalling cascade downstream of IRE1 activation. Cells were examined for the abundance of ASK1 (**A**) and TRAF2 (**B**) by western blot following transfection with membrane-integral PAM. These markers are involved in the pathway described in FIG. 4.3.8.1. Bands were quantified by lane densitometry and normalised to WT-transfected cells (dotted line). Analysis by one-way ANOVA on transformed (log) values, comparing either S539W or R36S to WT values. Data are representative of 3-4 independent experiments and error bars are SEM. * P<0.05.

4.3.10 R36S-PAM leads to reduced cell culture proliferation

Given the increase in the abundance of both UPR and pro-apoptotic markers following transfection of membrane-integral R36S-PAM, there may be a consequent loss of cell viability and proliferation. To address this, I examined the abundance of nuclear antigen Ki-67 (Ki67), which is used as a marker for cellular proliferation (Sales Gil and Vagnarelli 2018). However, I observed no impact of either R36S or S539W on proliferation compared to WT at 72hr post-transfection (FIG. 4.3.10.1).

I next investigated the caspase-3 activity in cell lysates. Caspase-3, a terminal effector caspase responsible for initiating apoptosis, would be activated downstream of ASK1-TRAF2 and can thereby be used as a marker of apoptosis activation (Porter and Janicke 1999). At 72hr post-transfection, I had observed elevated ASK1 and reduced TRAF2 levels – consistent with a pro-apoptotic cellular fate. Despite this, there was no difference observed in the caspase-3 activity for R36S- or S539W-transfected cells compared to WT-transfected cells (R36S; +38%, n=3, P=0.6; S539W: -28.3%, n=3, P=0.18) (FIG. 4.3.10.2). Interestingly, R36S-PAM and S539W-PAM do appear to mediate trends towards different directions of effect of caspase-3 activity.

Despite the observations that neither R36S nor S539W appear to affect apoptosis or proliferation at 72hr post-transfection, the elevated abundance of UPR and pro-apoptotic ASK1-TRAF2 markers for R36S still suggest a negative impact of R36S on cellular function and survival. I hypothesised that by looking at a general measure of cell viability, I may have greater resolution to pick up differences between WT-PAM transfected and S539W- or R36S-PAM transfected cells. In addition, this method may

identify a progressive loss of cell viability, which I cannot detect by looking at one time point alone.

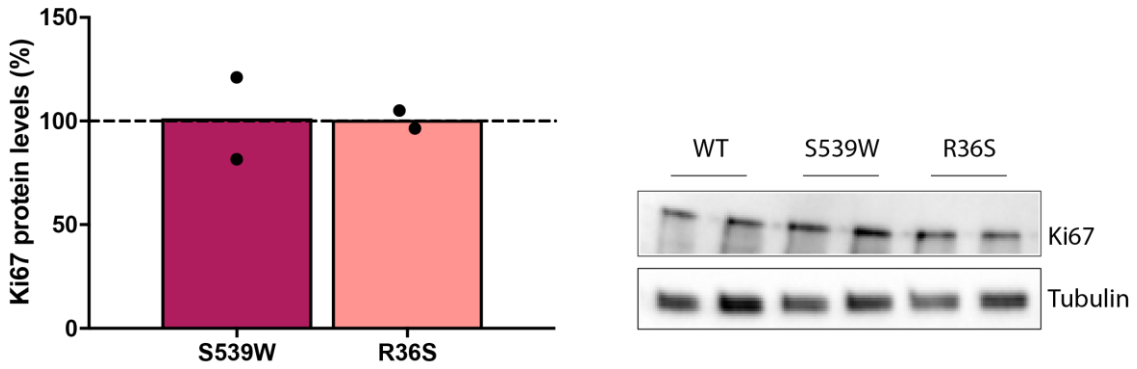

FIG. 4.3.10.1 R36S does not affect cellular proliferation. Cells transfected with membrane-integral PAM were examined for Ki67 abundance using western blot analysis. Bands were quantified by lane densitometry and intensities normalised to WT-transfected cells (dotted line). Data are representative of 2 independent experiments.

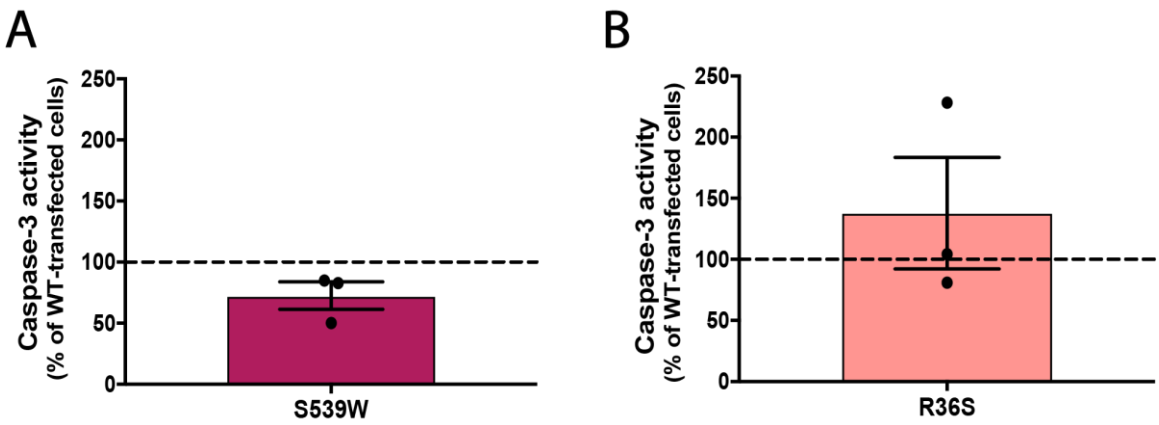

FIG. 4.3.10.2 Caspase-3 activity is not affected by either R36S-PAM or S539W-PAM. PAM KO cells were transfected with membrane-integral PAM for 72hr, prior to lysates being examined for caspase-3 activity by the EnzChek Caspase-3 Z-DEVD-AMC assay kit (Thermo Fisher). Reads for S539W (**A**) or R36S (**B**) were normalised to the WT-condition (dotted line) and analysed by unpaired Student's t-test with Welch's correction on transformed (log) data. Graphs are representative of 3 independent experiments and error bars are SEM.

I performed CyQUANT (Invitrogen) staining at 0 days, 3 days (72hrs) and 6 days post-transfection of either membrane-integral PAM or empty vector (EV) in PAM KO EndoC-βH1 cells. At day 6, I observed that R36S-PAM results in a trend towards reduced cell proliferation compared to WT-transfected cells (-32.4%, n=3, P=0.09), whereas S539W has no impact on cell viability at compared to WT-transfected cells (-13.4%, n=3, P=0.48) (FIG. 4.3.10.3). At day 3, there were no differences for either R36S-PAM or S539W-PAM compared to WT-PAM transfected cells (R36S: P=0.36; S539W: P=0.82).

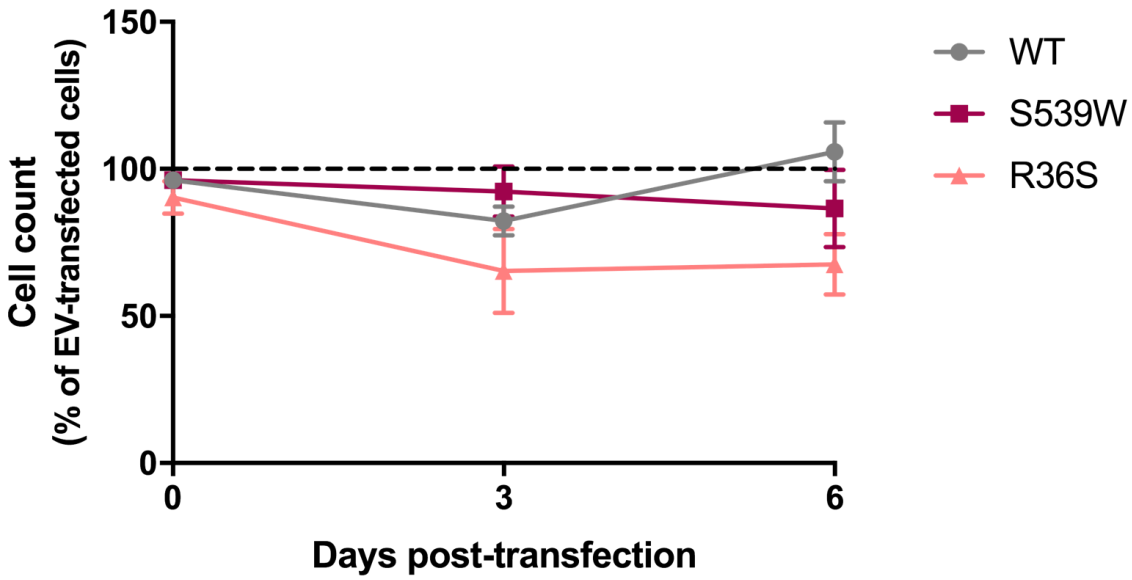

FIG. 4.3.10.3 R36S-PAM results in reduced cell growth in CRISPR PAM KO EndoC-βH1 cells. Cells were transfected with either Empty vector (EV) or membrane-integral PAM (WT, S539W, R36S) and CyQUANT performed at 0, 3, 6 days post-transfection. Raw fluorescent reads were normalised to EV-transfected cells (dotted line) within each time point. Transformed (log) normalised values (within each time point) have been compared to WT using one-way ANOVA. Data is representative of 3 independent experiments and error bars are SEM.

This data suggests that whilst UPR activation and a pro-apoptotic signalling cascade is initiated at 3 days (72hr) post-transfection, it takes a further 3 days to observe an impact on cellular viability and proliferation.

4.3.11 Impact of R36S-PAM on β cell secretory capacity

Much of the characterisation thus far has been surrounding the impact of R36S-PAM on cellular function and survival, but has not tackled a direct impact on insulin secretion. I observed previously that depletion of PAM function resulted in trends towards reduced insulin secretion and content (chapter 3), so there may be a direct impact of R36S on insulin secretion. To address this, I performed insulin secretion assays and examined the (intracellular) insulin content-adjusted secretion measures between variants (FIG. 4.3.11.1). In GFP-transfected cells, which represent the baseline condition (PAM KO), there was minimal glucose-stimulated insulin secretion (GSIS) between 2.8mM and 16.7mM glucose (+41%, n=9, P=0.86). The stimulation index was rescued by introduction of PAM in WT-transfected cells (+90%, n=9, P=0.058). By contrast, overexpression of S539W-PAM did not restore GSIS and a glucose-blunted response was observed (+13%, n=3, P=0.99), which was similar to GFP-transfected cells representing the baseline PAM KO condition and has been previously observed (Thomsen, Raimondo et al. 2018). Surprisingly, overexpression of R36S-PAM also restored the glucose stimulation index (+127%, n=6, P=0.016) in a manner similar to WT-transfected cells. There were no differences between genotypes under maximally stimulated conditions (16.7mM glucose + 100µM tolbutamide + 10 µM Forskolin), which reflects the maximally secretory capacity of the cell.

From these data, it is apparent that S539W-PAM and R36S-PAM can modulate insulin secretion by different mechanisms, despite sharing similar cellular localisation phenotypes in the β cell. This deviation is consistent with many of the cellular mechanisms characterised: S539W-PAM and R36S-PAM results in different directions and magnitude of effect. However, I have not pursued investigations to characterise exactly how or why these differences occur.

As observed with cell viability, I considered whether 72hr was sufficient time to observe the full effects of these variants PAM proteins on insulin content. To address this, I harvested cells at 6 days post-transfection to elucidate differences in the intracellular insulin content between WT-PAM and S539W- or R36S-PAM transfected cells. Mean insulin content was adjusted to cell count data (displayed in FIG. 4.3.10.3) to give an estimation of insulin content per cell and normalised to empty vector (EV) transfected cells to control for seeding discrepancies. Insulin-containing secretory granules are stored following packaging and maturation, and so a modest effect may not be visible in the short-term.

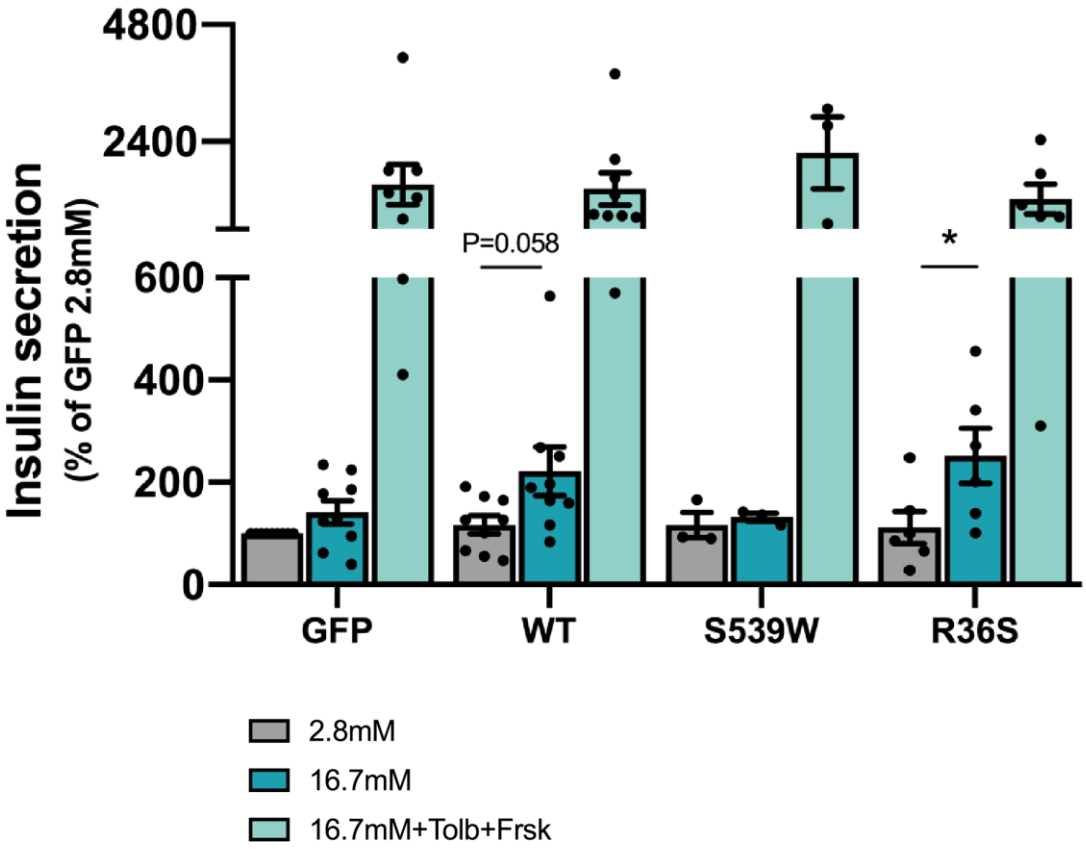

FIG. 4.3.11.1 Insulin secretion is affected differently by membrane-integral R36S- and S539W-PAM. Insulin secretion assays at 72hr post-transfection demonstrate that S539W and R36S have very different effects on insulin secretion, despite similar cellular localisation phenotypes. (**A**) Insulin secretion values were adjusted for intracellular content, then normalised to GFP-transfected cells at 2.8mM glucose within each experiment and have been plotted as mean percentages across the experiments. Data are representative of 3-9 independent experiments, consisting of 3-6 technical replicates within each experiment. Statistical analysis performed by 2-way ANOVA and error bars are SEM. * denotes P>0.05. GFP: green fluorescent protein; WT: wild-type, Tolb: Tolbutamide (100μM); Frsk: Forskolin (10μM).

I observed that there were no significant differences in insulin content between R36S- or S539W-PAM transfected cells compared to WT-PAM transfected cells (R36S: +8%, n=3, P=0.92; S539W: +20%, n=3, P=0.76). However, I did observe a reduced insulin content in WT-PAM transfected cells compared to empty vector (EV) transfected cells (-29.9%, n=3, P=0.03) (FIG. 4.3.11.2). This suggests there may be an additional mechanism at play in the "rescue" state, suggesting a role for the residual C-terminal fragment in insulin content in PAM KO EndoC-βH1 cells; although this has not been characterised further. Therefore, it is difficult to distinguish an impact of R36S- or S539W-PAM on the insulin content as compared to WT-PAM in this model.

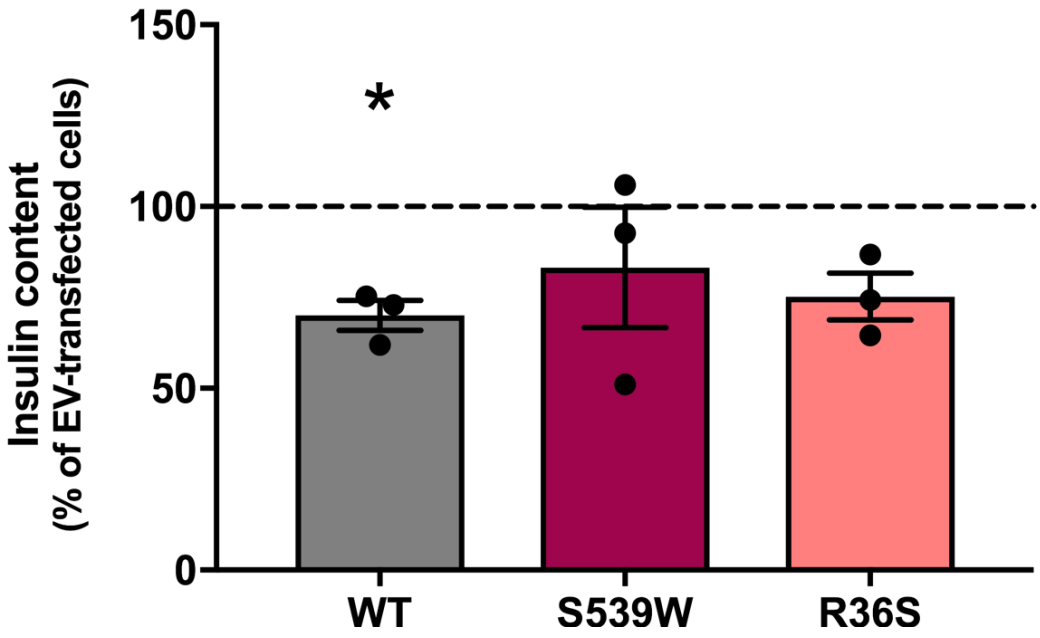

FIG. 4.3.11.2 Insulin content assayed at 6 days post-transfection of membrane-integral PAM in CRISPR PAM KO EndoC-βH1 cells. Raw insulin content measures are adjusted for cell counts (CyQUANT), normalised to content in EV-transfected cells (dotted line) and plotted as mean values (%) across experiments. Analysis by one-way ANOVA on transformed (log) data, comparing to EV-transfected cells (100%). Data is representative of 3 independent experiments and error bars are SEM. WT: wild-type; EV: empty vector. * denotes P<0.05.

4.4 DISCUSSION

In this chapter, I have performed functional characterisation of a novel, *de novo* candidate causal variant, p.R36S in the T2D-associated gene *PAM*, identified in a patient with early onset diabetes mellitus (DM). This characterisation was performed predominantly in PAM KO EndoC-βH1 cells, which were generated through CRISPR-Cas9 genome editing using 2 sgRNAs. Validation of this cell line identified that, whilst the predominant membrane-integral isoform had been successfully deleted, there was a residual expression of a short, C-terminal fragment. This fragment has not been functionally characterised, but does lack both catalytic domains (PHM and PAL). Therefore, these cells are not true PAM KOs and this is important to consider when interpreting my data.

Exome sequencing of a family trio, where the parents are unaffected, identified 4 candidate causal *de novo* variants in the proband (Table 4.3.2.2). *PAM* holds an established causal role for T2D risk, harbouring 2 independent coding variant signals of different effect sizes and a gene burden signal for T2D risk (Huyghe, Jackson et al. 2013, Steinthorsdottir, Thorleifsson et al. 2014, Fuchsberger, Flannick et al. 2016, Flannick, Mercader et al. 2019). This provides strong support for R36S-PAM to be contributing to diabetes in this proband. Cellular studies using CRISPR knockout of *CDH2* identified no conclusive defects to insulin secretion in Ins1 cells, suggesting that Q448*-CDH2 is unlikely to be driving diabetes causality in this proband. The variants in *CDH3* and *DOCK3* have not been functionally characterised, but they contain no putative links to β cell function or insulin secretion. Without further characterisation of these other *de novo* variants, I cannot rule out their relative contribution to the

pathogenesis in this proband. It is also not possible to exclude the possibility that early onset DM in this proband is caused by a non-coding regulatory element, since this has not been investigated.

The proband presented with no extra-pancreatic clinical features (personal communication, Dr. Hanna Huopio, University of Eastern Finland), and as such the pathological mechanisms are likely to be restricted to the pancreatic islet. The proband presented at 8 months of age, which is later than classically described for NDM; this is consistent with the normal birth weight observed suggesting normal foetal development *in utero*. Cases of NDM presenting at >6 months of age have been documented, and the absence of autoantibodies detected in this proband indicate that permanent early onset/neonatal diabetes mellitus is still a likely diagnosis (Rubio-Cabezas, Flanagan et al. 2012). However, the presentation at >6 months of age does increase the possibility of T1D, and this should be further investigated (Edghill, Dix et al. 2006). Application of genetic risk scores for T1D, which take into account more than HLA risk genotypes, would be beneficial in distinguishing the likelihood of T1D compared with monogenic diabetes in this proband (Patel, Oram et al. 2016, Perry, Wasserfall et al. 2018, Sharp, Rich et al. 2019).

Unlike known T2D-associated variants in *PAM*, R36S sits outside the catalytic domains and is located between the N-terminal ER signal peptide and the proregion (Mains, Milgram et al. 1995, Huyghe, Jackson et al. 2013, Steinthorsdottir, Thorleifsson et al. 2014, Thomsen, Raimondo et al. 2018). This gives the first indication that any altered function mediated by R36S is likely to occur through mechanisms, which are distinct

from those at play for the T2D-associated variants in *PAM*. Many T2D susceptibility loci harbour variants that are causal for monogenic diabetes (e.g. *WFS1,* wolframin; *LMNA,* Lamin A) (Sandhu, Weedon et al. 2007, Wegner, Andersen et al. 2007). Characterisation of these rare variants can indicate the role for that protein in the β cell and help to distinguish the mechanisms that dictate the difference in clinical severity between T2D risk and monogenic diabetes pathogenesis.

I initially characterised the impact of the R36S substitution on PAM protein function as it is currently understood. Characterisation of known T2D-associated variants, D563G and S539W, identified defects in PAM catalytic activity and protein expression respectively so I began by using a similar approach to characterise R36S (Thomsen, Raimondo et al. 2018). I observed a mild reduction in the serum amidation profile for both the proband and her unaffected father. This is at a level similar to the heterozygous carriers of the D563G allele, but since neither the proband nor father are carriers of either T2D risk allele, there must be additional (unreported) *PAM* alleles present in both individuals mediating this effect.

Whilst the serum amidation assay provides a physiological measure of amidating capacity, it is not possible to determine the PAM content in the serum samples. Therefore, I cannot distinguish between impaired activity and altered PAM abundance. In addition, there are multiple tissues in which PAM may be secreted from (e.g. pituitary) and there are no studies confirming PAM is secreted from the pancreatic islet. As such, the serum amidating profile may not provide a direct readout of PAM function in the β cell.

However, the *in vitro* amidation assay, which generates amidating activity measures normalised to PAM content, also supports normal amidating activity for R36S-PAM, although this is only an n=1. Recombinant luminal PAM is generated through the stable transfection of HEK293 cells, which constitutively secrete luminal PAM into the supernatant. This is an efficient manner of generating recombinant luminal PAM in a mammalian expression system, but does not provide the tissue-specific processing present in β cells (Thiele, Marek et al. 1989, Eipper, Green et al. 1992). One key aspect is the expression of prohormone convertases such as PC1/3, which are expressed in neuroendocrine cells but not in HEK293 cells and are pertinent to the post-translational cleavage of PAM (Eipper, Green et al. 1992). Using information deduced from both amidation assays, it is possible to isolate the true effect of a variant on amidating activity. The D563G allele demonstrates nominal impact in the serum amidation assay, but displays ~25% reduction compared to WT-PAM in the *in vitro* assay (Thomsen, Raimondo et al. 2018). This suggests that D563G has a defect in amidating activity rather than expression. By contrast, the nominal reduction in amidating profile for R36S-PAM (proband) in the serum amidation assay is not explained by the *in vitro* amidation assay, which is normal. This suggests that R36S-PAM is more likely to be affected by a protein stability defect rather than a defect in catalytic activity.

I hypothesised that given its position; R36S was likely to have greatest impact on protein stability and/or the correct trafficking of PAM to the secretory pathway in β cells. I characterised the impact of R36S-PAM on the abundance and stability of

luminal and membrane-integral PAM isoforms in EndoC-βH1 cells. Whilst both isoforms should be trafficked to the secretory pathway and packaged into insulin-containing secretory granules, they are likely to be handled differently by the cell, owing to the C-terminal transmembrane domain in membrane integral PAM. Luminal PAM, which is soluble and so can be secreted, sits within the granular lumen whilst membrane-integral PAM straddles the granule membrane (Martinez, Montuenga et al. 1993, Steveson, Zhao et al. 2001, Francone, Ifrim et al. 2010). Therefore, there will be putative differences in how each isoform is trafficked by the cell. I identified a protein stability defect with luminal R36S-PAM, which resulted in degradation by the proteosome. This is likely driven through ER-associated degradation (ERAD), a physiological mechanism which acts to clear the ER lumen of misfolded proteins, although this was not investigated further (Meusser, Hirsch et al. 2005). As I observed efficient clearance of luminal R36S-PAM, it suggests that clearance through ERAD ensured that there was minimal retention of R36S-PAM within the ER and may result in no activation of the UPR. In line with this, I observed no elevations in any of the UPR markers for R36S-PAM transfected cells as compared to WT-PAM transfected cells. To confirm the presence of active ERAD, I would examine the abundance of ERAD machinery such as the E3 ubiquitin ligase HRD1 following transfection of luminal R36S-PAM (Jung, Hong et al. 2015).

In chapter 3, I identified that the major isoform of PAM in human β cells is membrane-integral (FIG. 3.3.1.1, 3.3.2.1). Luminal PAM, which in β cells would be generated through endoproteolytic cleavage at dibasic motifs by PC1/3, occurs in mature secretory granules. Therefore, luminal PAM is unlikely to be exposed as an

independent isoform until late into the secretory pathway. It is therefore much more informative to characterise the impact following transfection of membrane-integral R36S-PAM on cellular dysfunction and disease pathogenesis in β cells. As such, the majority of my functional characterisation centres on membrane-integral R36S-PAM.

Membrane-integral R36S-PAM, by contrast, displayed no difference in PAM abundance compared to WT-PAM, but did show an ER-like staining pattern in fixed EndoC-βH1 cells suggesting that it may be ER-retained. This is similar to the T2D-risk allele S539W in *PAM* (Thomsen, Raimondo et al. 2018). Efforts to further characterise this ER-retention using transmission electron microscopy were unsuccessful as antibodies targeting the C-terminal Myc- and FLAG-tags did not provide any specific labelling. In addition, given the low transfection efficiency in EndoC-βH1 cells, it was difficult to identify transfected cells; as such, this was not pursued further. One option would be to sort cells using fluorescence-activated cell sorting (FACS) to enrich for transfected cells prior to processing/immunogold labelling.

Since both R36S-PAM and S539W-PAM demonstrated similar cellular localisation phenotypes, I hypothesised that there were specific mechanisms further to ER-retention that dictated the severity of cellular dysfunction and associated clinical phenotype. Since I observed no reduction in PAM abundance for R36S, I considered that the accumulation of PAM in the ER may result in activation of the UPR. I observed phospho-activation of IRE1 and a trend towards increased abundance of CHOP in R36S-PAM transfected cells compared to WT-transfected cells. There was no elevation in any of these UPR markers for S539W-transfected cells, suggesting that, despite

being similarly ER-retained, the two variant PAM proteins could be handled very differently by the ER, thereby resulting in different mechanisms and degrees of cellular dysfunction.

I speculated that hyper-activation of the UPR by R36S-PAM may be prolonged, which may in turn lead to pro-apoptotic signalling downstream of IRE1 (FIG. 4.3.8.1). I observed elevations in the abundance of ASK1 in R36S-PAM transfected cells. ASK1 is a pro-apoptotic kinase involved in the phosphorylation of JNK and the subsequent activation of caspase-12 in rodents (Yoneda, Imaizumi et al. 2001, Nishitoh, Matsuzawa et al. 2002). I also observed a reduction in TRAF2 abundance for R36S-PAM transfected cells (FIG 4.3.9.2), which has been shown to occur under ER stress and been linked to the activation of apoptosis through dampening of the NF-κβ pathway; although this is largely characterised in type 1 diabetes (Tang, Minemoto et al. 2001, Chang, Kim et al. 2003, Hu, Han et al. 2006). In addition, TRAF2 can function as an ASK1 inhibitor and when downregulated enables ASK1 to be activated, thereby promoting the phosphorylation of JNK (Tang, Minemoto et al. 2001, Hu, Han et al. 2006). These data collectively suggest that expression of membrane-integral R36S-PAM for 72hr results in hyper-activation of the UPR and a pro-apoptotic signalling cascade consistent with a cell fate travelling towards apoptosis. The activation of ASK1 and downregulation of TRAF2 are likely to favour the switch to a pro-apoptotic cellular fate. However, additional mechanisms might be at play that I have not surveyed, such as the production of reactive oxygen species (ROS) or pro-inflammatory cytokines (e.g. tumour necrosis factor, TNFα) as terminal effectors of β cell death (Chang, Kim et al. 2003, Brezniceanu, Lau et al. 2010).

I did not observe significant differences in either the caspase-3 activity or abundance of Ki67 for R36S-PAM as compared to WT-PAM transfected cells at 72hr post-transfection. Caspase-3 activity has been previously assessed in EndoC-βH1 cells at a similar time point (Tsonkova, Sand et al. 2017). This is compounded by a lack of resolution from the low transfection efficiency, but may also reflect a time delay between induction of pro-apoptotic signalling and the activation of terminal effectors such as caspase-3. This hypothesis is supported by the clear deviation in cell viability for R36S-PAM transfected cells as compared to WT- or S539W-PAM transfected cells observed at 6 days but not at 3 days (72hr) post-transfection. Whilst length of IRE1 induction is important in dictating the type of cellular response mediated, it has been shown that JNK activation at 12hr or later following UPR activation is critical in determining whether apoptosis is initiated (Zhang, Kawauchi et al. 2001, Arshad, Ye et al. 2013, Brown, Strudwick et al. 2016). Follow-up experiments are necessary to measure the caspase-3 activity and Ki67 abundance day 6 to determine whether these effectors contribute to the reduced cell viability observed for R36S-PAM at this time point. In addition to caspase activation, the IRE1α-ASK1-JNK pathway has been proposed to phosphorylate the pro-apoptotic mitochondrial proteins, BAX and BAK (Hetz, Bernasconi et al. 2006). Based on the evidence I have shown in this chapter, my proposed mechanism for R36S-mediated cellular dysfunction underlying NDM pathogenesis in the β cell is working largely through ER stress, apoptosis and reduced β cell viability mechanisms (FIG. 4.4.1).

Despite robust evidence suggesting that membrane-integral R36S-PAM results in ER-stress, I have not observed convincing evidence to suggest that R36S-PAM impacts insulin secretion or that it is causal for neonatal diabetes in a systemic model. Investigation of R36S-PAM in insulin secretion and content in PAM KO EndoC-βH1 cells, showed a trend towards an increased fold change (FC: 2.54) between basal (2.8mM) and stimulated (16.7mM) glucose conditions as compared to WT-transfected cells (FC: 1.96). The resolution is lacking to fully characterise this effect in an isolated cellular model: the low transfection efficiency means the majority of cells are untransfected and therefore will dilute out any effect making them difficult to separate from background.

In collaboration with MRC Harwell, a p.R36S (p.R40S in mouse) global knock-in mouse is being generated, which should answer whether global expression of R36S-PAM is pathogenic to the extent of causing a glucose-intolerant diabetic phenotype consistent with the clinical presentation in this human proband. It will also allow investigation of whether the variant leads to β cell loss and diabetes. The identification and characterisation of extra-pancreatic phenotypes can also be explored in this model, but it will be important to consider the species differences when comparing phenotypes between the mouse model and human proband. Finally, the identification of a second family with an additional R36S-PAM carrier is required for confidence of *PAM* as a cause of NDM. This would also provide further clarification on the relative contribution of genetic background to the causal mechanisms underlying R36S-PAM mediated NDM pathogenesis.

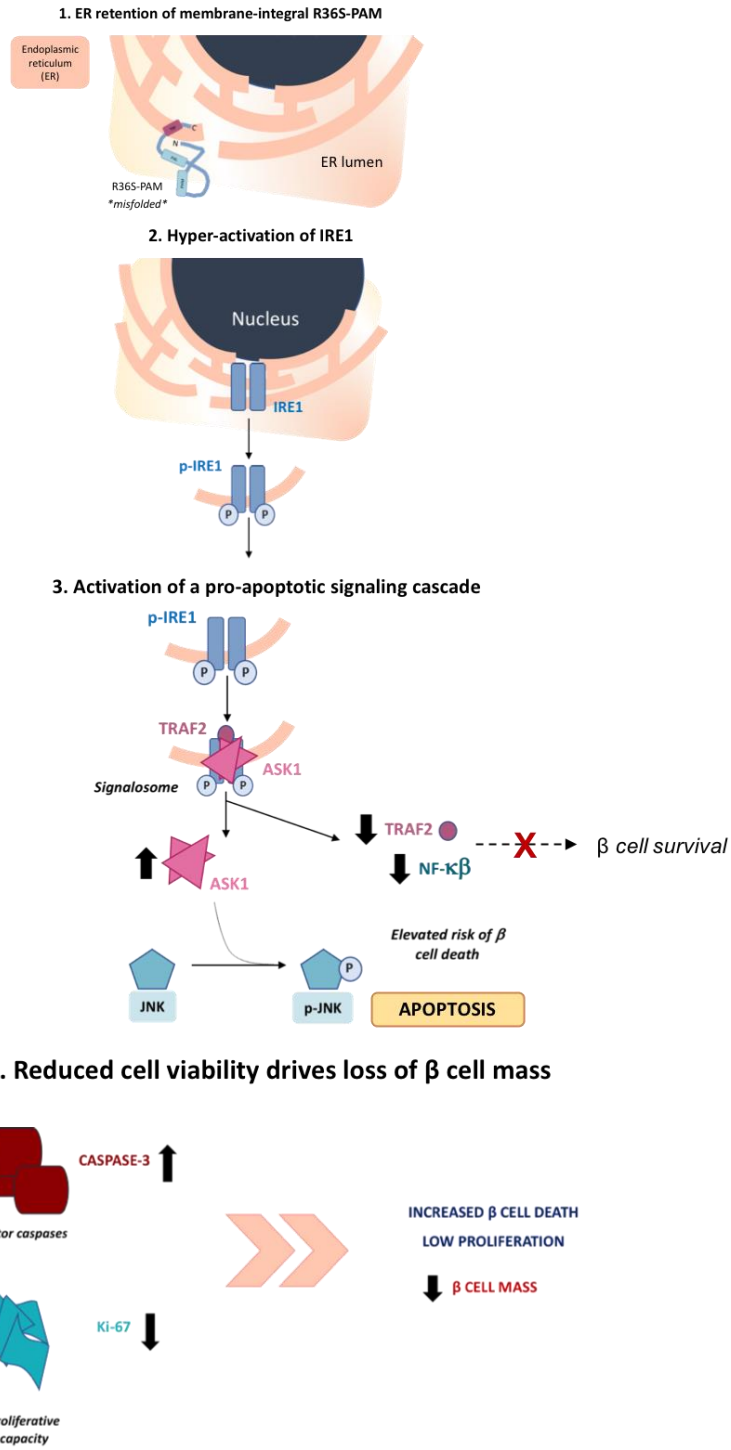

FIG. 4.4.1 Proposed mechanism for R36S-mediated cellular dysfunction underlying early onset diabetes pathogenesis. Mechanism can be divided into 4 main stages of dysfunction: (**1**) ER-retention of misfolded membrane-integral R36S-PAM, (**2**) activation of the UPR, (**3**) switch to pro-apoptotic cellular fate, (**4**) Reduced cell viability leading to loss of β cell mass in cell line and NDM. This mechanism is proposed and therefore further investigation is required to confirm whether all components play a role in the pathogenesis in this proband.

CHAPTER 5

Establishing a strategy for the characterisation of rare coding *PAM* variants in T2D risk

5.1 INTRODUCTION

5.1.1 Gene-level associations for complex disease risk

Genetic studies in complex disease seek to characterise the heritability underlying disease risk (Manolio, Collins et al. 2009). Rare coding variants have been proposed to harbour a higher proportion of disease heritability than their common counterparts (McClellan and King 2010, Tennessen, Bigham et al. 2012). According to the common disease-rare variant hypothesis, highly penetrant alleles are likely to be rare in frequency, and are perceived to be a better pharmacological target than common variants with modest effect sizes (Eyre-Walker 2010, Zuk, Schaffner et al. 2014, Saleheen, Natarajan et al. 2017).

Genome wide association studies (GWAS) and imputation strategies are underpowered to detect rare variants, unless the samples size is infinite (Zuk, Schaffner et al. 2014). Whole genome or exome array-based sequencing studies, which screen for known variants at a large scale, can provide means of identifying rare, highly penetrant disease-associated variants (Ng, Turner et al. 2009, Kiezun, Garimella et al. 2012, Fuchsberger, Flannick et al. 2016, Jun, Manning et al. 2018, Flannick, Mercader et al. 2019). However, most variants are too rare (MAF<0.05%) to be individually associated with disease risk. To circumnavigate this, multiple variants can be collapsed into a single statistic to assess the aggregate or gene-level association across the gene (Moutsianas, Agarwala et al. 2015, Guo, Plummer et al. 2018). This circumnavigates the low absolute power but does also increase the likelihood of false positives (Moutsianas, Agarwala et al. 2015).

Exome sequencing of 20,791 T2D cases and 24,440 controls spanning 5 ancestries (Hispanic/Latino, European, African-American, East Asian and South Asian) identified a gene-level association signal for T2D risk in *PAM* (Flannick, Mercader et al. 2019). This is contributed to by the known independent T2D-risk alleles and 35 novel rare coding alleles. The gene-based signal even remained nominally significant (P<0.05) following conditional analyses for these 35 rare variants and the aforementioned independent T2D-coding variants, emphasizing the extent to which *PAM* is involved in T2D risk. *PAM* is one of only three gene-based signals identified in association with T2D risk (along with *SLC30A8*, *MC4R*) (Flannick, Mercader et al. 2019).

5.1.2 Characterisation of multiple, rare coding variants in *PAM*

Multiple coding alleles spanning one gene provide the opportunity to interrogate the relationship between altered gene function and disease risk. That is, the functional characterisation of multiple variants at different penetrance, direction of effect (e.g. LoF or GoF) and therefore causal functional mechanisms. Elucidation of variants along an allelic series can therefore provide a more comprehensive understanding of the how modulation of gene function affects protein function and disease risk. This is exemplified by characterisation of variants in *MNTR1B* and *SLC30A8,* which provided further insight into this gene-dose relationship than by characterisation of one variant alone (Bonnefond, Clement et al. 2012, Flannick, Thorleifsson et al. 2014).

Functional characterisation is notoriously low throughput and it is not feasible to characterise multiple variants without strong genetic evidence for their contribution to disease risk. The known T2D-risk alleles in *PAM*, D563G (MAF=4.98%) and S539W

(MAF=0.65%), are independently associated with increased T2D risk and reduced insulinogenic index (Huyghe, Jackson et al. 2013, Steinthorsdottir, Thorleifsson et al. 2014). This means that their role in T2D risk is supported by strong genetic evidence and these alleles are therefore likely to affect PAM function. This was confirmed by functional characterisation performed in human β cells (Thomsen, Raimondo et al. 2018). However, the 35 rare variants identified through the gene-based signal are not individually supported by strong genetic evidence, since their association with T2D risk is a cumulative measure of their aggregate risk (Flannick, Mercader et al. 2019). This means that without comprehensive characterisation of all variants, it is difficult to distinguish deleterious/protective variants from neutral variants. To do so, requires a robust phenotyping strategy capable of elucidating the key aspects of PAM function.

5.1.3 Experimental aims

In this chapter, I perform a pilot study to characterise 5 variants shortlisted from the gene-level association signal for T2D risk in *PAM.* I am aided by my previous work in this thesis to inform on the relevant mechanisms in which PAM function is likely to be key for β cell function and may be affected. This study will establish a strategy for the comprehensive characterisation of rare coding variants (as an allelic series) underlying T2D risk in *PAM*. In this pilot study, I have shortlisted a subset of variants and pursued the following experimental avenues:

- Impact of variants on abundance of membrane-integral PAM in EndoC-βH1 cells and HEK293 cells (non-neuroendocrine control)
- Impact on protein stability through modulation of the proteasomal pathway
- Cellular localisation in PAM KO EndoC-βH1 cells

- Impact on insulin secretion and content in PAM KO EndoC-βH1 cells

5.2 METHODS

Since many of the methods have been described in previous chapters, only those unique to this chapter are described below. Other sections can be found in:

- Generation and validation of PAM KO EndoC-βH1 cells (section 2.5, 4.3.1)
- Routine maintenance of HEK293 and PAM KO EndoC-βH1 cells (sections 2.6, 4.2.2)
- Transfection of cells (section 2.7) and harvest using RIPA buffer (section 2.9)
- Protein degradation assays using MG132 (section 4.2.4)
- Western blotting (section 2.9), primary antibodies (Table 4.2.3.1) and secondary antibodies listed (Table 2.9.4.2)
- Immunofluorescence (section 2.11), primary antibodies (Table 4.2.3) and secondary antibodies (Table 2.11.1.2)
- Insulin secretion assays (sections 2.19, 4.2.10)

5.2.1 *In silico* analysis

Shortlisted variants for study were subject to analysis by SIFT, PolyPhen2 and Condel, as described in section 2.17. Variants with predicted pathogenic impact in all three analyses were prioritised for study.

5.2.2 Generation of constructs

Constructs containing membrane-integral PAM for each shortlisted variant were generated in collaboration with Dr. Fernando Abaitua (WTCHG, University of Oxford) using site directed mutagenesis of pCMV6 wild-type (WT) membrane-integral PAM

constructs as detailed in section 2.4. Primers used for the generation of each variant are listed in Table 5.2.2.1. V803M was shortlisted as one of the variants for study, but was removed following technical challenges in performing the site-directed mutagenesis and in the interests of time. Sanger sequencing was performed to ensure that only the desired nucleotide changes were introduced. Several control constructs to elucidate the impact of these shortlisted variants on PAM protein function, namely WT-PAM, S539W-PAM (low-frequency T2D-risk allele) and GFP as a marker of transfection efficiency (Table 5.2.2.2).

Variant (p.)	Primer sequence (5' to 3')
A137V	GGGGGAGCATTTCTC**A**CCCAGGCATACAGAA
R144Q	CCGGCTCCCCAAA**A**GTGTTGGATTCAG
R260W	GAGGGCTCTGCC**A**TCCAATCAGTGTCCACT
G531C	ATCCCAGACATGGTCACATC**T**GTGGAAAATCACCA
R776C	GCATATCAAAGTGCT**T**GCACACTGGCTTGAAGATGTC
V803M	**GCTCATACCAACACC<u>A</u>TGTGGAAGTTCACC**

Table 5.2.2.1 Primer sequences to generate variant PAM construct using site-directed mutagenesis. Primers were purchased from Eurofins Genomics. Nucleotide changes are highlighted in **red**. V803M (**bold**) was initially shortlisted for study, but was excluded following challenges with site-directed mutagenesis.

Construct
pCMV6-hPAM WT
pCMV6-hPAM S539W
pCMV6-GFP
pCMV6-hPAM A137V
pCMV6-hPAM R144Q
pCMV6-hPAM R260W
pCMV6-hPAM G531C
pCMV6-hPAM R776C

Table 5.2.2.2 Constructs used to assess impact on PAM function. All PAM constructs encode membrane-integral PAM; pCMV6-GFP was used as a marker of transfection efficiency. S539W is a known low frequency, T2D-associated allele that has been included as a positive control for some of the functional analyses. WT: wild-type; GFP: green fluorescent protein. All constructs were verified by Sanger sequencing.

5.2.3 Statistical analysis

Unless otherwise stated, data were normalised to control samples (WT-PAM transfected cells) within each experiment and collated as mean normalised values (%) across an experiment. For statistical analysis, mean data points were transformed into log and analysed using the one-way ANOVA when comparing one variable or the two-way ANOVA when there were multiple variables. Statistical analysis used per experiment is listed within the figure legend. For ease, data have been plotted as the normalised averages (%) rather than log values.

5.3 RESULTS

5.3.1 Shortlisting of variants for study

In addition to the two known independent variants, D563G and S539W, the gene-level association signal for T2D risk in *PAM* uncovered a >35 rare coding variants (MAF<0.1%) (Huyghe, Jackson et al. 2013, Steinthorsdottir, Thorleifsson et al. 2014, Flannick, Mercader et al. 2018).

I prioritised the variants of interest on the basis of their association with T2D risk, using frequency in T2D cases as compared to control. The odds ratio (OR) is not an accurate measure of penetrance at this rare allelic frequency (Table 5.3.1). I next selected variants which were located in the catalytic domains. My characterisation followed an approach similar to that used in the functional characterisation of D563G and S539W, which are both located in the PAL catalytic domain (Thomsen, Raimondo et al. 2018). I found that there were 24 variants that were enriched in T2D cases rather than controls, of which 12 are located in the catalytic domains of PAM (Table 5.3.3.1).

Variant	Allele		Position of variant	Frequency of minor allele	
	Reference	Minor		Case	Control
p.H240Y	C	T	Catalytic - PHM	8	1
p.I52V	A	G	Proregion	4	0
p.G531C	G	T	Catalytic - PAL	5	1
p.G728R	G	C	Catalytic - PAL	13	6
p.V803M	G	A	Catalytic - PAL	13	3
p.G65E	G	A	Proregion	3	0
p.F31Y	T	A	Proregion	294	220
p.S113P	T	C	Catalytic - PAL	11	6
p.I737T	T	C	Catalytic - PHM	2	0
p.I452V	A	G,T	Exon A	5	2
p.S892L	C	T	C-terminus	4	1
p.N428Y	A	T	Exon A	3	1
p.S958L	C	T	C-terminus	3	1
p.R144Q	G	A	Catalytic - PHM	19	13
p.S395C	C	G	Exon A	2	0
p.R260W	C	T	Catalytic - PHM	3	1
p.S972F	C	T	C-terminus	8	5
p.T905M	C	T	C-terminus	3	1
p.E380K	G	*,A	Exon A	11	5
p.K396R	A	G	Exon A	5	3
p.R776C	C	T	Catalytic - PAL	6	4
p.P521L	C	T	Catalytic - PAL	18	13
p.A137V	C	T	Catalytic - PHM	2	0
p.V507A	T	C	Catalytic - PAL	2	0

Table 5.3.1.1 Variants that are enriched in T2D cases compared to controls. Variants identified from exome sequencing data generated in (Flannick, Mercader et al. 2019). Variants listed were selected based for their enriched frequency in T2D cases compared with controls, as a measure of likely pathogenicity. Those highlighted indicate variants also positioned in the catalytic domains and therefore prioritised for study. * denotes a premature stop codon resulting in a nonsense mutation at the same position (p.380) that also results in a missense variant. PHM: peptidylglycine α-hydroxylating monooxygenase; PAL: peptidyl α-hydroxyglycine α-amidating lyase. Exon A: linker region present between PHM and PAL catalytic domains. T2D: type 2 diabetes.

5.3.2 *In silico* analysis of shortlisted variants

To maximise the chances that my selected variants were likely to be deleterious to PAM function, I used the *in silico* tool SIFT, PolyPhen2 and Condel to assess for predicted pathogenicity. Of the 12 variants I had identified to be enriched in T2D cases and located in the catalytic domains, 6 variants were predicted to be deleterious to PAM function, which were prioritised for cellular characterisation (Table 5.2.3.1). However, due to technical challenges with the efficiency of the SDM reaction, only 5 of these were taken beyond site directed mutagenesis: A137V, R144Q, R260W, G531C and R776C.

Variant	SIFT		PolyPhen2		Condel	
	Prediction	Score	Prediction	Score	Prediction	Score
p.H240Y	tolerated	0.06	benign	0.182	neutral	0.339
p.G531C	**deleterious**	**0**	**Possibly damaging**	**0.925**	**deleterious**	**0.818**
p.G728R	tolerated	0.06	Possibly damaging	1	deleterious	0.823
p.V803M	deleterious	0	Possibly damaging	0.998	deleterious	0.919
p.S113P	tolerated	0.27	benign	0.205	neutral	0.054
p.I737T	tolerated	1	Possibly damaging	0.476	neutral	0.157
p.R144Q	**deleterious**	**0.04**	**Possibly damaging**	**0.915**	**deleterious**	**0.716**
p.R260W	**deleterious**	**0**	**Possibly damaging**	**0.996**	**deleterious**	**0.906**
p.R776C	**deleterious**	**0**	**Possibly damaging**	**0.91**	**deleterious**	**0.807**
p.P521L	tolerated	0.17	benign	0.311	neutral	0.156
p.A137V	**deleterious**	**0**	**Possibly damaging**	**0.99**	**deleterious**	**0.886**
p.V507A	tolerated	0.28	benign	0.404	neutral	0.162

Table 5.3.2.1 Prioritisation of variants for study using *in silico* prediction tools. Variants found to be enriched in cases rather than controls and located in the catalytic domains were subject to analysis by SIFT, PolyPhen2 and Condel. Those highlighted indicate variants predicted to be deleterious in all 3 analyses. Only those in **bold** have been taken forward for study. p. = position in protein sequence

5.3.3 Abundance of membrane-integral PAM isoforms in HEK293 cells

To investigate the impact of these variants on membrane-integral PAM, I transiently transfected HEK293 cells for 48hr and assessed the abundance of membrane-integral PAM using western blotting analysis and lane densitometry. I found that, whilst there were no significant differences in any of the variant PAM levels (FIG. 5.3.3.1), there was a trend towards reduced abundance of G531C-PAM as compared to WT-PAM (-34.5%, n=3, P=0.252).

I considered that this trend towards reduced G531C abundance may be due to differences in transfection efficiency. My constructs generate PAM-Myc-FLAG fusion proteins, which in themselves cannot be quantified for differences in transfection efficiency (between variants). To investigate this, I co-transfected HEK293 cells with pCMV6-PAM (WT or variant) constructs and an independent pCMV6-GFP construct at a ratio of 65:35, and then examined GFP protein abundance at 48hrs post-transfection as a measure of the transfection efficiency. To this end, I identified no significant differences in the transfection efficiency of any variants as compared to WT-PAM transfected cells (A137V: P>0.999; R144Q: P=0.997; R260W: P>0.999; G531C: P=0.626; R776C: P>0.999) (FIG. 5.3.3.2). Whilst it is not statistically significant, there is a modest impact on transfection efficiency for G531C (-25.4%). This may explain the trend towards reduced PAM protein abundance observed in FIG. 5.3.3.1.

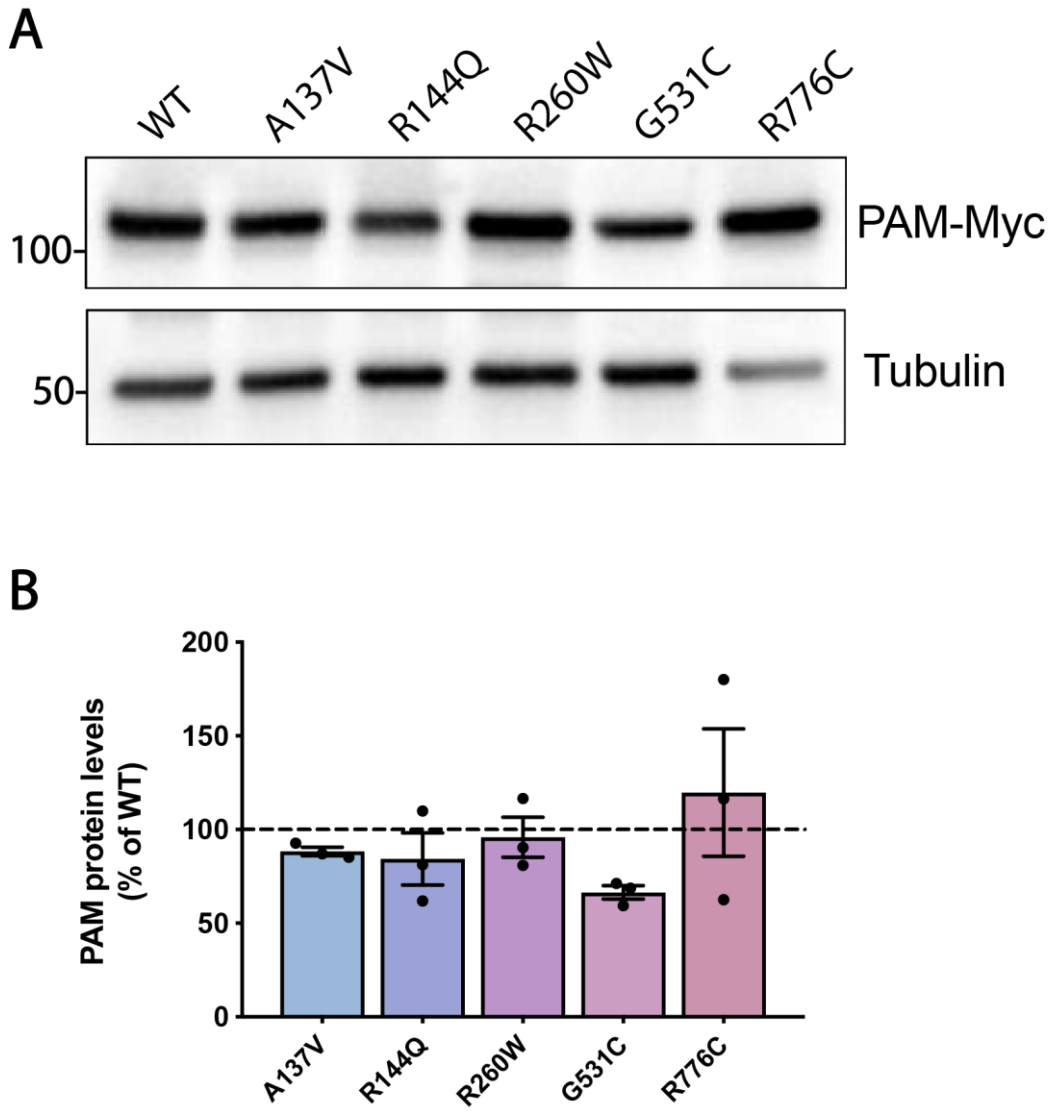

FIG. 5.3.3.1 Impact of rare *PAM* variants on membrane-integral PAM abundance in HEK293 cells. Transient overexpression of variants in HEK293 cells as assessed by western analysis blot (**A**) and quantified by lane densitometry (**B**). Band intensities were adjusted for tubulin intensities and normalised to wild type (WT) PAM (dotted line). Statistical analysis was performed by one-way ANOVA on transformed (log) values, data is representative of 3 independent experiments and error bars are SEM.

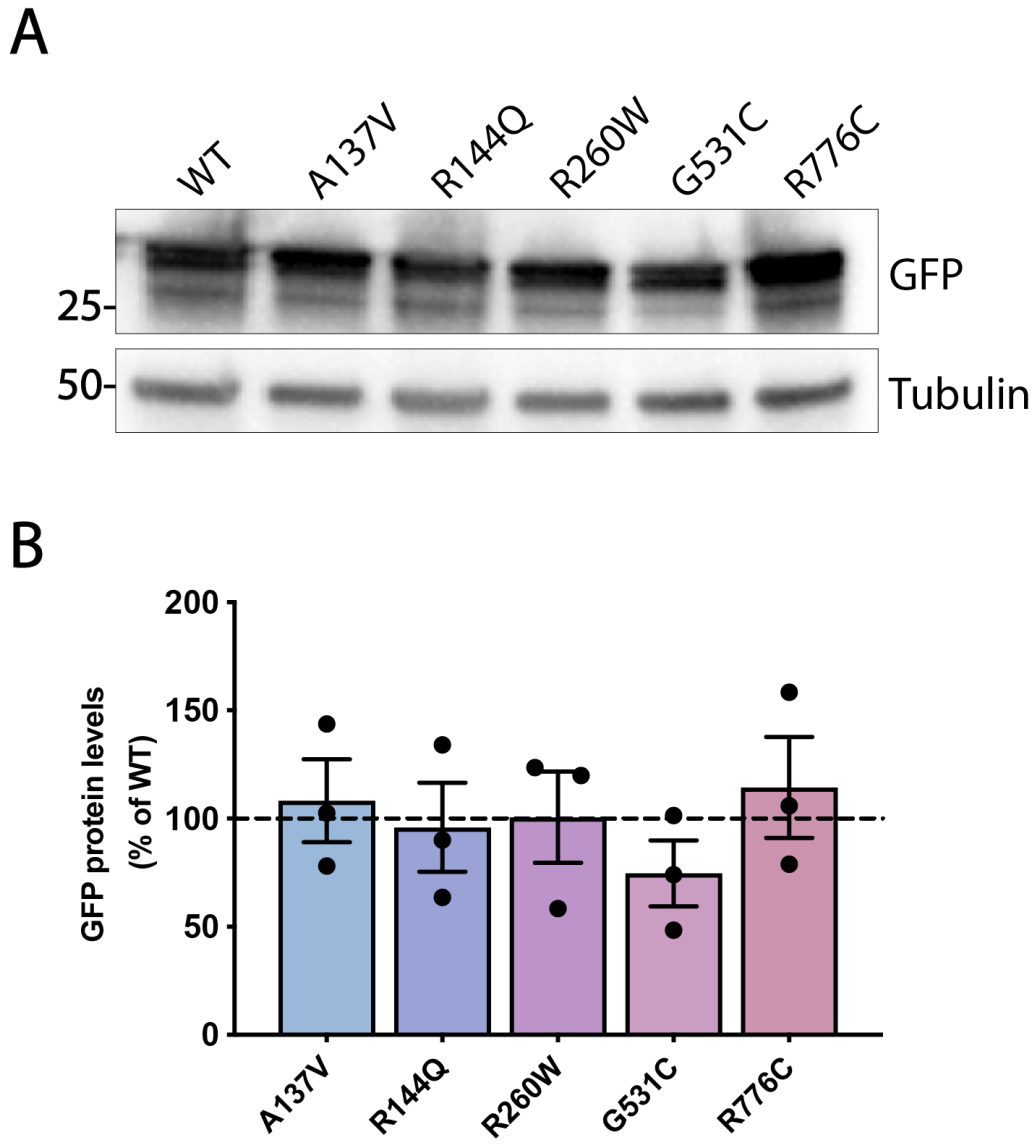

FIG. 5.3.3.2 Estimation of transfection efficiency for variants compared to WT-PAM. To gauge an estimate of the (differences) in transfection efficiency between the different variants and WT-PAM, HEK293 cells were co-transfected with pCMV6-PAM and pCMV6-GFP constructs in a 65:35 ratio (%) and GFP protein abundance level assessed by western blotting analysis (A). Band intensities were quantified using lane densitometry and normalised to WT-PAM (dotted line). Statistical analysis performed using one-way ANOVA on transformed (log) values. Data are representative of 3 independent experiments and error bars are SEM.

5.3.4 Modulation of the proteasomal pathway to investigate protein instability and degradation

My characterisation of *PAM* variants (R36S in chapter 4) highlighted the role that the proteasome plays in protein stability and biosynthesis. I transfected HEK293 cells with either WT-PAM or G531C-PAM, and then incubated cells with MG132 (inhibitor of the 26S subunit of proteasome) at 72hr post-transfection for a period of either 1hr or 3hr (or DMSO as a control). I found that whilst I observed non-significant trends towards an increased abundance of WT-PAM following incubation with MG132, there was no increase ('rescue') in the abundance of G531C-PAM following MG132 incubation (FIG. 5.3.4.1). In fact, there are significant differences in abundance between G531C-PAM (in all 3 conditions) and WT_DMSO (G531C_DMSO: -65.7%, P=0.008; G531C_MG132_1hr: -58.8% P=0.03; G531C_MG132_3hr: -70.2%, P=0.001) This suggests that the reduced abundance observed for G531C-PAM is not due to proteasomal degradation.

For both A137V and G531C, there are trends towards an increase in PAM abundance following 1hr of MG132 incubation. Incubation with MG132 for 3hr, has no impact on PAM abundance for either variant compared to control-treated cells (DMSO). This is a trend not observed for WT-PAM.

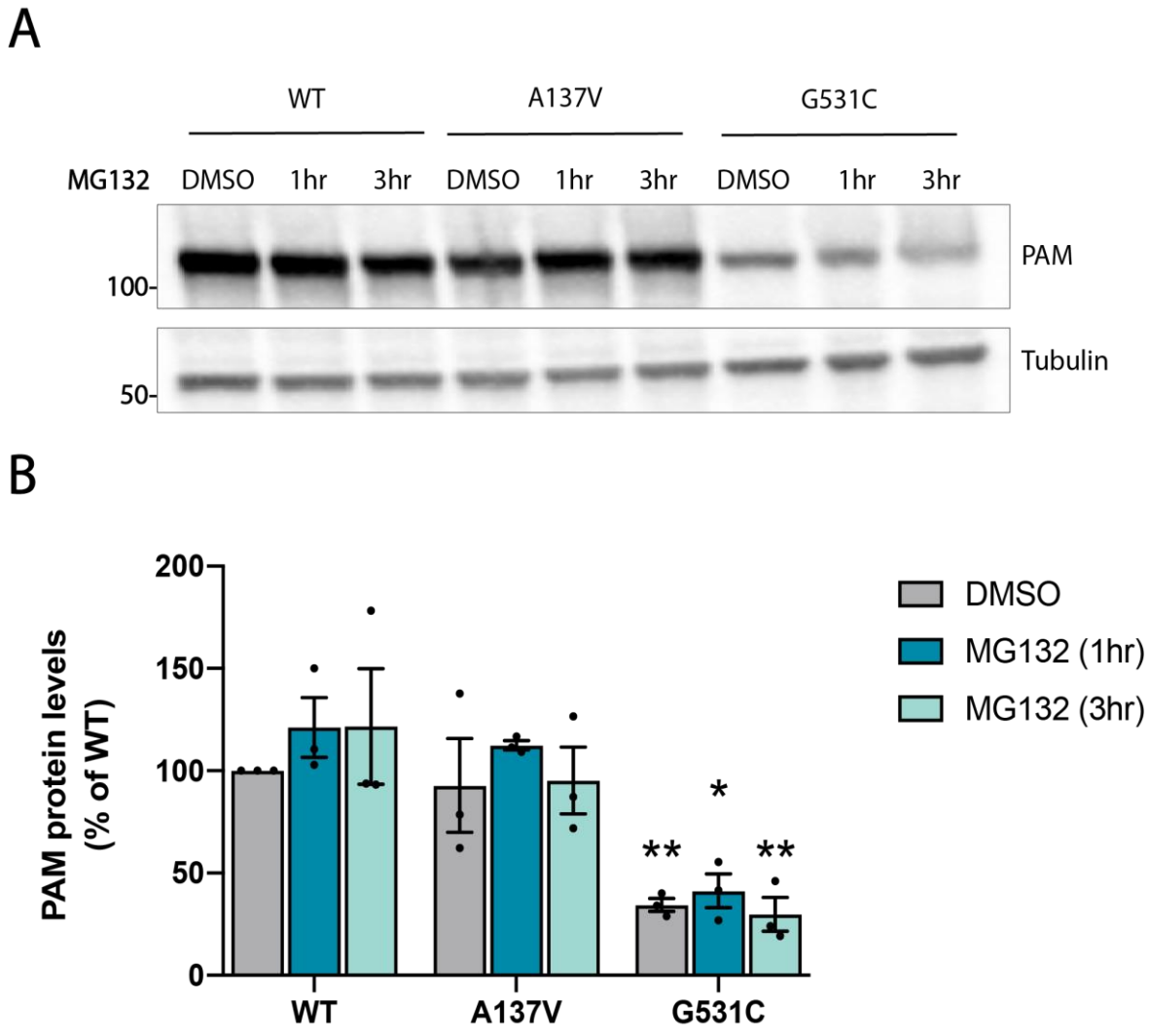

FIG. 5.3.4.1 Inhibition of the proteasomal pathways by MG132 does not rescue abundance of G531C-PAM. HEK293 cells transfected with membrane-integral PAM for 48hr were incubated with 10μM MG132 for 1-3hr or DMSO. Impact on protein abundance was assessed by western blotting analysis (**A**) Bands were quantified by lane densitometry, normalising values to WT-PAM (DMSO) within each experiment (**B**). Statistical analysis was performed by two-way ANOVA on transformed (log) data, comparing each variant/condition to WT (DMSO). Data are representative of 3 independent experiments and error bars are SEM. * denotes P<0.05, ** denotes P<0.01.

5.3.5 Abundance of membrane-integral PAM in PAM KO EndoC-βH1 cells

The expression and post-translational processing of PAM is perceived to be tissue-specific, and broadly restricted to neuroendocrine cells. Therefore, I postulated that there may be a different pattern observed in PAM KO EndoC-βH1 cells. HEK293 cells, therefore serve as a non-neuroendocrine control, and have been shown to lack endogenous expression of processing enzymes such as prohormone convertases important for PAM biosynthesis (Eipper, Green et al. 1992).

Using a similar approach to section 5.3.3, I transiently transfected PAM KO EndoC-βH1 cells with membrane-integral PAM and then examined the abundance of the membrane-integral isoform using western blotting analysis. I observed a reduction in the G531C-PAM as compared to WT-PAM (-53.1%, n=4, P=0.047), which is similar in effect size to what I observed in HEK293 cells. There were no significant differences in the abundance of any of the other variant proteins compared to WT-PAM (A137V, R144Q, R260W, R776C: P>0.97). However, there is much more variation (compared to similar analysis in HEK293 cells, FIG. 5.3.3.1). This has made it much more challenging to identify variants that affect PAM abundance.

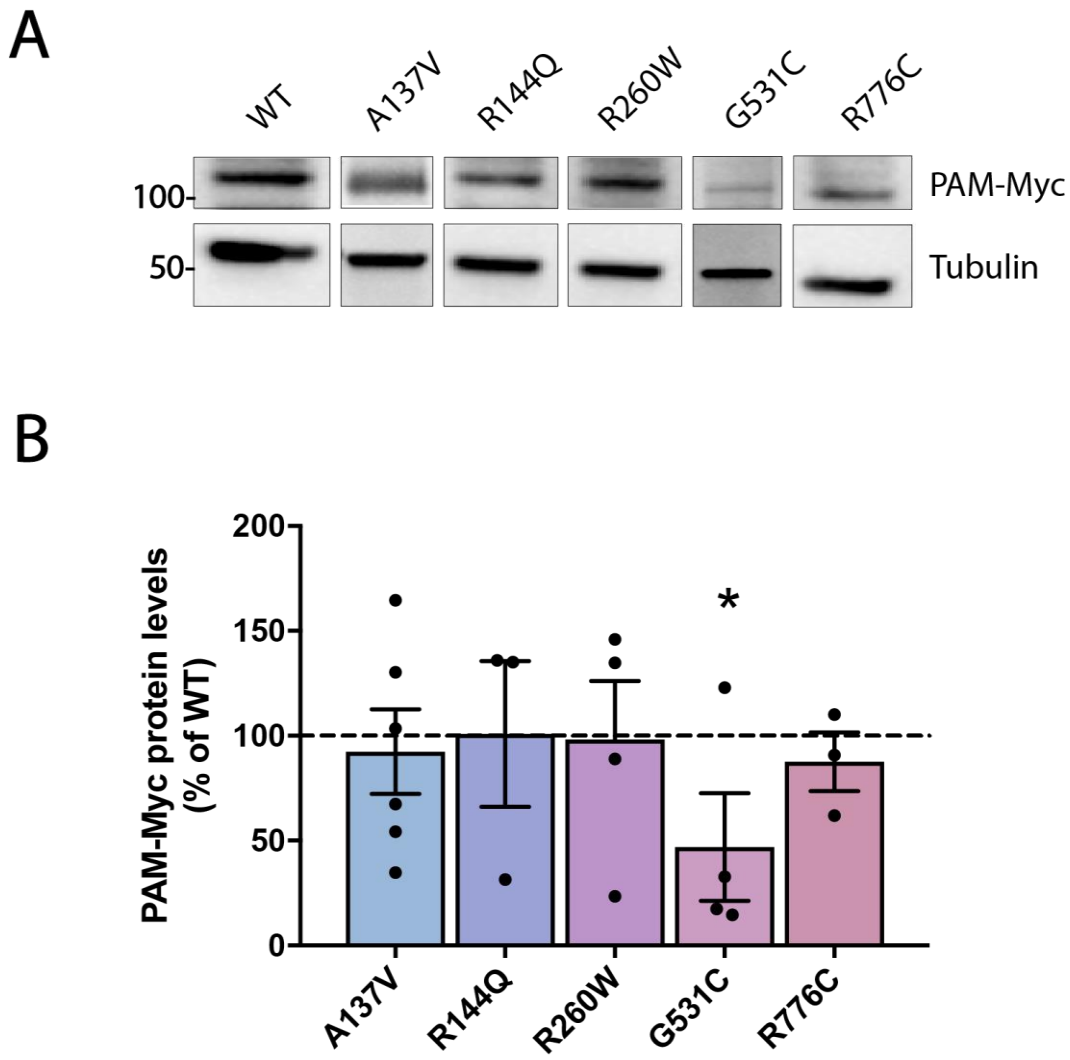

FIG. 5.3.5.1 Abundance of membrane-integral variant PAM in PAM KO EndoC-βH1 cells. Cells were transiently transfected with membrane-integral PAM for 72hr and then PAM abundance assessed by western blotting analysis (A). Bands were quantified by lane densitometry and normalised to WT-PAM (dotted line) (B). Data analysed by one-way ANOVA on transformed (log) values. Data are representative of 3-6 independent experiments and error bars are SEM. * denotes $P<0.05$.

5.3.6 Localisation of A137V-PAM and G531C-PAM in PAM KO EndoC-βH1 cells

I observed that the only the G531C allele resulted in reduced abundance of membrane-integral PAM in PAM KO EndoC-βH1 cells. However, variants exhibiting normal PAM abundance may still impact protein function through alternative mechanisms, such as mislocalisation or impaired trafficking, as was the case with the low frequency T2D-associated allele S539W (Thomsen, Raimondo et al. 2018). To test this, I performed immunostaining in fixed PAM KO EndoC-βH1 cells transfected with membrane integral PAM. I had issues with generating reliable immunostaining and as such, prioritised A137V-PAM and G531C-PAM (WT and S539W were included as controls).

As expected, I observed that WT-PAM displayed co-localisation with the *trans*-Golgi marker (TGN46). By contrast and as previously demonstrated, S539W did not co-localise with the *trans*-Golgi (FIG. 5.3.6.1) (Thomsen, Raimondo et al. 2018). For A137V-PAM, I observed a staining pattern similar to that of S539W-PAM, there was no co-localisation between PAM-FLAG (green) and TGN46 (red) staining, suggesting that A137V-PAM may be mislocalised. However, without further staining using markers targeting alternative subcellular organelles (e.g. ER), I cannot decipher exactly where A137V-PAM is targeted to. By contrast, the staining for G531C-PAM proved to be far more unreliable and non-descript (FIG. 5.3.6.1). There is no visible staining for TGN46 in the transfected cell (identified by anti-FLAG staining in green). This makes it difficult to ascertain the exact cellular localisation of G531C-PAM, without further optimisation of the staining and/or transfection conditions.

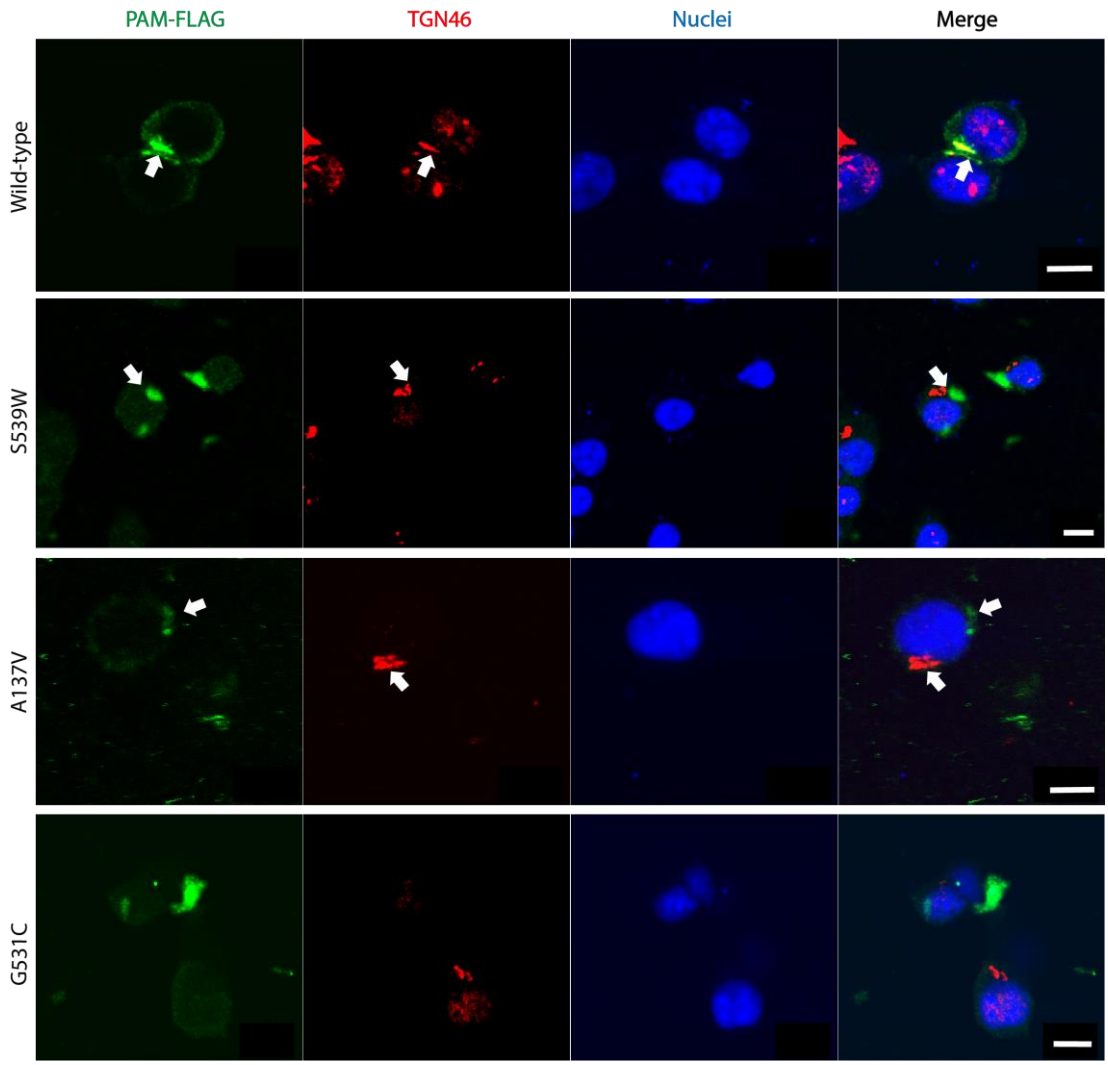

FIG. 5.3.6.1 Cellular localisation studies of A137V-PAM and G531C-PAM as assessed by immunofluorescence. Fixed PAM KO EndoC-βH1 cells transfected with membrane-integral PAM were double immunostained with PAM-FLAG (green) and TGN46 (red) and red dot2 (blue) applied as a nuclear marker. Images were visualised using a BioRad Radiance 2100 confocal microscope and arranged using Abobe Illustrator. Scale bar indicates 50µm.

I found that transfected cells (denoted by anti-FLAG staining, green) did not co-stain well for the *trans*-Golgi marker (TGN46, red) when the membrane-integral PAM was localised to the *trans*-Golgi (TGN). This is highlighted when comparing transfected and non-transfected cells in parallel in the same image (FIG. 5.3.6.2). These data suggest that the overexpression model may result in technical limitations, which in this experiment block the binding of anti-TGN46 antibodies.

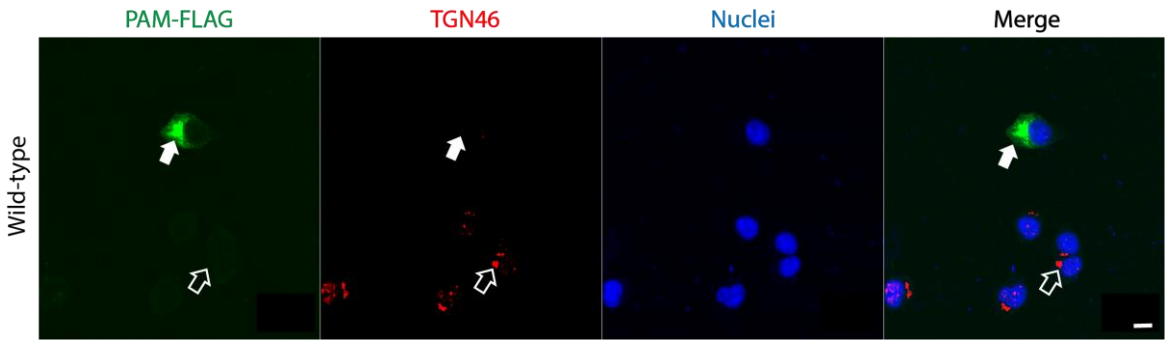

FIG. 5.3.6.2 Cells transfected with membrane-integral PAM do not co-stain well for TGN46. PAM KO EndoC-βH1 cells that are transfected with membrane-integral WT-PAM (white filled arrow) display poor co-staining for TGN46 (red) as compared with untransfected cells in the same slide (white hollow arrow). Cells were imaged using a BioRad Radiance 2100 confocal microscope; images were arranged using Adobe Illustrator. Scale bar indicates 50µm.

5.3.7 Impact on insulin secretion and content

I hypothesised that the reduced abundance observed for G531C and the mislocalisation observed for A137V-PAM may have a consequential impact on insulin secretion from β cells (EndoC-βH1). The low frequency T2D-associated allele S539W, which mediates both mislocalisation and protein instability, resulted in glucose-blunted insulin secretion when transiently overexpressed in PAM KO EndoC-βH1 (FIG. 4.3.9.1). I considered whether a similar effect may be observed for these variants.

To investigate this, I transiently transfected PAM KO EndoC-βH1 cells with membrane integral PAM (WT, S539W or G531C) and performed insulin secretion assays at 72hr post-transfection. Raw insulin secretion values are normalised to intracellular insulin content. Although preliminary, I observed trends towards a reduced basal and glucose-stimulated insulin secretion for G531C-PAM transfected cells as compared to WT-PAM transfected cells (2.8mM glucose: -68%; 16.7mM glucose: -59.5%; n=2). There were no obvious differences under maximally stimulated conditions (16.7mM glucose + 100μM tolbutamide + 10μM Forskolin), which indicates no impact on the number of granules (i.e. maximal secretory capacity).

I also performed insulin secretion assays in PAM KO EndoC-βH1 cells transfected with A137V-PAM, since although I didn't observe any defects to abundance of PAM protein, the immunostaining suggests that A137V may be mislocalised. There were no obvious defects in insulin secretion in A137V-transfected cells, although given the variable data points; it is difficult to ascertain the true effect. Unlike G531C, A137V has a basal insulin secretion similar to that of WT-PAM transfected cells (+106.2%, n=2), but does

appear to improve glucose stimulated insulin secretion at 16.7mM glucose as compared to WT-PAM transfected cells (+156.9%, n=2). There is similarly no defect at maximally stimulated conditions (16.7mM glucose + 100µM tolbutamide + 10µM Forskolin).

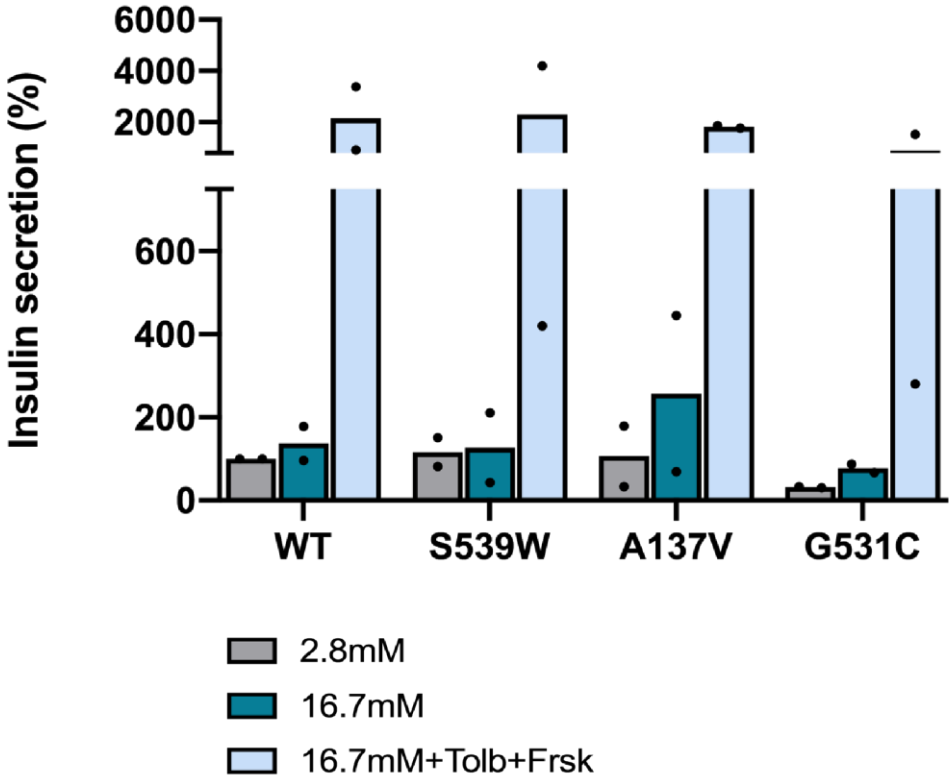

FIG. 5.3.7.1 Impact of A137V-PAM and G531C-PAM on insulin secretion in PAM KO EndoC-βH1 cells. Cells were transfected with membrane-integral PAM and insulin secretion assays performed at 72hr post-transfection. Quantified insulin secretion measures were adjusted to insulin content and then normalised to secretion for WT-transfected cells at 2.8mM glucose within each experiment. Data are representative of 2 independent experiments. WT: wild-type; Tolb: tolbutamide (100µM); Frsk: Forskolin (10µM).

5.4 DISCUSSION

In this chapter, I have established a strategy for the characterisation of a subset of rare coding variants in *PAM* (Flannick, Mercader et al. 2019). The cellular pathways included as part of this strategy are informed by my investigations on the endogenous protein in β cells (chapter 3) and for other *PAM* variants (chapter 4) and therefore include only those to be perceived to be critical to PAM function. This pilot study is the first step to a comprehensive characterisation of all 35 rare coding variants underlying the gene burden for T2D risk in *PAM,* which may help provide insights into additional functional mechanisms implicated in disease risk and may inform future efforts in drug development (Bonnefond, Clement et al. 2012, Flannick, Thorleifsson et al. 2014, Flannick, Mercader et al. 2018).

Of the variants identified in this gene-level association, two were previously reported through single variant analyses, D563G (MAF=4.98%) and S539W (MAF=0.65%), and were reported in association with T2D risk and insulinogenic index (Huyghe, Jackson et al. 2013, Steinthorsdottir, Thorleifsson et al. 2014, Fuchsberger, Flannick et al. 2016, Flannick, Mercader et al. 2018). Cellular characterisation of these variants identified causal mechanisms pertaining to impaired amidation activity (D563G) and altered protein stability and trafficking (S539W) in EndoC-βH1 cells (Thomsen, Raimondo et al. 2018). I prioritised 5 of the rare coding variants for study (A137V, R144Q, R260W, G531C and R776C) based on their enrichment in T2D cases, predicted pathogenicity on PAM function and position in catalytic domains. However, due to time constraints I did not investigate the impact of these variants on amidation activity for PAM. Instead, I

prioritised characterising the impact on PAM protein abundance, stability, cellular localisation and insulin secretion.

Overexpression of variant membrane-integral PAM constructs in HEK293 and PAM KO EndoC-βH1 cells identified only G531C as displaying reduced PAM protein abundance in both PAM KO EndoC-βH1 cells (FIG. 5.3.5.1). I considered that differences in the transfection efficiency may have an impact on the PAM abundance observed, so I proceeded to characterise the relative transfection efficiency between each variant compared to WT-PAM, using co-transfection with an independent GFP construct. Whilst I identified no significant differences in the abundance of GFP levels between variants, I observed that there was a minor trend towards reduced GFP abundance in G531C-PAM transfected cells as compared to WT-PAM transfected cells (-25.4%, P=0.626). This may partially contribute to the reduced PAM abundance that I observe in G531C-PAM transfected cells and would need to be further investigated.

My previous characterisation highlighted that the proteasome is involved in PAM biosynthesis and can imply a defect in protein stability. Whilst there was a nominal increase in the abundance of WT-PAM following incubation with MG132, there were no differences in the protein abundance observed for either A137V-PAM or G531C-PAM (FIG. 5.3.4.1). However, there may be an alternative pattern observed in PAM KO EndoC-βH1 cells which was not performed. As membrane-integral PAM contains a C-terminal transmembrane domain, it may be preferentially degraded through either the lysosomal pathway or by autophagy, particularly if it is trafficked to beyond the ER (Schulze, Kolter et al. 2009, Tooze, Abada et al. 2014, Galluzzi, Baehrecke et al. 2017).

Treatment of cells transfected with membrane-integral PAM with an inhibitor of autophagy, such as chloroquine, may investigate whether non-proteasomal degradation of G531C-PAM explains the reduced PAM abundance observed in PAM KO EndoC-βH1 cells. Chloroquine inhibits autophagy through preventing the fusion of autophagosome with lysosomes, thereby blocking the terminal degradation reactions (Mauthe, Orhon et al. 2018).

I next characterised the impact of these variants on the localisation of membrane-integral PAM in PAM KO EndoC-βH1 cells, where I prioritised A137V-PAM and G531C-PAM for investigation. As previously shown, membrane-integral WT-PAM localised to the *trans*-Golgi, as shown by co-localisation with TGN46. By contrast, membrane-integral S539W-PAM was mislocalised as it did not co-localise with TGN46 (Thomsen, Raimondo et al. 2018). For A137V-PAM, I observed that PAM-FLAG staining did not co-localise with TGN46, suggesting that it A137V-PAM may also be mislocalised (FIG. 5.3.6.1). I had run into a number of technical difficulties that made performing immunofluorescence a challenge, and meant that I cannot deduce the localisation of G531C-PAM with the quality of images I generated. The first of these issues, which are an inherent problem for EndoC-βH1 cells, is the low transfection efficiency. This means that I am limited in the number of cells that I can image in a given slide, and can greatly reduce the likelihood of successful staining. Secondly, I observed that TGN46 staining was poor or absent in transfected cells as compared to untransfected cells in the same slide (FIG. 5.3.6.2). This is likely an artefact mediated through the overexpression system. Markers of other cellular compartments, such as the ER and

insulin-containing granules, should also be included to provide resolution on where mislocalised PAM is targeted if not to the *trans*-Golgi network.

I lastly investigated the impact of A137V-PAM and G531C-PAM on insulin secretion and content in PAM KO EndoC-βH1 cells. In chapter 4, I observed a defect in insulin secretion for cells transfected with S539W-PAM, where there was a glucose-blunted response (FIG. 4.3.9.1). I hypothesised that I may be able to observe similar defects following transient transfection of A137V-PAM and G531C-PAM. The preliminary data show that G531C-PAM results in reduced basal insulin secretion (2.8mM glucose) and stimulated insulin secretion (16.7mM glucose), despite the stimulation index remaining within normal range (Fold change [16.7mM-2.8mM]= 2.4). There also appeared to be no stark effect under maximal stimulation suggesting that there is no impact on the total number of insulin-containing granules. A137V-PAM, by contrast, had a basal insulin secretion similar to that of WT-PAM transfected and S539W-PAM transfected cells. If confirmed, these data may reflect a mechanism for G531C-PAM involving a defect in a glucose-dependent regulatory element of exocytosis, without affecting the total secretory capacity of these cells. The normal stimulation index suggests that the G531C-PAM transfected cells are capable of sensing an increase in glucose, but cannot release the same amount of insulin as WT-PAM transfected cells. These data highlight the potentially diverse mechanisms in which altered PAM function can affect insulin secretion and is an important avenue that should be interrogated further.

Future experiments should prioritise the characterisation of these variants in amidating activity; this was not included in my pilot study in the interests of time. However, carboxyamidation is an established function for PAM, which is impaired for the independent T2D-associated alleles (Eipper and Mains 1988, Thomsen, Raimondo et al. 2018). In addition, these variants are all located within the catalytic domains so it is likely that some may affect amidating activity, even if protein stability and localisation remain unaffected. This was observed with the common T2D-associated allele, D563G, which displayed a ~25% reduction in amidating activity in the *in vitro* assay, despite normal cellular localisation (Thomsen, Raimondo et al. 2018). This application of this characterisation strategy to all 35 rare coding variants identified as part of the gene-level association will help to fully elucidate the complex relationship between altered *PAM* function and T2D risk. It may also uncover further biological roles for PAM in glucose homeostasis and metabolism.

CHAPTER 6

Discussion and Future Directions

6.1 Using coding variation in *PAM* to elucidate its roles in β cell function and diabetes pathogenesis

6.2 Distinctions in the causal mechanisms between T2D risk and PNDM

6.3 Challenges in defining novel genetic aetiology for disease

6.4 Future directions

6.1 USING CODING VARIATION IN *PAM* TO ELUCIDATE ITS ROLES IN β CELL FUNCTION AND DIABETES PATHOGENESIS

The general aim of my thesis has been to identify the functional roles for PAM in pancreatic β cells and the cellular mechanisms by which dysfunction leads to either T2D risk or early onset diabetes pathogenesis.

PAM was identified as a T2D susceptibility gene following the identification of two independent coding alleles, D563G (MAF=4.98%) and S539W (MAF=0.65%), which were associated with increased T2D risk and reduced insulinogenic index (Huyghe, Jackson et al. 2013, Steinthorsdottir, Thorleifsson et al. 2014, Fuchsberger, Flannick et al. 2016). Functional characterisation of these risk alleles uncovered causal mechanisms in the β cell pertaining to impaired amidating activity and reduced PAM protein stability and trafficking respectively (Thomsen, Raimondo et al. 2018).

In chapter 3, I characterised the role of (endogenous) PAM in the β cell and the mechanisms by which altered PAM function (through impaired carboxyamidation) may precipitate T2D risk. Characterisation of PAM highlighted roles in glucose-stimulated insulin secretion (GSIS) and showed that this is impaired in islets isolated for T2D risk allele carriers (Thomsen, Raimondo et al. 2018). Using both *PAM* silencing and catalytic inhibition (4P3BA), I identified that reduced insulin content observed following PAM depletion was mediated through reduced proinsulin biosynthesis, which was observed both at mRNA and protein level (FIG. 3.3.5.1, FIG. 3.3.5.2, FIG. 3.3.5.3) (Thomsen,

Raimondo et al. 2018). This correlates with fewer insulin particles per granule as seen by electron microscopy (Thomsen, Raimondo et al. 2018).

Impaired PAM carboxyamidation also affected the amidation of chromogranin A (FIG. 3.3.7.3, FIG. 3.3.7.4), an auxiliary protein present in the secretory granule and implicated in granule biogenesis (Kim, Tao-Cheng et al. 2001). Reduced carboxyamidation led to downregulated chromogranin A (CgA) expression at both mRNA and protein level (FIG. 3.3.7.4, FIG 3.3.7.5). In addition, I observed an altered localisation pattern for CgA in PAM KO EndoC-βH1 cells (FIG. 3.3.8.1), which lack catalytically active PAM but retain a short C-terminal fragment of unknown function (FIG. 3.3.1.2, FIG. 3.3.1.3). Whilst this work is preliminary, these data may suggest an important role for PAM (carboxyamidation) in governing the correct stability and trafficking of CgA and its subsequent impact on granule biogenesis, which is currently undescribed. The maturation of chromogranin A and its post-translational processing is largely undefined, so determining the relative impact of impaired PAM function in this is difficult (Huttner and Benedum 1987, Loh, Koshimizu et al. 2012). The secretory granule as a scaffold of proteins, cofactors and cargo is a complex microenvironment to characterise, and there are limited techniques with which sufficient resolution can be generated. However, the role of PAM in insulin secretion through interactions made within the granule is relevant and important to investigate. These data may imply a currently undescribed feedback loop between PAM carboxyamidation, *CHGA* expression and trafficking and *INS* expression and content. The relative contribution of non-catalytic PAM functions in the β cell has not been investigated. It would be

important to decipher these roles in order to understand the full scope of PAM function in the β cell.

The identification of a gene-based signal for T2D risk provides further evidence for *PAM* as a relevant locus for study in T2D susceptibility (Flannick, Mercader et al. 2018, Mahajan, Taliun et al. 2018, Mahajan, Wessel et al. 2018). Identification of aggregate burdens of rare, coding variants in specific genes are proposed to help uncover disease susceptibility, which has been a challenge in complex disease (Gusev, Lee et al. 2014, MacArthur, Manolio et al. 2014, Guo, Plummer et al. 2018). Gene-level signals for T2D risk have only been identified for a handful of loci: *PAM*, *SLC30A8* and *MC4R* (Flannick, Mercader et al. 2018). For *PAM,* this is composed of 35 rare coding variants, which individually are too rare to produce independent association signals (Flannick, Mercader et al. 2018). This approach has successfully identified disease genes for both Mendelian (e.g. *ACTN1* in congenital macrothrombocytopenia) and complex diseases (e.g. *LDLR* in myocardial infarction) (Kunishima, Okuno et al. 2013, Cirulli, Lasseigne et al. 2015, Do, Stitziel et al. 2015). Functional characterisation of rare variants identified by this approach can also help to elucidate the relationship between altered gene function and disease risk (i.e. the gene-dose effect or therapeutic window), but to do so would require a screening strategy that is high throughput, efficient and comprehensive.

In chapter 5, I set up a functional screening strategy for the characterisation of multiple rare variants in *PAM* identified in the gene-level signal for T2D risk (Flannick, Mercader et al. 2018). As a pilot study, I prioritised a subset of these variants, selecting

those that were located in the catalytic domains and enriched in T2D cases as compared with controls. This was further stratified to only include variants predicted to be deleterious by all *in silico* prediction tools (SIFT, PolyPhen2 and Condel). The functional components included in this strategy were informed by my characterisation of endogenous PAM function (in chapter 3) and other *PAM* variants (in chapter 4), which allowed for prioritisation of relevant cellular mechanisms for PAM in the β cell. For example, mechanisms pertaining to PAM protein abundance and stability, cellular localisation, amidating activity and insulin secretion in β cells proved to be important in the causal mechanisms underlying T2D risk alleles, D563G and S539W, and in R36S (Thomsen, Raimondo et al. 2018). Due to time constraints, I was unable to investigate the impact on amidating activity in my pilot study (chapter 5), but this should be performed in the future. For the D563G and S539W alleles, an aspect of their impaired function could be quantified using the serum amidation assay (FIG. 4.3.2.1). Serum samples collected from carriers of all of the 35 rare variants underlying the gene burden for T2D risk, consistent with a recruit-by-genotype (RBG) approach, may allow for the high throughput interrogation of the physiological impact of these variants on amidating function. The main challenge would be in finding carriers of these variants, given the rare frequency of the risk allele. For completeness, this characterisation strategy should also be replicated for variants that are enriched in control individuals and therefore may infer protection to disease risk. This will elucidate the full gene-dose relationship and may highlight the prevalence of both loss of function (LoF) and gain of function (GoF) variants in disease risk. LoF alleles in *SLC30A8* examined in an allelic series were found to be protective against T2D risk (Flannick, Thorleifsson et al. 2014).

6.2 DISTINCTIONS IN THE CAUSAL MECHANISMS BETWEEN T2D RISK AND MONOGENIC DIABETES ALLELES

Genes identified as causal in monogenic diabetes are often identified to harbour T2D-associated variants (e.g. in *HNF1A*) (Stanescu, Hughes et al. 2012, Estrada, Aukrust et al. 2014). It is therefore not surprising to capture similar signals in *PAM*. However, characterisation of the low frequency T2D risk allele (S539W) demonstrated that even with a cellular phenotype resembling a *PAM* LoF allele, this was not sufficient to be causal for monogenic disease (Thomsen, Raimondo et al. 2018). This implicates that mechanisms disrupted in monogenic disease for *PAM* variants, are likely to encompass additional or alternative mechanisms to those identified underlying T2D risk.

My characterisation of the R36S-PAM variant (Chapter 4) has allowed for the direct comparison in cellular mechanisms between a rare allele in monogenic disease (R36S) and a low-frequency LoF T2D risk allele (S539W). Therefore, I have been able to elucidate differences in the cellular mechanisms between the alleles. Since the predominant transcript in islets encodes a membrane-integral PAM isoform (FIG. 3.3.1.1), this comparative characterisation has been performed on this isoform only.

I observed that, similar to the T2D risk allele S539W, membrane-integral R36S-PAM is also mislocalised and likely ER-retained. With the understanding that loss of PAM function itself does not mediate monogenic diabetes (as exhibited by S539W), I anticipated that this mislocalisation could not be the causal mechanism driving early onset diabetes mellitus (DM) pathogenesis and that there must be additional processes that deviate between S539W- and R36S-PAM alleles. As my characterisation

progressed, I began to see further deviations in the effects of S539W-PAM and R36S-PAM proteins on β cell physiology, which explain the differences in associated clinical severity. For R36S-PAM, sequestration in the ER led to hyperactivation of the UPR and activation of a pro-apoptotic signalling cascade, which resulted in reduced β cell viability and proliferation in culture. In contrast, S539W-PAM did not result in activation of these UPR or pro-apoptotic markers and had a cellular proliferation similar to that of WT-PAM transfected cells. This suggests that there may be distinctions in the level of misfolding that occurs in the ER between these two different *PAM* alleles. However, since the human PAM structure is unresolved, it is not possible to elucidate the true impact of these alleles on misfolding without crystallography and/or *in silico* modelling.

Chaperone proteins (e.g. BiP) are ready available in the ER lumen and are proposed to bind and sequester misfolded proteins to minimise aggregation capable of elevating stress (Merksamer, Trusina et al. 2008, Jessop, Watkins et al. 2009). Minimal deviations in the native conformation of a protein during folding is sufficient to instigate the binding of these chaperone proteins (Fra, Fagioli et al. 1993, Hellman, Vanhove et al. 1999). Generally, proteins transiting through the ER will undergo transient, dynamic switches between their native (folded) and unfolded states (FIG. 6.2.1) (Ellgaard and Helenius 2003). The more unstable a protein is, the longer it spends in its unfolded state and the more likely it is to be ER-retained (Kowalski, Parekh et al. 1998). This may be an additional difference mediated by R36S and S539W substitutions, which promote greater instability for R36S and thereby promote UPR activation. The binding affinity of chaperones for a misfolded protein may be key in

determining the level of UPR activation and cellular dysfunction caused (Powers, Morimoto et al. 2009, Onn and Ron 2010). For example, the R36S substitution may have a negative impact on the binding affinity to chaperones, which without optimal interaction are unable to efficiently sequester misfolded R36S-PAM (FIG. 6.2.2). This mediates greater exposure of misfolded R36S-PAM aggregates to UPR sensors, thereby resulting in hyperactivation of the UPR and associated pathways. This is likely exacerbated by the overexpression system, which has been used for all of the characterisation presented. Not all misfolded proteins have been observed to result in UPR activation, as is the case for S539W-PAM (Hidvegi, Schmidt et al. 2005). In fact, some have proposed that it is the availability of BiP itself that can determine whether UPR activation is mediated following protein misfolding (Bertolotti, Zhang et al. 2000, Credle, Finer-Moore et al. 2005, Oikawa, Kimata et al. 2009). BiP has also been proposed to act as a regulator of IRE1-mediated activation, where it monitors the duration of UPR activation and can thereby assess whether ER homeostasis can be restored in a timely manner (Pincus, Chevalier et al. 2010). It is therefore possible that BiP functions to mediate different cellular outcomes for S539W- and R36S-PAM alleles, thereby promoting different causal mechanisms.

The crystal structure for human PAM is currently unresolved, but may highlight important motifs important for global protein folding. This may also identify structural features within the proregion or PAL domain that are disrupted by S539W and R36S substitutions respectively and thereby result in altered conformations in those domains. The proregion of PAM, near which R36S is located, has been proposed to influence the solubility and secretion of PAM protein (Mains, Milgram et al. 1995).

Mutants in the N-terminal proregion of prothryotropin-releasing hormone, mediated defects in the exit of folded protein from the ER to the Golgi (Romero, Çakir et al. 2008).

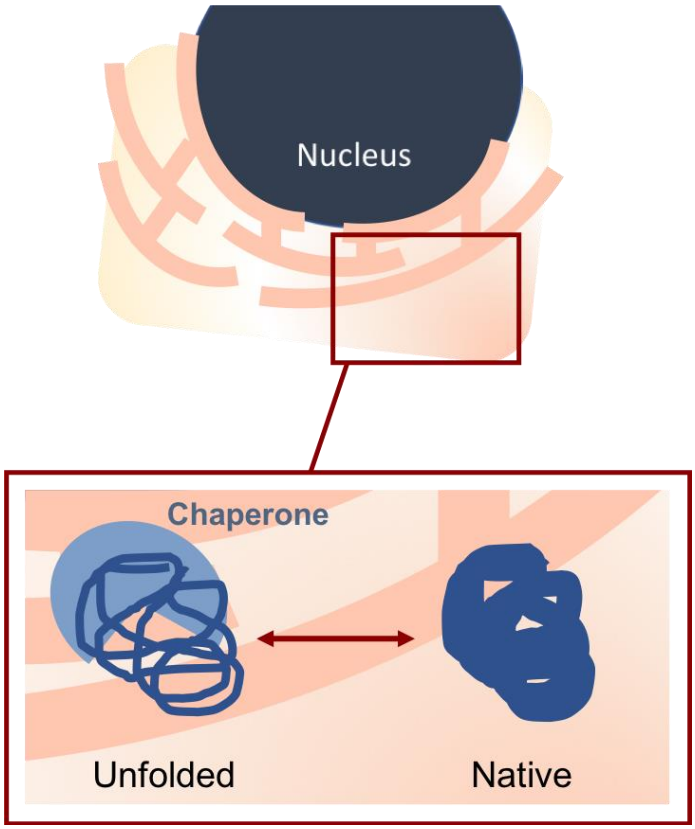

FIG. 6.2.1 Highly unstable proteins are more likely to be convert to their unfolded conformations during transit through the ER. This results in elevated probability of chaperone binding and UPR activation and may provide a distinct mechanism between S539W- and R36S-PAM alleles.

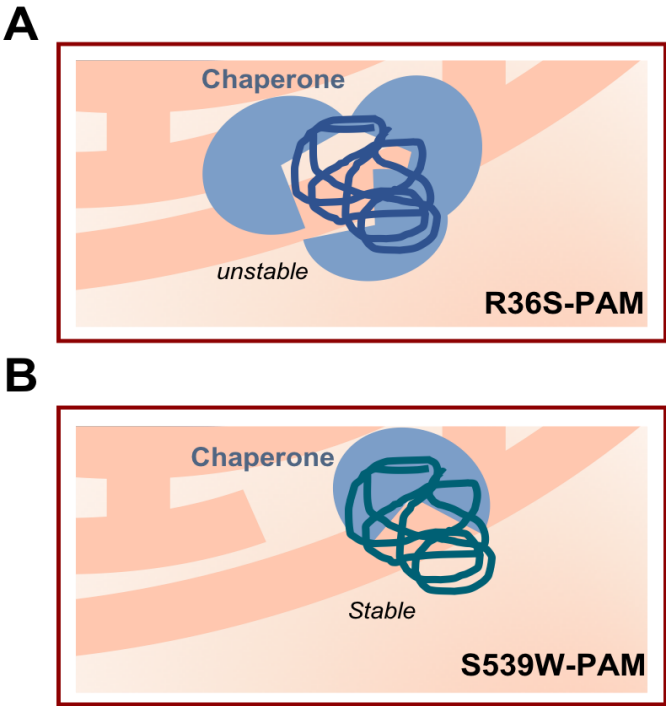

FIG. 6.2.2 The binding efficiency of misfolded proteins with ER-resident chaperones may dictate the outcome on UPR activation and cellular stress. (**A**) Binding of chaperones to R36S-PAM may not stabilise the structure sufficiently, thereby leading to UPR activation. (**B**) By contrast, the structure of S539W-PAM may be efficiently stabilised by chaperone binding, thereby limiting the need for UPR activation.

Exploration of the impact on insulin secretion revealed further deviations between the S539W- and R36S-PAM alleles. I observed that S539W-PAM transfected cells displayed a glucose-blunted response at 16.7mM glucose (FIG. 4.3.9.1) which was not observed for either WT-PAM or R36S-PAM transfected cells. In fact, the effects observed for S539W-PAM are consistent with the GFP-transfected cells which lack catalytically-active PAM, providing further evidence that S539W-PAM results in a *PAM* knockout-like phenotype in the β cell. For R36S-PAM, there appears to be no acute effect on glucose-stimulated insulin secretion, as shown by a similar fold change to WT-PAM transfected cells between 2.8mM and 16.7mM glucose conditions (FIG. 4.3.10.1). However, the impact of prolonged hyperactivation of the UPR on insulin secretion may not be apparent at 3 days after transfection (72hr). Prolonged UPR activation has also been shown to impede the correct folding of proinsulin, which over time may result in reduced insulin content and be an additional mechanism altered in R36S-mediated dysfunction (Lipson, Fonseca et al. 2006, Feng, Wei et al. 2009).

6.3 CHALLENGES IN DEFINING NOVEL GENETIC AETIOLOGY FOR DISEASE

Identification and characterisation of novel variants in disease is the first step to establishing novel genetic aetiologies for disease. However, to consider a variant (e.g. R36S-PAM) to be pathogenic, pre-defined criteria set out by the American College of Medical Genetics (ACMG) need to be met. This algorithm considers various lines of evidence (FIG. 6.3.1) as indicators of whether a given mutation is pathogenic or benign (Richards, Aziz et al. 2015).

	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	MAF is too high for disorder *BA1/BS1* **OR** observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product *BP4* Missense in gene where only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		Co-segregation with disease in multiple affected family members *PP1*	Increased segregation data →		
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			

FIG. 6.3.1 ACMG criteria for determining whether a variant is considered pathogenic for disease (adapted from Richards et al, 2015). Variants can be designated as either pathogenic, likely pathogenic, benign or likely benign.

The use of next generation sequencing (NGS) in genetic laboratories has been critical to increasing the accuracy, efficiency and the number of novel variants underlying disease (Vrijenhoek, Kraaijeveld et al. 2015). However, this shift from Sanger sequencing to NGS has brought about challenges, such as increased incidental findings (Cooper and Shendure 2011, Frebourg 2014). The identification of many variants of unknown significance, makes finding causal pathogenic variants a challenge and tricky to apply to a clinical setting (Katsanis and Katsanis 2013, Matthijs, Souche et al. 2016).

Identification of additional families can provide support to a variant being causal for disease. For R36S-PAM (currently an N=1), identification of another family would strengthen the functional evidence provided by *in vitro* characterisation. The application of a genetic risk score for T1D in combination with clinical indications would also help differentiate between diagnosis of monogenic diabetes and T1D in this proband (Patel, Oram et al. 2016). This information could provide further strength to the argument that a variant is causal for (monogenic) disease.

6.4 FUTURE DIRECTIONS

Through the investigation of multiple aspects, my work has uncovered a more complex role for PAM in β cells. I have shown important roles for PAM in maintaining insulin content and β cell function through the modulation of auxiliary proteins, such as chromogranin A. I have also shown how varying impairment of PAM protein function by coding variants may lead to different functional outcomes, which in turn could mediate either T2D risk or early onset DM pathogenesis.

Following the identification and characterisation of R36S-PAM, there should be a continued search for additional cases of monogenic diabetes. It is likely to be only certain alleles in *PAM* that give rise to such a severe clinical phenotype, and therefore cases may be variant-specific. Another key direction would be characterisation of p.R36S (p.R40S in mouse) in a knock-in mouse model, which is currently being generated in collaboration with MRC Harwell. Whilst I have provided evidence to suggest that causal mechanisms underlying early onset DM in this proband can be attributed to R36S-PAM and its activation of ER stress and β cell death; it is important to confirm whether global expression of this allele results in insulin-deficient glucose intolerance phenotypes at birth consistent with the early onset DM pathogenesis present in the proband. This model would also allow investigation into extra-pancreatic phenotypes caused by this variant. It is important to consider the limitations with using a mouse model, particularly the functional differences between rodent and human physiology. The proband has not presented with any extra pancreatic phenotypes but this is still important to consider. An additional follow up study is the implementation of the characterisation strategy I described in chapter 5. This would allow characterisation of all 35 rare coding variants underlying the gene-level signal for T2D risk, and provide both confirmation of my preliminary work and elucidation of novel mechanisms underlying *PAM* variants in T2D risk.

Bibliography

ADA (2010). "Diagnosis and classification of diabetes mellitus." Diabetes Care **33 Suppl 1**: S62-69.

Adachi, Y., et al. (2008). "ATF6 is a transcription factor specializing in the regulation of quality control proteins in the endoplasmic reticulum." Cell Struct Funct **33**(1): 75-89.

Aguilar-Bryan, L., et al. (1998). "Toward understanding the assembly and structure of KATP channels." Physiol Rev **78**(1): 227-245.

Ahren, B., et al. (1988). "Pancreastatin inhibits insulin secretion and stimulates glucagon secretion in mice." Diabetes **37**(3): 281-285.

Alam, M. R., et al. (1996). "Novel proteins that interact with the COOH-terminal cytosolic routing determinants of an integral membrane peptide-processing enzyme." J Biol Chem **271**(45): 28636-28640.

Alam, M. R., et al. (2001). "Signaling mediated by the cytosolic domain of peptidylglycine alpha-amidating monooxygenase." Mol Biol Cell **12**(3): 629-644.

Allen, H. L., et al. (2011). "GATA6 haploinsufficiency causes pancreatic agenesis in humans." Nat Genet **44**(1): 20-22.

Altshuler, D., et al. (2000). "The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes." Nat Genet **26**(1): 76-80.

Apgar, V. (1953). "A proposal for a new method of evaluation of the newborn infant." Curr Res Anesth Analg **32**(4): 260-267.

Arshad, M., et al. (2013). "RNF13, a RING finger protein, mediates endoplasmic reticulum stress-induced apoptosis through the inositol-requiring enzyme (IRE1alpha)/c-Jun NH2-terminal kinase pathway." J Biol Chem **288**(12): 8726-8736.

Arvan, P. and D. Castle (1998). "Sorting and storage during secretory granule biogenesis: looking backward and looking forward." The Biochemical journal **332 (Pt 3)**(Pt 3): 593-610.

Ashcroft, F. M. (1996). "Mechanisms of the glycaemic effects of sulfonylureas." Horm Metab Res **28**(9): 456-463.

Ashcroft, F. M. and F. M. Gribble (1998). "Correlating structure and function in ATP-sensitive K+ channels." Trends Neurosci **21**(7): 288-294.

Ashcroft, F. M. and P. Rorsman (1989). "Electrophysiology of the pancreatic beta-cell." Prog Biophys Mol Biol **54**(2): 87-143.

Babenko, A. P., et al. (2006). "Activating mutations in the ABCC8 gene in neonatal diabetes mellitus." N Engl J Med **355**(5): 456-466.

Barg, S., et al. (2001). "Priming of insulin granules for exocytosis by granular Cl(-) uptake and acidification." J Cell Sci **114**(Pt 11): 2145-2154.

Barroso, I., et al. (2003). "Candidate gene association study in type 2 diabetes indicates a role for genes involved in beta-cell function as well as insulin action." PLoS biology **1**(1): E20-E20.

Belkadi, A., et al. (2016). "Whole-exome sequencing to analyze population structure, parental inbreeding, and familial linkage." Proceedings of the National Academy of Sciences **113**(24): 6713.

Bellanne-Chantelot, C., et al. (2004). "Clinical spectrum associated with hepatocyte nuclear factor-1beta mutations." Ann Intern Med **140**(7): 510-517.

Bertolotti, A., et al. (2000). "Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response." Nat Cell Biol **2**(6): 326-332.

Bhattacharya, N., et al. (2012). "The structure of the Sec13/31 COPII cage bound to Sec23." J Mol Biol **420**(4-5): 324-334.

Bingham, C., et al. (2001). "Mutations in the hepatocyte nuclear factor-1beta gene are associated with familial hypoplastic glomerulocystic kidney disease." Am J Hum Genet **68**(1): 219-224.

Blundell, T. L. and R. E. Humbel (1980). "Hormone families: pancreatic hormones and homologous growth factors." Nature **287**(5785): 781-787.

Boesgaard, T. W., et al. (2010). "Further evidence that mutations in INS can be a rare cause of Maturity-Onset Diabetes of the Young (MODY)." BMC Med Genet **11**: 42.

Bonifacio, E., et al. (2010). "Harmonization of glutamic acid decarboxylase and islet antigen-2 autoantibody assays for national institute of diabetes and digestive and kidney diseases consortia." J Clin Endocrinol Metab **95**(7): 3360-3367.

Bonnefond, A., et al. (2012). "Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes." Nat Genet **44**(3): 297-301.

Bonnefond, A. and P. Froguel (2015). "Rare and Common Genetic Events in Type 2 Diabetes: What Should Biologists Know?" Cell metabolism **21**(3): 357-368.

Bonnefond, A., et al. (2012). "Whole-exome sequencing and high throughput genotyping identified KCNJ11 as the thirteenth MODY gene." PLoS One **7**(6): e37423.

Bonnefond, A., et al. (2013). "Transcription factor gene MNX1 is a novel cause of permanent neonatal diabetes in a consanguineous family." Diabetes Metab **39**(3): 276-280.

Bonnemaison, M., et al. (2014). "AP-1A controls secretory granule biogenesis and trafficking of membrane secretory granule proteins." Traffic (Copenhagen, Denmark) **15**(10): 1099-1121.

Botstein, D. and N. Risch (2003). "Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease." Nature Genetics **33**: 228.

Bowman, P., et al. (2012). "Heterozygous ABCC8 mutations are a cause of MODY." Diabetologia **55**(1): 123-127.

Bradbury, A. F., et al. (1982). "Mechanism of C-terminal amide formation by pituitary enzymes." Nature **298**(5875): 686-688.

Brezniceanu, M.-L., et al. (2010). "Reactive oxygen species promote caspase-12 expression and tubular apoptosis in diabetic nephropathy." Journal of the American Society of Nephrology : JASN **21**(6): 943-954.

Brissova, M., et al. (2005). "Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy." J Histochem Cytochem **53**(9): 1087-1097.

Brown, M., et al. (2016). "An initial phase of JNK activation inhibits cell death early in the endoplasmic reticulum stress response." J Cell Sci **129**(12): 2317-2328.

Brunton, S. (2014). "GLP-1 receptor agonists vs. DPP-4 inhibitors for type 2 diabetes: is one approach more successful or preferable than the other?" Int J Clin Pract **68**(5): 557-567.

Butler, A. E., et al. (2003). "Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes." Diabetes **52**(1): 102-110.

Cabrera, O., et al. (2006). "The unique cytoarchitecture of human pancreatic islets has implications for islet cell function." Proc Natl Acad Sci U S A **103**(7): 2334-2339.

Calfon, M., et al. (2002). "IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA." Nature **415**(6867): 92-96.

Carpenter, M. C. and D. E. Wilcox (2014). "Thermodynamics of formation of the insulin hexamer: metal-stabilized proton-coupled assembly of quaternary structure." Biochemistry **53**(8): 1296-1301.

Chakera, A. J., et al. (2014). "The 0.1% of the population with glucokinase monogenic diabetes can be recognized by clinical characteristics in pregnancy: the Atlantic Diabetes in Pregnancy cohort." Diabetes Care **37**(5): 1230-1236.

Chang, I., et al. (2003). "Nuclear Factor κB Protects Pancreatic β-Cells From Tumor Necrosis Factor-α-Mediated Apoptosis." Diabetes **52**(5): 1169.

Chang, S. G., et al. (2003). "Role of disulfide bonds in the structure and activity of human insulin." Mol Cells **16**(3): 323-330.

Chatila, T. A., et al. (2000). "JM2, encoding a fork head-related protein, is mutated in X-linked autoimmunity-allergic disregulation syndrome." J Clin Invest **106**(12): R75-81.

Chatterjee, S., et al. (2017). "Type 2 diabetes." Lancet **389**(10085): 2239-2251.

Chen, C., et al. (2017). "Human beta cell mass and function in diabetes: Recent advances in knowledge and technologies to understand disease pathogenesis." Molecular metabolism **6**(9): 943-957.

Chufan, E. E., et al. (2009). "Amidation of bioactive peptides: the structure of the lyase domain of the amidating enzyme." Structure **17**(7): 965-973.

Ciccotosto, G. D., et al. (1999). "Induction of integral membrane PAM expression in AtT-20 cells alters the storage and trafficking of POMC and PC1." The Journal of cell biology **144**(3): 459-471.

Cirulli, E. T., et al. (2015). "Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways." Science **347**(6229): 1436-1441.

Cnop, M., et al. (2012). "Endoplasmic reticulum stress, obesity and diabetes." Trends in Molecular Medicine **18**(1): 59-68.

Cohen, J. C., et al. (2006). "Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels." Proc Natl Acad Sci U S A **103**(6): 1810-1815.

Colclough, K., et al. (2013). "Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha and 4 alpha in maturity-onset diabetes of the young and hyperinsulinemic hypoglycemia." Hum Mutat **34**(5): 669-685.

Collier, C. A., et al. (2006). "Metformin counters the insulin-induced suppression of fatty acid oxidation and stimulation of triacylglycerol storage in rodent skeletal muscle." Am J Physiol Endocrinol Metab **291**(1): E182-189.

Colombo, C., et al. (2008). "Seven mutations in the human insulin gene linked to permanent neonatal/infancy-onset diabetes mellitus." J Clin Invest **118**(6): 2148-2156.

Cooper, G. M. and J. Shendure (2011). "Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data." Nat Rev Genet **12**(9): 628-640.

Credle, J. J., et al. (2005). "On the mechanism of sensing unfolded protein in the endoplasmic reticulum." Proc Natl Acad Sci U S A **102**(52): 18773-18784.

Cunningham, C. N., et al. (2017). "Chaperone-Driven Degradation of a Misfolded Proinsulin Mutant in Parallel With Restoration of Wild-Type Insulin Secretion." Diabetes **66**(3): 741-753.

Czyzyk, T. A., et al. (2005). "Deletion of peptide amidation enzymatic activity leads to edema and embryonic lethality in the mouse." Dev Biol **287**(2): 301-313.

d'Hennezel, E., et al. (2012). "The immunogenetics of immune dysregulation, polyendocrinopathy, enteropathy, X linked (IPEX) syndrome." J Med Genet **49**(5): 291-302.

Da Silva Xavier, G. (2018). "The Cells of the Islets of Langerhans." Journal of clinical medicine **7**(3): 54.

Davidson, H. W., et al. (1988). "Intraorganellar calcium and pH control proinsulin cleavage in the pancreatic beta cell via two distinct site-specific endopeptidases." Nature **333**(6168): 93-96.

De Franco, E., et al. (2015). "The effect of early, comprehensive genomic testing on clinical care in neonatal diabetes: an international cohort study." Lancet **386**(9997): 957-963.

De Franco, E., et al. (2017). "Dominant ER Stress-Inducing WFS1 Mutations Underlie a Genetic Syndrome of Neonatal/Infancy-Onset Diabetes, Congenital Sensorineural Deafness, and Congenital Cataracts." Diabetes **66**(7): 2044-2053.

De Franco, E., et al. (2013). "Biallelic PDX1 (insulin promoter factor 1) mutations causing neonatal diabetes without exocrine pancreatic insufficiency." Diabet Med **30**(5): e197-200.

De Franco, E., et al. (2019). "A Specific CNOT1 Mutation Results in a Novel Syndrome of Pancreatic Agenesis and Holoprosencephaly through Impaired Pancreatic and Neurological Development." Am J Hum Genet **104**(5): 985-989.

Dean, P. M. and E. K. Matthews (1970). "Glucose-induced electrical activity in pancreatic islet cells." J Physiol **210**(2): 255-264.

Deeb, S. S., et al. (1998). "A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity." Nat Genet **20**(3): 284-287.

Deeg, M. A., et al. (2007). "Pioglitazone and rosiglitazone have different effects on serum lipoprotein particle concentrations and sizes in patients with type 2 diabetes and dyslipidemia." Diabetes Care **30**(10): 2458-2464.

Del Prato, S., et al. (2009). "Changing the treatment paradigm for type 2 diabetes." Diabetes care **32 Suppl 2**(Suppl 2): S217-S222.

Delepine, M., et al. (2000). "EIF2AK3, encoding translation initiation factor 2-alpha kinase 3, is mutated in patients with Wolcott-Rallison syndrome." Nat Genet **25**(4): 406-409.

Diabetes-UK (2014). The Cost of Diabetes. https://www.diabetes.org.uk.

Do, R., et al. (2015). "Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction." Nature **518**(7537): 102-106.

Dodson, G. and D. Steiner (1998). "The role of assembly in insulin's biosynthesis." Curr Opin Struct Biol **8**(2): 189-194.

Donath, M. Y., et al. (2005). "Mechanisms of beta-cell death in type 2 diabetes." Diabetes **54 Suppl 2**: S108-113.

Dormandy, J. A., et al. (2005). "Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitAzone Clinical Trial In macroVascular Events): a randomised controlled trial." Lancet **366**(9493): 1279-1289.

Dufurrena, Q., et al. (2018). "Kalirin/Trio Rho GDP/GTP exchange factors regulate proinsulin and insulin secretion." J Mol Endocrinol.

Edghill, E. L., et al. (2006). "HLA genotyping supports a nonautoimmune etiology in patients diagnosed with diabetes under the age of 6 months." Diabetes **55**(6): 1895-1898.

Edghill, E. L., et al. (2008). "Insulin mutation screening in 1,044 patients with diabetes: mutations in the INS gene are a common cause of neonatal diabetes but a rare cause of diabetes diagnosed in childhood or adulthood." Diabetes **57**(4): 1034-1042.

Eipper, B. A., et al. (1992). "Alternative splicing and endoproteolytic processing generate tissue-specific forms of pituitary peptidylglycine alpha-amidating monooxygenase (PAM)." The Journal of biological chemistry **267**(6): 4008-4015.

Eipper, B. A. and R. E. Mains (1988). "Peptide alpha-amidation." Annu Rev Physiol **50**: 333-344.

Eipper, B. A., et al. (1983). "Identification in pituitary tissue of a peptide alpha-amidation activity that acts on glycine-extended peptides and requires molecular oxygen, copper, and ascorbic acid." Proc Natl Acad Sci U S A **80**(16): 5144-5148.

Eipper, B. A., et al. (1993). "Peptidylglycine alpha-amidating monooxygenase: a multifunctional protein with catalytic, processing, and routing domains." Protein Sci **2**(4): 489-497.

Eisenbarth, G. S. (1986). "Type I Diabetes Mellitus." New England Journal of Medicine **314**(21): 1360-1368.

El-Shanti, H., et al. (2000). "Homozygosity mapping identifies an additional locus for Wolfram syndrome on chromosome 4q." Am J Hum Genet **66**(4): 1229-1236.

El Meskini, R., et al. (2001). "Targeting of membrane proteins to the regulated secretory pathway in anterior pituitary endocrine cells." J Biol Chem **276**(5): 3384-3393.

El Meskini, R., et al. (2000). "Cell type-specific metabolism of peptidylglycine alpha-amidating monooxygenase in anterior pituitary." Endocrinology **141**(8): 3020-3034.

Ellard, S., et al. (2008). "Best practice guidelines for the molecular genetic diagnosis of maturity-onset diabetes of the young." Diabetologia **51**(4): 546-553.

Ellard, S., et al. (2013). "Improved genetic testing for monogenic diabetes using targeted next-generation sequencing." Diabetologia **56**(9): 1958-1963.

Ellgaard, L. and A. Helenius (2003). "Quality control in the endoplasmic reticulum." Nature Reviews Molecular Cell Biology **4**(3): 181-191.

Estrada, K., et al. (2014). "Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population." Jama **311**(22): 2305-2314.

Eyre-Walker, A. (2010). "Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies." Proceedings of the National Academy of Sciences **107**(suppl 1): 1752.

Faguer, S., et al. (2011). "Diagnosis, management, and prognosis of HNF1B nephropathy in adulthood." Kidney Int **80**(7): 768-776.

Fajans, S. S., et al. (2010). "Obesity and hyperinsulinemia in a family with pancreatic agenesis and MODY caused by the IPF1 mutation Pro63fsX60." Transl Res **156**(1): 7-14.

Fajans, S. S., et al. (2001). "Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young." N Engl J Med **345**(13): 971-980.

Fantus, I. G. and R. Brosseau (1986). "Mechanism of action of metformin: insulin receptor and postreceptor effects in vitro and in vivo." J Clin Endocrinol Metab **63**(4): 898-905.

Farquhar, M. G. and G. E. Palade (1981). "The Golgi apparatus (complex)-(1954-1981)-from artifact to center stage." J Cell Biol **91**(3 Pt 2): 77s-103s.

Fass, D. (2010). "Hunting for alternative disulfide bond formation pathways: endoplasmic reticulum janitor turns professor and teaches a lesson." Mol Cell **40**(5): 685-686.

Feng, D., et al. (2009). "Acute ablation of PERK results in ER dysfunctions followed by reduced insulin secretion and cell proliferation." BMC Cell Biol **10**: 61.

Ferraro, F., et al. (2007). "Kalirin/Trio Rho guanine nucleotide exchange factors regulate a novel step in secretory granule maturation." Molecular biology of the cell **18**(12): 4813-4825.

Flanagan, S. E., et al. (2014). "Analysis of transcription factors key for mouse pancreatic development establishes NKX2-2 and MNX1 mutations as causes of neonatal diabetes in man." Cell Metab **19**(1): 146-154.

Flanagan, S. E., et al. (2006). "Mutations in KCNJ11, which encodes Kir6.2, are a common cause of diabetes diagnosed in the first 6 months of life, with the phenotype determined by genotype." Diabetologia **49**(6): 1190-1197.

Flanagan, S. E., et al. (2014). "Activating germline mutations in STAT3 cause early-onset multi-organ autoimmune disease." Nat Genet **46**(8): 812-814.

Flanagan, S. E., et al. (2007). "Mutations in ATP-sensitive K+ channel genes cause transient neonatal diabetes and permanent diabetes in childhood or adulthood." Diabetes **56**(7): 1930-1937.

Flannick, J., et al. (2019). "Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls." Nature.

Flannick, J., et al. (2018). "Genetic discovery and translational decision support from exome sequencing of 20,791 type 2 diabetes cases and 24,440 controls from five ancestries." bioRxiv: 371450.

Flannick, J., et al. (2014). "Loss-of-function mutations in SLC30A8 protect against type 2 diabetes." Nature genetics **46**(4): 357-363.

Florez, J. C., et al. (2004). "Haplotype structure and genotype-phenotype correlations of the sulfonylurea receptor and the islet ATP-sensitive potassium channel gene region." Diabetes **53**(5): 1360-1368.

Fonseca, S. G., et al. (2009). "Endoplasmic reticulum stress in beta-cells and development of diabetes." Current opinion in pharmacology **9**(6): 763-770.

Fonseca, S. G., et al. (2005). "WFS1 is a novel component of the unfolded protein response and maintains homeostasis of the endoplasmic reticulum in pancreatic beta-cells." J Biol Chem **280**(47): 39609-39615.

Fra, A. M., et al. (1993). "Quality control of ER synthesized proteins: an exposed thiol group as a three-way switch mediating assembly, retention and degradation." Embo j **12**(12): 4755-4761.

Francioli, L. C., et al. (2015). "Genome-wide patterns and properties of de novo mutations in humans." Nat Genet **47**(7): 822-826.

Francone, V. P., et al. (2010). "Signaling from the secretory granule to the nucleus: Uhmk1 and PAM." Mol Endocrinol **24**(8): 1543-1558.

Fraser, F. C. and T. Gunn (1977). "Diabetes mellitus, diabetes insipidus, and optic atrophy. An autosomal recessive syndrome?" J Med Genet **14**(3): 190-193.

Frebourg, T. (2014). "The challenge for the next generation of medical geneticists." Hum Mutat **35**(8): 909-911.

Froguel, P., et al. (1993). "Familial hyperglycemia due to mutations in glucokinase. Definition of a subtype of diabetes mellitus." N Engl J Med **328**(10): 697-702.

Fuchsberger, C., et al. (2016). "The genetic architecture of type 2 diabetes." Nature **536**(7614): 41-47.

Fuchsberger, C., et al. (2016). "The genetic architecture of type 2 diabetes." Nature **536**(7614): 41-47.

Galan, M., et al. (2011). "Differential effects of HNF-1alpha mutations associated with familial young-onset diabetes on target gene regulation." Mol Med **17**(3-4): 256-265.

Galluzzi, L., et al. (2017). "Molecular definitions of autophagy and related processes." Embo j **36**(13): 1811-1836.

Gao, R., et al. (2014). "Evaluation of a target region capture sequencing platform using monogenic diabetes as a study-model." BMC Genet **15**: 13.

Garin, I., et al. (2010). "Recessive mutations in the INS gene result in neonatal diabetes through reduced insulin biosynthesis." Proceedings of the National Academy of Sciences of the United States of America **107**(7): 3105-3110.

Gastaldelli, A., et al. (2004). "Beta-cell dysfunction and glucose intolerance: results from the San Antonio metabolism (SAM) study." Diabetologia **47**(1): 31-39.

Gauss, R., et al. (2006). "A complex of Yos9p and the HRD ligase integrates endoplasmic reticulum quality control into the degradation machinery." Nat Cell Biol **8**(8): 849-854.

Gembal, M., et al. (1992). "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 **89**(4): 1288-1295.

Gepts, W. (1965). "Pathologic anatomy of the pancreas in juvenile diabetes mellitus." Diabetes **14**(10): 619-633.

Gloyn, A. L., et al. (2006). "KCNJ11 activating mutations are associated with developmental delay, epilepsy and neonatal diabetes syndrome and other neurological features." Eur J Hum Genet **14**(7): 824-830.

Gloyn, A. L., et al. (2001). "Association studies of variants in promoter and coding regions of beta-cell ATP-sensitive K-channel genes SUR1 and Kir6.2 with Type 2 diabetes mellitus (UKPDS 53)." Diabet Med **18**(3): 206-212.

Gloyn, A. L., et al. (2004). "Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes." N Engl J Med **350**(18): 1838-1849.

Gloyn, A. L., et al. (2003). "Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes." Diabetes **52**(2): 568-572.

Goldberg, R. B., et al. (2005). "A comparison of lipid and glycemic effects of pioglitazone and rosiglitazone in patients with type 2 diabetes and dyslipidemia." Diabetes Care **28**(7): 1547-1554.

Gorlov, I. P., et al. (2008). "Shifting paradigm of association studies: value of rare single-nucleotide polymorphisms." Am J Hum Genet **82**(1): 100-112.

Grant, S. F., et al. (2006). "Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes." Nat Genet **38**(3): 320-323.

Greeley, S. A. W., et al. (2011). "Neonatal diabetes: an expanding list of genes allows for improved diagnosis and treatment." Current diabetes reports **11**(6): 519-532.

Guan, W., et al. (2008). "Meta-analysis of 23 type 2 diabetes linkage studies from the International Type 2 Diabetes Linkage Analysis Consortium." Human heredity **66**(1): 35-49.

Guo, M. H., et al. (2018). "Burden Testing of Rare Variants Identified through Exome Sequencing via Publicly Available Control Data." Am J Hum Genet **103**(4): 522-534.

Gupta, S., et al. (2010). "PERK (EIF2AK3) regulates proinsulin trafficking and quality control in the secretory pathway." Diabetes **59**(8): 1937-1947.

Gurzov, E. N., et al. (2009). "Signaling by IL-1beta+IFN-gamma and ER stress converge on DP5/Hrk activation: a novel mechanism for pancreatic beta-cell apoptosis." Cell Death Differ **16**(11): 1539-1550.

Gusev, A., et al. (2014). "Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases." Am J Hum Genet **95**(5): 535-552.

Harding, H. P., et al. (2001). "Diabetes mellitus and exocrine pancreatic dysfunction in perk-/- mice reveals a role for translational control in secretory cell survival." Mol Cell **7**(6): 1153-1163.

Harding, H. P., et al. (2000). "Perk is essential for translational regulation and cell survival during the unfolded protein response." Mol Cell **5**(5): 897-904.

Hattersley, A., et al. (2006). "ISPAD Clinical Practice Consensus Guidelines 2006-2007. The diagnosis and management of monogenic diabetes in children." Pediatr Diabetes **7**(6): 352-360.

Hattersley, A., et al. (2009). "The diagnosis and management of monogenic diabetes in children and adolescents." Pediatr Diabetes **10 Suppl 12**: 33-42.

Hattersley, A. T. and F. M. Ashcroft (2005). "Activating mutations in Kir6.2 and neonatal diabetes: new clinical syndromes, new scientific insights, and new therapy." Diabetes **54**(9): 2503-2513.

Haynes, C. M., et al. (2004). "Degradation of misfolded proteins prevents ER-derived oxidative stress and cell death." Mol Cell **15**(5): 767-776.

Haze, K., et al. (1999). "Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress." Mol Biol Cell **10**(11): 3787-3799.

Hellman, R., et al. (1999). "The in vivo association of BiP with newly synthesized proteins is dependent on the rate and stability of folding and not simply on the presence of sequences that can bind to BiP." J Cell Biol **144**(1): 21-30.

Hemminki, K., et al. (2008). "The 'common disease-common variant' hypothesis and familial risks." PloS one **3**(6): e2504-e2504.

Henquin, J.-C. and M. Nenquin (2014). "Activators of PKA and Epac distinctly influence insulin secretion and cytosolic Ca2+ in female mouse islets stimulated by glucose and tolbutamide." Endocrinology **155**(9): 3274-3287.

Henquin, J. C. (2000). "Triggering and amplifying pathways of regulation of insulin secretion by glucose." Diabetes **49**(11): 1751-1760.

Henquin, J. C. (2004). "Pathways in beta-cell stimulus-secretion coupling as targets for therapeutic insulin secretagogues." Diabetes **53 Suppl 3**: S48-58.

Hermansen, K. (1985). "Forskolin, an activator of adenylate cyclase, stimulates pancreatic insulin, glucagon, and somatostatin release in the dog: studies in vitro." Endocrinology **116**(6): 2251-2258.

Hetz, C., et al. (2006). "Proapoptotic BAX and BAK modulate the unfolded protein response by a direct interaction with IRE1alpha." Science **312**(5773): 572-576.

Hidvegi, T., et al. (2005). "Accumulation of mutant alpha1-antitrypsin Z in the endoplasmic reticulum activates caspases-4 and -12, NFkappaB, and BAP31 but not the unfolded protein response." J Biol Chem **280**(47): 39002-39015.

Hill, C. P., et al. (1991). "X-ray structure of an unusual Ca2+ site and the roles of Zn2+ and Ca2+ in the assembly, stability, and storage of the insulin hexamer." Biochemistry **30**(4): 917-924.

Horikawa, Y., et al. (1997). "Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY." Nat Genet **17**(4): 384-385.

Horn, S. C., et al. (2009). "Usa1 functions as a scaffold of the HRD-ubiquitin ligase." Mol Cell **36**(5): 782-793.

Hosaka, M., et al. (2004). "Secretogranin III binds to cholesterol in the secretory granule membrane as an adapter for chromogranin A." J Biol Chem **279**(5): 3627-3634.

Hosaka, M., et al. (2002). "Identification of a chromogranin A domain that mediates binding to secretogranin III and targeting to secretory granules in pituitary cells and pancreatic beta-cells." Molecular biology of the cell **13**(10): 3388-3399.

Hoseki, J., et al. (2010). "Mechanism and components of endoplasmic reticulum-associated degradation." J Biochem **147**(1): 19-25.

Hoskinson, D. C., et al. (2017). "The current state of clinical interpretation of sequence variants." Current Opinion in Genetics & Development **42**: 33-39.

Houghton, J. A., et al. (2016). "Isolated Pancreatic Aplasia Due to a Hypomorphic PTF1A Mutation." Diabetes **65**(9): 2810-2815.

Houghton, J. A. L., et al. (2016). "Isolated Pancreatic Aplasia Due to a Hypomorphic PTF1A Mutation." Diabetes **65**(9): 2810-2815.

Hu, P., et al. (2006). "Autocrine tumor necrosis factor alpha links endoplasmic reticulum stress to the membrane death receptor pathway through IRE1alpha-mediated NF-kappaB activation and down-regulation of TRAF2 expression." Molecular and cellular biology **26**(8): 3071-3084.

Huang, X. F. and P. Arvan (1995). "Intracellular transport of proinsulin in pancreatic beta-cells. Structural maturation probed by disulfide accessibility." J Biol Chem **270**(35): 20417-20423.

Hundal, R. S., et al. (2000). "Mechanism by which metformin reduces glucose production in type 2 diabetes." Diabetes **49**(12): 2063-2069.

Huttner, W. B. and U. M. Benedum (1987). "Chromogranin A and pancreastatin." Nature **325**(6102): 305.

Hutton, J. C. (1982). "The internal pH and membrane potential of the insulin-secretory granule." The Biochemical journal **204**(1): 171-178.

Huyghe, J. R., et al. (2013). "Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion." Nat Genet **45**(2): 197-201.

Huyghe, J. R., et al. (2013). "Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion." Nature genetics **45**(2): 197-201.

Ichii, H., et al. (2005). "A novel method for the assessment of cellular composition and beta-cell viability in human islet preparations." Am J Transplant **5**(7): 1635-1645.

IDF (2017). IDF DIABETES ATLAS. http://diabetesatlas.org/resources/2017-atlas.html.

Ilonen, J., et al. (2013). "Patterns of beta-cell autoantibody appearance and genetic associations during the first years of life." Diabetes **62**(10): 3636-3640.

Iltz, J. L., et al. (2006). "Exenatide: an incretin mimetic for the treatment of type 2 diabetes mellitus." Clin Ther **28**(5): 652-665.

Inoue, H., et al. (1998). "A gene encoding a transmembrane protein is mutated in patients with diabetes mellitus and optic atrophy (Wolfram syndrome)." Nat Genet **20**(2): 143-148.

Ionescu-Tirgoviste, C., et al. (2015). "A 3D map of the islet routes throughout the healthy human pancreas." Sci Rep **5**: 14634.

Iyengar, S. K. and R. C. Elston (2007). "The genetic basis of complex traits: rare variants or "common gene, common disease"?" Methods Mol Biol **376**: 71-84.

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

Izumi, T., et al. (2003). "Dominant Negative Pathogenesis by Mutant Proinsulin in the Akita Diabetic Mouse." Diabetes **52**(2): 409.

Jessop, C. E., et al. (2009). "Protein disulphide isomerase family members show distinct substrate specificity: P5 is targeted to BiP client proteins." J Cell Sci **122**(Pt 23): 4287-4295.

Johansson, B. B., et al. (2011). "Diabetes and pancreatic exocrine dysfunction due to mutations in the carboxyl ester lipase gene-maturity onset diabetes of the young (CEL-MODY): a protein misfolding disease." J Biol Chem **286**(40): 34593-34605.

Johansson, J. K., et al. (2010). "N-cadherin is dispensable for pancreas development but required for beta-cell granule turnover." Genesis (New York, N.Y. : 2000) **48**(6): 374-381.

Johnson, M. B., et al. (2019). "Trisomy 21 is a Cause of Permanent Neonatal Diabetes that is Autoimmune but not HLA Associated." Diabetes.

Johnson, M. B., et al. (2017). "Recessively Inherited LRBA Mutations Cause Autoimmunity Presenting as Neonatal Diabetes." Diabetes **66**(8): 2316-2322.

Jones, P. M. and S. J. Persaud (1998). "Protein kinases, protein phosphorylation, and the regulation of insulin secretion from pancreatic beta-cells." Endocr Rev **19**(4): 429-461.

Jonikas, M. C., et al. (2009). "Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum." Science **323**(5922): 1693-1697.

Jun, G., et al. (2018). "Evaluating the contribution of rare variants to type 2 diabetes and related traits using pedigrees." Proc Natl Acad Sci U S A **115**(2): 379-384.

Jung, E. S., et al. (2015). "Acute ER stress regulates amyloid precursor protein processing through ubiquitin-dependent degradation." Scientific Reports **5**: 8805.

Juszczak, A., et al. (2016). "When to consider a diagnosis of MODY at the presentation of diabetes: aetiology matters for correct management." British Journal of General Practice **66**(647): e457.

Kahn, C. B., et al. (1969). "Clinical and chemical diabetes in offspring of diabetic couples." N Engl J Med **281**(7): 343-347.

Kahn, S. E., et al. (2008). "Rosiglitazone-associated fractures in type 2 diabetes: an Analysis from A Diabetes Outcome Progression Trial (ADOPT)." Diabetes Care **31**(5): 845-851.

Kakiuchi, C., et al. (2006). "XBP1 induces WFS1 through an endoplasmic reticulum stress response element-like motif in SH-SY5Y cells." J Neurochem **97**(2): 545-555.

Kaprio, J., et al. (1992). "Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland." Diabetologia **35**(11): 1060-1067.

Katopodis, A. G., et al. (1990). "A novel enzyme from bovine neurointermediate pituitary catalyzes dealkylation of alpha-hydroxyglycine derivatives, thereby functioning sequentially with peptidylglycine alpha-amidating monooxygenase in peptide amidation." Biochemistry **29**(26): 6115-6120.

Katsanis, S. H. and N. Katsanis (2013). "Molecular genetic testing and the future of clinical genomics." Nat Rev Genet **14**(6): 415-426.

Kiezun, A., et al. (2012). "Exome sequencing and the genetic basis of complex traits." Nat Genet **44**(6): 623-630.

Kim, T., et al. (2001). "Chromogranin A, an "on/off" switch controlling dense-core secretory granule biogenesis." Cell **106**(4): 499-509.

Klupa, T., et al. (2010). "Efficacy and safety of sulfonylurea use in permanent neonatal diabetes due to KCNJ11 gene mutations: 34-month median follow-up." Diabetes Technol Ther **12**(5): 387-391.

Knowler, W. C., et al. (2002). "Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin." N Engl J Med **346**(6): 393-403.

Koshimizu, H., et al. (2011). "Serpinin: a novel chromogranin A-derived, secreted peptide up-regulates protease nexin-1 expression and granule biogenesis in endocrine cells." Molecular endocrinology **25**(5): 732-744.

Koshimizu, H., et al. (2010). "Chromogranin A: a new proposal for trafficking, processing and induction of granule biogenesis." Regulatory peptides **160**(1-3): 153-159.

Kowalski, J. M., et al. (1998). "Protein folding stability can determine the efficiency of escape from endoplasmic reticulum quality control." J Biol Chem **273**(31): 19453-19458.

Kowarik, M., et al. (2002). "Protein folding during cotranslational translocation in the endoplasmic reticulum." Mol Cell **10**(4): 769-778.

Krischer, J. P., et al. (2015). "The 6 year incidence of diabetes-associated autoantibodies in genetically at-risk children: the TEDDY study." Diabetologia **58**(5): 980-987.

Kristinsson, S. Y., et al. (2001). "MODY in Iceland is associated with mutations in HNF-1alpha and a novel mutation in NeuroD1." Diabetologia **44**(11): 2098-2103.

Kumar, D., et al. (2016). "60 YEARS OF POMC: From POMC and α-MSH to PAM, molecular oxygen, copper, and vitamin C." Journal of molecular endocrinology **56**(4): T63-T76.

Kunishima, S., et al. (2013). "ACTN1 mutations cause congenital macrothrombocytopenia." Am J Hum Genet **92**(3): 431-438.

Lander, E. S. and D. Botstein (1987). "Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children." Science **236**(4808): 1567-1570.

Lee, A. H., et al. (2003). "XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response." Mol Cell Biol **23**(21): 7448-7459.

Lee, J., et al. (2011). "p38 MAPK-mediated regulation of Xbp1s is crucial for glucose homeostasis." Nat Med **17**(10): 1251-1260.

Lee, K. P. K., et al. (2008). "Structure of the dual enzyme Ire1 reveals the basis for catalysis and regulation in nonconventional RNA splicing." Cell **132**(1): 89-100.

Lee, M. C., et al. (2004). "Bi-directional protein transport between the ER and Golgi." Annu Rev Cell Dev Biol **20**: 87-123.

Li, B. and S. M. Leal (2008). "Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data." Am J Hum Genet **83**(3): 311-321.

Lipson, K. L., et al. (2006). "Regulation of insulin biosynthesis in pancreatic beta cells by an endoplasmic reticulum-resident protein kinase IRE1." Cell Metab **4**(3): 245-254.

Liu, L., et al. (2013). "Mutations in KCNJ11 are associated with the development of autosomal dominant, early-onset type 2 diabetes." Diabetologia **56**(12): 2609-2618.

Liu, M., et al. (2007). "Proinsulin maturation, misfolding, and proteotoxicity." Proc Natl Acad Sci U S A **104**(40): 15841-15846.

Loh, Y. P., et al. (2012). "Serpinins: role in granule biogenesis, inhibition of cell death and cardiac function." Curr Med Chem **19**(24): 4086-4092.

Ludwig, M. G., et al. (2003). "Proton-sensing G-protein-coupled receptors." Nature **425**(6953): 93-98.

MacArthur, D. G., et al. (2014). "Guidelines for investigating causality of sequence variants in human disease." Nature **508**(7497): 469-476.

Mackay, D. J. G., et al. (2008). "Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57." Nature Genetics **40**: 949.

Madsbad, S., et al. (2011). "An overview of once-weekly glucagon-like peptide-1 receptor agonists--available efficacy and safety data and perspectives for the future." Diabetes Obes Metab **13**(5): 394-407.

Mahajan, A., et al. (2014). "Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility." Nat Genet **46**(3): 234-244.

Mahajan, A., et al. (2015). "Identification and functional characterization of G6PC2 coding variants influencing glycemic traits define an effector transcript at the G6PC2-ABCB11 locus." PLoS Genet **11**(1): e1004876.

Mahajan, A., et al. (2018). "Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps." Nat Genet **50**(11): 1505-1513.

Mahajan, A., et al. (2018). "Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes." Nature Genetics **50**(4): 559-571.

Mains, R. E., et al. (1995). "The NH2-terminal proregion of peptidylglycine alpha-amidating monooxygenase facilitates the secretion of soluble proteins." Mol Endocrinol **9**(1): 3-13.

Malecki, M. T., et al. (1999). "Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus." Nat Genet **23**(3): 323-328.

Manolio, T. A., et al. (2009). "Finding the missing heritability of complex diseases." Nature **461**(7265): 747-753.

Marciniak, S. J., et al. (2004). "CHOP induces death by promoting protein synthesis and oxidation in the stressed endoplasmic reticulum." Genes Dev **18**(24): 3066-3077.

Marie, P. J., et al. (2014). "Cadherin-mediated cell-cell adhesion and signaling in the skeleton." Calcif Tissue Int **94**(1): 46-54.

Marouli, E., et al. (2017). "Rare and low-frequency coding variants alter human adult height." Nature **542**(7640): 186-190.

Martinez, A., et al. (1993). "Localization of amidating enzymes (PAM) in rat gastrointestinal tract." J Histochem Cytochem **41**(11): 1617-1622.

Martinez, A., et al. (1993). "Immunocytochemical localization of peptidylglycine alpha-amidating monooxygenase enzymes (PAM) in human endocrine pancreas." J Histochem Cytochem **41**(3): 375-380.

Matschinsky, F. M. (1996). "Banting Lecture 1995. A lesson in metabolic regulation inspired by the glucokinase glucose sensor paradigm." Diabetes **45**(2): 223-241.

Matschinsky, F. M., et al. (1998). "Pancreatic beta-cell glucokinase: closing the gap between theoretical concepts and experimental realities." Diabetes **47**(3): 307-315.

Matthijs, G., et al. (2016). "Guidelines for diagnostic next-generation sequencing." Eur J Hum Genet **24**(1): 2-5.

Mauthe, M., et al. (2018). "Chloroquine inhibits autophagic flux by decreasing autophagosome-lysosome fusion." Autophagy **14**(8): 1435-1455.

Mauvais-Jarvis, F., et al. (2004). "PAX4 gene variations predispose to ketosis-prone diabetes." Hum Mol Genet **13**(24): 3151-3159.

Mayor, S. (2019). "Diabetes: precision approach will improve outcomes, says research director." BMJ **364**: l186.

McCarthy, M. I. (2017). "Painting a new picture of personalised medicine for diabetes." Diabetologia **60**(5): 793-799.

McClellan, J. and M. C. King (2010). "Genetic heterogeneity in human disease." Cell **141**(2): 210-217.

McCulloch, L. J., et al. (2011). "GLUT2 (SLC2A2) is not the principal glucose transporter in human pancreatic beta cells: implications for understanding genetic association signals at this locus." Mol Genet Metab **104**(4): 648-653.

Meigs, J. B., et al. (2000). "Parental transmission of type 2 diabetes: the Framingham Offspring Study." Diabetes **49**(12): 2201-2207.

Merkler, D. J. (1994). "C-terminal amidated peptides: production by the in vitro enzymatic amidation of glycine-extended peptides and the importance of the amide to bioactivity." Enzyme and microbial technology **16**(6): 450-456.

Merksamer, P. I., et al. (2008). "Real-time redox measurements during endoplasmic reticulum stress reveal interlinked protein folding functions." Cell **135**(5): 933-947.

Meur, G., et al. (2010). "Insulin gene mutations resulting in early-onset diabetes: marked differences in clinical presentation, metabolic status, and pathogenic effect through endoplasmic reticulum retention." Diabetes **59**(3): 653-661.

Meusser, B., et al. (2005). "ERAD: the long road to destruction." Nat Cell Biol **7**(8): 766-772.

Milgram, S. L., et al. (1992). "Expression of individual forms of peptidylglycine alpha-amidating monooxygenase in AtT-20 cells: endoproteolytic processing and routing to secretory granules." J Cell Biol **117**(4): 717-728.

Milgram, S. L., et al. (1997). "Localization of integral membrane peptidylglycine alpha-amidating monooxygenase in neuroendocrine cells." Journal of cell science **110 (Pt 6)**: 695-706.

Milgram, S. L., et al. (1996). "Identification of routing determinants in the cytosolic domain of a secretory granule-associated integral membrane protein." The Journal of biological chemistry **271**(29): 17526-17535.

Miller, R. A., et al. (2013). "Biguanides suppress hepatic glucagon signalling by decreasing production of cyclic AMP." Nature **494**(7436): 256-260.

Montoli, A., et al. (2002). "Renal cysts and diabetes syndrome linked to mutations of the hepatocyte nuclear factor-1 beta gene: description of a new family with associated liver involvement." Am J Kidney Dis **40**(2): 397-402.

Morris, A. P., et al. (2012). "Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes." Nat Genet **44**(9): 981-990.

Moutsianas, L., et al. (2015). "The power of gene-based rare variant methods to detect disease-associated variation and test hypotheses about complex disease." PLoS Genet **11**(4): e1005165.

Nathan, D. M., et al. (2009). "Medical management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy: a consensus statement of the American Diabetes Association and the European Association for the Study of Diabetes." Diabetes care **32**(1): 193-203.

Nejentsev, S., et al. (1999). "Population-based genetic screening for the estimation of Type 1 diabetes mellitus risk in Finland: selective genotyping of markers in the HLA-DQB1, HLA-DQA1 and HLA-DRB1 loci." Diabet Med **16**(12): 985-992.

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

Newman, B., et al. (1987). "Concordance for type 2 (non-insulin-dependent) diabetes mellitus in male twins." Diabetologia **30**(10): 763-768.

Ng, S. B., et al. (2009). "Targeted capture and massively parallel sequencing of 12 human exomes." Nature **461**(7261): 272-276.

Nicolino, M., et al. (2010). "A novel hypomorphic PDX1 mutation responsible for permanent neonatal diabetes with subclinical exocrine deficiency." Diabetes **59**(3): 733-740.

Nieto-Estevez, V. and J. Hsieh (2018). "CHD2: One Gene, Many Roles." Neuron **100**(5): 1014-1016.

Nishitoh, H., et al. (2002). "ASK1 is essential for endoplasmic reticulum stress-induced neuronal cell death triggered by expanded polyglutamine repeats." Genes Dev **16**(11): 1345-1355.

Nissen, S. E. and K. Wolski (2007). "Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes." N Engl J Med **356**(24): 2457-2471.

Njolstad, P. R., et al. (2001). "Neonatal diabetes mellitus due to complete glucokinase deficiency." N Engl J Med **344**(21): 1588-1592.

Odom, D. T., et al. (2004). "Control of pancreas and liver gene expression by HNF transcription factors." Science **303**(5662): 1378-1381.

Oikawa, D., et al. (2009). "Activation of mammalian IRE1alpha upon ER stress depends on dissociation of BiP rather than on direct interaction with unfolded proteins." Exp Cell Res **315**(15): 2496-2504.

Onn, A. and D. Ron (2010). "Modeling the endoplasmic reticulum unfolded protein response." Nat Struct Mol Biol **17**(8): 924-925.

Orci, L., et al. (1985). "Direct identification of prohormone conversion site in insulin-secreting cells." Cell **42**(2): 671-681.

Osbak, K. K., et al. (2009). "Update on mutations in glucokinase (GCK), which cause maturity-onset diabetes of the young, permanent neonatal diabetes, and hyperinsulinemic hypoglycemia." Hum Mutat **30**(11): 1512-1526.

Ouafik, L. H., et al. (1992). "The multifunctional peptidylglycine alpha-amidating monooxygenase gene: exon/intron organization of catalytic, processing, and routing domains." Mol Endocrinol **6**(10): 1571-1584.

Oyadomari, S., et al. (2002). "Targeted disruption of the Chop gene delays endoplasmic reticulum stress-mediated diabetes." J Clin Invest **109**(4): 525-532.

Pace, C. S. and G. Sachs (1982). "Glucose-induced proton uptake in secretory granules of beta-cells in monolayer culture." American Journal of Physiology-Cell Physiology **242**(5): C382-C387.

Patel, K. A., et al. (2016). "Type 1 Diabetes Genetic Risk Score: A Novel Tool to Discriminate Monogenic and Type 1 Diabetes." Diabetes **65**(7): 2094-2099.

Pearson, E. R., et al. (2006). "Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations." N Engl J Med **355**(5): 467-477.

Pearson, E. R., et al. (2003). "Genetic cause of hyperglycaemia and response to treatment in diabetes." Lancet **362**(9392): 1275-1281.

Pearson, E. R., et al. (2001). "beta-cell genes and diabetes: quantitative and qualitative differences in the pathophysiology of hepatic nuclear factor-1alpha and glucokinase mutations." Diabetes **50 Suppl 1**: S101-107.

Perkins, S. N., et al. (1990). "The 108-kDA peptidylglycine alpha-amidating monooxygenase precursor contains two separable enzymatic activities involved in peptide amidation." Biochem Biophys Res Commun **171**(3): 926-932.

Perry, D. J., et al. (2018). "Application of a Genetic Risk Score to Racially Diverse Type 1 Diabetes Populations Demonstrates the Need for Diversity in Risk-Modeling." Sci Rep **8**(1): 4529.

Phani, N. M., et al. (2014). "Population specific impact of genetic variants in KCNJ11 gene to type 2 diabetes: a case-control and meta-analysis study." PLoS One **9**(9): e107021.

Pincus, D., et al. (2010). "BiP binding to the ER-stress sensor Ire1 tunes the homeostatic behavior of the unfolded protein response." PLoS Biol **8**(7): e1000415.

Plengvidhya, N., et al. (2007). "PAX4 mutations in Thais with maturity onset diabetes of the young." J Clin Endocrinol Metab **92**(7): 2821-2826.

Polak, M. and H. Cavé (2007). "Neonatal diabetes mellitus: a disease linked to multiple mechanisms." Orphanet journal of rare diseases **2**: 12-12.

Porter, A. G. and R. U. Janicke (1999). "Emerging roles of caspase-3 in apoptosis." Cell Death Differ **6**(2): 99-104.

Poulsen, P., et al. (1999). "Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study." Diabetologia **42**(2): 139-145.

Poulton, C. J., et al. (2011). "Microcephaly with simplified gyration, epilepsy, and infantile diabetes linked to inappropriate apoptosis of neural progenitors." Am J Hum Genet **89**(2): 265-276.

Powers, E. T., et al. (2009). "Biological and chemical approaches to diseases of proteostasis deficiency." Annu Rev Biochem **78**: 959-991.

Prasad, R. B. and L. Groop (2015). "Genetics of type 2 diabetes-pitfalls and possibilities." Genes **6**(1): 87-123.

Price, A. L., et al. (2010). "Pooled association tests for rare variants in exon-resequencing studies." Am J Hum Genet **86**(6): 832-838.

Raeder, H., et al. (2006). "Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction." Nat Genet **38**(1): 54-62.

Rafiq, M., et al. (2008). "Effective Treatment With Oral Sulfonylureas in Patients With Diabetes Due to Sulfonylurea Receptor 1 (SUR1) Mutations." Diabetes Care **31**(2): 204.

Rafiq, M., et al. (2008). "Effective treatment with oral sulfonylureas in patients with diabetes due to sulfonylurea receptor 1 (SUR1) mutations." Diabetes Care **31**(2): 204-209.

Raile, K., et al. (2015). "Treatment of young patients with HNF1A mutations (HNF1A-MODY)." Diabet Med **32**(4): 526-530.

Rajagopal, C., et al. (2010). "Secretion stimulates intramembrane proteolysis of a secretory granule membrane enzyme." J Biol Chem **285**(45): 34632-34642.

Ravassard, P., et al. (2011). "A genetically engineered human pancreatic beta cell line exhibiting glucose-inducible insulin secretion." J Clin Invest **121**(9): 3589-3597.

Reich, D. E. and E. S. Lander (2001). "On the allelic spectrum of human disease." Trends Genet **17**(9): 502-510.

Rewers, M., et al. (1996). "Newborn screening for HLA markers associated with IDDM: diabetes autoimmunity study in the young (DAISY)." Diabetologia **39**(7): 807-812.

Rhoads, A. and K. F. Au (2015). "PacBio Sequencing and Its Applications." Genomics Proteomics Bioinformatics **13**(5): 278-289.

Richards, S., et al. (2015). "Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology." Genetics in medicine : official journal of the American College of Medical Genetics **17**(5): 405-424.

Rodriguez-Castaneda, F., et al. (2018). "The SUMO protease SENP1 and the chromatin remodeler CHD3 interact and jointly affect chromatin accessibility and gene expression." J Biol Chem **293**(40): 15439-15454.

Romero, A., et al. (2008). "Role of a Pro-sequence in the Secretory Pathway of Prothyrotropin-releasing Hormone." **283**(46): 31438-31448.

Ron, D. and P. Walter (2007). "Signal integration in the endoplasmic reticulum unfolded protein response." Nat Rev Mol Cell Biol **8**(7): 519-529.

Rubio-Cabezas, O. and S. Ellard (2013). "Diabetes mellitus in neonates and infants: genetic heterogeneity, clinical approach to diagnosis, and therapeutic options." Horm Res Paediatr **80**(3): 137-146.

Rubio-Cabezas, O., et al. (2012). "KATP channel mutations in infants with permanent diabetes diagnosed after 6 months of life." Pediatr Diabetes **13**(4): 322-325.

Rubio-Cabezas, O., et al. (2011). "Permanent Neonatal Diabetes and Enteric Anendocrinosis Associated With Biallelic Mutations in NEUROG3." Diabetes **60**(4): 1349-1353.

Rubio-Cabezas, O., et al. (2010). "Homozygous mutations in NEUROD1 are responsible for a novel syndrome of permanent neonatal diabetes and neurological abnormalities." Diabetes **59**(9): 2326-2331.

Rubio-Cabezas, O., et al. (2009). "Wolcott-Rallison syndrome is the most common genetic cause of permanent neonatal diabetes in consanguineous families." J Clin Endocrinol Metab **94**(11): 4162-4170.

Rutkowski, D. T. and R. J. Kaufman (2007). "That which does not kill me makes me stronger: adapting to chronic ER stress." Trends Biochem Sci **32**(10): 469-476.

Sakuraba, H., et al. (2002). "Reduced beta-cell mass and expression of oxidative stress-related DNA damage in the islet of Japanese Type II diabetic patients." Diabetologia **45**(1): 85-96.

Saleheen, D., et al. (2017). "Human knockouts and phenotypic analysis in a cohort with a high rate of consanguinity." Nature **544**(7649): 235-239.

Sales Gil, R. and P. Vagnarelli (2018). "Ki-67: More Hidden behind a 'Classic Proliferation Marker'." Trends Biochem Sci **43**(10): 747-748.

Sandhu, M. S., et al. (2007). "Common variants in WFS1 confer risk of type 2 diabetes." Nat Genet **39**(8): 951-953.

Sano, R. and J. C. Reed (2013). "ER stress-induced cell death mechanisms." Biochim Biophys Acta **1833**(12): 3460-3470.

Santer, R., et al. (1997). "Mutations in GLUT2, the gene for the liver-type glucose transporter, in patients with Fanconi-Bickel syndrome." Nat Genet **17**(3): 324-326.

Satin LS, S. P. (1994). "Electrical bursting in beta cells of the pancreatic islets of Langerhans." Endocrine **2**: 677-687.

Sato, Y., et al. (1992). "Dual functional role of membrane depolarization/Ca2+ influx in rat pancreatic B-cell." Diabetes **41**(4): 438-443.

Saxena, R., et al. (2007). "Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels." Science **316**(5829): 1331-1336.

Scharfmann, R., et al. (1988). "Persistence of peptidylglycine alpha-amidating monooxygenase activity and elevated thyrotropin-releasing hormone concentrations in fetal rat islets in culture." Endocrinology **123**(3): 1329-1334.

Scheen, A. J. and N. Paquot (2013). "Metformin revisited: a critical review of the benefit-risk balance in at-risk patients with type 2 diabetes." Diabetes Metab **39**(3): 179-190.

Schindler, A. J. and R. Schekman (2009). "In vitro reconstitution of ER-stress induced ATF6 transport in COPII vesicles." Proc Natl Acad Sci U S A **106**(42): 17775-17780.

Schuit, F. C., et al. (1988). "Glucose stimulates proinsulin biosynthesis by a dose-dependent recruitment of pancreatic beta cells." Proc Natl Acad Sci U S A **85**(11): 3865-3869.

Schulz, J., et al. (2017). "Conserved cytoplasmic domains promote Hrd1 ubiquitin ligase complex formation for ER-associated degradation (ERAD)." J Cell Sci **130**(19): 3322-3335.

Schulze, H., et al. (2009). "Principles of lysosomal membrane degradation: Cellular topology and biochemistry of lysosomal lipid degradation." Biochim Biophys Acta **1793**(4): 674-683.

Scott, L. J., et al. (2007). "A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants." Science **316**(5829): 1341-1345.

Seamon, K. B. and J. W. Daly (1981). "Forskolin: a unique diterpene activator of cyclic AMP-generating systems." J Cyclic Nucleotide Res **7**(4): 201-224.

Seino, S. (1999). "ATP-sensitive potassium channels: a model of heteromultimeric potassium channel/receptor assemblies." Annu Rev Physiol **61**: 337-362.

Sellick, G. S., et al. (2004). "Mutations in PTF1A cause pancreatic and cerebellar agenesis." Nat Genet **36**(12): 1301-1305.

Senee, V., et al. (2006). "Mutations in GLIS3 are responsible for a rare syndrome with neonatal diabetes mellitus and congenital hypothyroidism." Nat Genet **38**(6): 682-687.

Sforca, M. L., et al. (2004). "How C-terminal carboxyamidation alters the biological activity of peptides from the venom of the eumenine solitary wasp." Biochemistry **43**(19): 5608-5617.

Shaat, N., et al. (2006). "Common variants in MODY genes increase the risk of gestational diabetes mellitus." Diabetologia **49**(7): 1545-1551.

Shalev, D. E., et al. (2002). "Structural consequences of carboxyamidation of dermaseptin S3." Biochemistry **41**(23): 7312-7317.

Sharp, G. W. (1979). "The adenylate cyclase-cyclic AMP system in islets of Langerhans and its role in the control of insulin release." Diabetologia **16**(5): 287-296.

Sharp, S. A., et al. (2019). "Development and Standardization of an Improved Type 1 Diabetes Genetic Risk Score for Use in Newborn Screening and Incident Diagnosis." Diabetes Care **42**(2): 200-207.

Shaw-Smith, C., et al. (2014). "GATA4 mutations are a cause of neonatal and childhood-onset diabetes." Diabetes **63**(8): 2888-2894.

Shield, J. P. (2000). "Neonatal diabetes: new insights into aetiology and implications." Horm Res **53 Suppl 1**: 7-11.

Shields, B. M., et al. (2010). "Maturity-onset diabetes of the young (MODY): how many cases are we missing?" Diabetologia **53**(12): 2504-2508.

Shinde, U. and M. Inouye (1993). "Intramolecular chaperones and protein folding." Trends Biochem Sci **18**(11): 442-446.

Sladek, R., et al. (2007). "A genome-wide association study identifies novel risk loci for type 2 diabetes." Nature **445**(7130): 881-885.

Slager, S. L., et al. (2000). "Effect of allelic heterogeneity on the power of the transmission disequilibrium test." Genet Epidemiol **18**(2): 143-156.

Smith, D. J. and A. J. Lusis (2002). "The allelic structure of common disease." Hum Mol Genet **11**(20): 2455-2461.

Smith, S. B., et al. (2010). "Rfx6 directs islet formation and insulin production in mice and humans." Nature **463**(7282): 775-780.

Snijders Blok, L., et al. (2018). "CHD3 helicase domain mutations cause a neurodevelopmental syndrome with macrocephaly and impaired speech and language." Nat Commun **9**(1): 4619.

Sobota, J. A., et al. (2006). "Not all secretory granules are created equal: Partitioning of soluble content proteins." Mol Biol Cell **17**(12): 5038-5052.

Solomon, B. D., et al. (2009). "Compound heterozygosity for mutations in PAX6 in a patient with complex brain anomaly, neonatal diabetes mellitus, and microophthalmia." Am J Med Genet A **149a**(11): 2543-2546.

Spijkerman, A. M., et al. (2003). "Microvascular complications at time of diagnosis of type 2 diabetes are similar among diabetic patients detected by targeted screening and patients newly diagnosed in general practice: the hoorn screening study." Diabetes Care **26**(9): 2604-2608.

Stagg, S. M., et al. (2006). "Structure of the Sec13/31 COPII coat cage." Nature **439**(7073): 234-238.

Stancakova, A. and M. Laakso (2016). "Genetics of Type 2 Diabetes." Endocr Dev **31**: 203-220.

Stanescu, D. E., et al. (2012). "Novel presentations of congenital hyperinsulinism due to mutations in the MODY genes: HNF1A and HNF4A." J Clin Endocrinol Metab **97**(10): E2026-2030.

Steele, A. M., et al. (2014). "Prevalence of vascular complications among patients with glucokinase mutations and prolonged, mild hyperglycemia." Jama **311**(3): 279-286.

Steiner, D. F. and D. E. James (1992). "Cellular and molecular biology of the Beta cell." Diabetologia **35**(2): S41-S48.

Steinthorsdottir, V., et al. (2014). "Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes." Nat Genet **46**(3): 294-298.

Steinthorsdottir, V., et al. (2014). "Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes." Nat Genet **46**(3): 294-298.

Steveson, T. C., et al. (1999). "Phosphorylation of cytosolic domain Ser(937) affects both biosynthetic and endocytic trafficking of peptidylglycine alpha-amidating monooxygenase." J Biol Chem **274**(30): 21128-21138.

Steveson, T. C., et al. (2001). "Access of a membrane protein to secretory granules is facilitated by phosphorylation." J Biol Chem **276**(43): 40326-40337.

Stewart, L. C. and J. P. Klinman (1988). "Dopamine beta-hydroxylase of adrenal chromaffin granules: structure and function." Annu Rev Biochem **57**: 551-592.

Stoffers, D. A., et al. (1997). "Early-onset type-II diabetes mellitus (MODY4) linked to IPF1." Nat Genet **17**(2): 138-139.

Stoffers, D. A., et al. (1997). "Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence." Nat Genet **15**(1): 106-110.

Stoy, J., et al. (2007). "Insulin gene mutations as a cause of permanent neonatal diabetes." Proc Natl Acad Sci U S A **104**(38): 15040-15044.

Stride, A., et al. (2005). "Beta-cell dysfunction, insulin sensitivity, and glycosuria precede diabetes in hepatocyte nuclear factor-1alpha mutation carriers." Diabetes Care **28**(7): 1751-1756.

Strom, T. M., et al. (1998). "Diabetes insipidus, diabetes mellitus, optic atrophy and deafness (DIDMOAD) caused by mutations in a novel gene (wolframin) coding for a predicted transmembrane protein." Hum Mol Genet **7**(13): 2021-2028.

Stumvoll, M., et al. (2005). "Type 2 diabetes: principles of pathogenesis and therapy." Lancet **365**(9467): 1333-1346.

Szegezdi, E., et al. (2006). "Mediators of endoplasmic reticulum stress-induced apoptosis." EMBO reports **7**(9): 880-885.

Tang, G., et al. (2001). "Inhibition of JNK activation through NF-kappaB target genes." Nature **414**(6861): 313-317.

Temple, I. K., et al. (1995). "An imprinted gene(s) for diabetes?" Nat Genet **9**(2): 110-112.

Tennessen, J. A., et al. (2012). "Evolution and functional impact of rare coding variation from deep sequencing of human exomes." Science **337**(6090): 64-69.

Thiele, E. A., et al. (1989). "Tissue-specific regulation of peptidyl-glycine alpha-amidating monooxygenase expression." Endocrinology **125**(5): 2279-2288.

Thomsen, S. K., et al. (2018). "Type 2 diabetes risk alleles in PAM impact insulin release from human pancreatic β-cells." Nature Genetics **50**(8): 1122-1131.

Tirasophon, W., et al. (2000). "The endoribonuclease activity of mammalian IRE1 autoregulates its mRNA and is required for the unfolded protein response." Genes Dev **14**(21): 2725-2736.

Tontonoz, P., et al. (1998). "PPARgamma promotes monocyte/macrophage differentiation and uptake of oxidized LDL." Cell **93**(2): 241-252.

Tooze, S. A., et al. (2014). "Endocytosis and autophagy: exploitation or cooperation?" Cold Spring Harb Perspect Biol **6**(5): a018358.

Travers, K. J., et al. (2000). "Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation." Cell **101**(3): 249-258.

Tsonkova, V. G., et al. (2017). "The EndoC-βH1 cell line is a valid model of human beta cells and applicable for screenings to identify novel drug target candidates." Molecular metabolism **8**: 144-157.

Tuomilehto, J., et al. (2001). "Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance." N Engl J Med **344**(18): 1343-1350.

Ulinski, T., et al. (2006). "Renal phenotypes related to hepatocyte nuclear factor-1beta (TCF2) mutations in a pediatric cohort." J Am Soc Nephrol **17**(2): 497-503.

Urano, F., et al. (2000). "Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1." Science **287**(5453): 664-666.

van de Bunt, M. and A. L. Gloyn (2012). "A tale of two glucose transporters: how GLUT2 re-emerged as a contender for glucose transport into the human beta cell." Diabetologia **55**(9): 2312-2315.

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

Vishwanatha, K., et al. (2014). "A histidine-rich linker region in peptidylglycine α-amidating monooxygenase has the properties of a pH sensor." The Journal of biological chemistry **289**(18): 12404-12420.

Voight, B. F., et al. (2010). "Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis." Nat Genet **42**(7): 579-589.

von Heijne, G. (1983). "Patterns of amino acids near signal-sequence cleavage sites." Eur J Biochem **133**(1): 17-21.

Vrijenhoek, T., et al. (2015). "Next-generation sequencing-based genome diagnostics across clinical genetics centers: implementation choices and their effects." Eur J Hum Genet **23**(9): 1142-1150.

Wakeling, M. N., et al. (2019). "Homozygosity mapping provides supporting evidence of pathogenicity in recessive Mendelian disease." Genetics in Medicine **21**(4): 982-986.

Walter, P. and D. Ron (2011). "The unfolded protein response: from stress pathway to homeostatic regulation." Science **334**(6059): 1081-1086.

Wang, J., et al. (1999). "A mutation in the insulin 2 gene induces diabetes with severe pancreatic beta-cell dysfunction in the Mody mouse." J Clin Invest **103**(1): 27-37.

Weedon, M. N., et al. (2014). "Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agenesis." Nat Genet **46**(1): 61-64.

Wegner, L., et al. (2007). "Common variation in LMNA increases susceptibility to type 2 diabetes and associates with elevated fasting glycemia and estimates of body fat and height in the general population: studies of 7,495 Danish whites." Diabetes **56**(3): 694-698.

WHO. (2016). "Diabetes." 2016, from http://www.who.int/mediacentre/factsheets/fs312/en/.

Wiedenkeller, D. E. and G. W. Sharp (1983). "Effects of forskolin on insulin release and cyclic AMP content in rat pancreatic islets." Endocrinology **113**(6): 2311-2313.

Wollam, J., et al. (2017). "Chromogranin A regulates vesicle storage and mitochondrial dynamics to influence insulin secretion." Cell Tissue Res **368**(3): 487-501.

Woolhead, C. A., et al. (2004). "Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins." Cell **116**(5): 725-736.

Wright, N. M., et al. (1993). "Permanent neonatal diabetes mellitus and pancreatic exocrine insufficiency resulting from congenital pancreatic agenesis." Am J Dis Child **147**(6): 607-609.

Yamada, T., et al. (2006). "WFS1-deficiency increases endoplasmic reticulum stress, impairs cell cycle progression and triggers the apoptotic pathway specifically in pancreatic beta-cells." Hum Mol Genet **15**(10): 1600-1609.

Yamagata, K., et al. (1996). "Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3)." Nature **384**(6608): 455-458.

Yau, J. W., et al. (2012). "Global prevalence and major risk factors of diabetic retinopathy." Diabetes Care **35**(3): 556-564.

Ye, J., et al. (2000). "ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs." Mol Cell **6**(6): 1355-1364.

Yesilkaya, E., et al. (2009). "A novel mutation in the SLC19A2 gene in a Turkish female with thiamine-responsive megaloblastic anemia syndrome." J Trop Pediatr **55**(4): 265-267.

Yoneda, T., et al. (2001). "Activation of caspase-12, an endoplastic reticulum (ER) resident caspase, through tumor necrosis factor receptor-associated factor 2-dependent mechanism in response to the ER stress." J Biol Chem **276**(17): 13935-13940.

Yoon, K. H., et al. (2003). "Selective beta-cell loss and alpha-cell expansion in patients with type 2 diabetes mellitus in Korea." J Clin Endocrinol Metab **88**(5): 2300-2308.

Yoshida, H., et al. (2003). "A time-dependent phase shift in the mammalian unfolded protein response." Dev Cell **4**(2): 265-271.

Yoshioka, M., et al. (1997). "A novel locus, Mody4, distal to D7Mit189 on chromosome 7 determines early-onset NIDDM in nonobese C57BL/6 (Akita) mutant mice." Diabetes **46**(5): 887-894.

Yun, H. Y., et al. (1995). "Phosphorylation of the cytosolic domain of peptidylglycine alpha-amidating monooxygenase." J Biol Chem **270**(50): 30075-30083.

Zeggini, E., et al. (2007). "Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes." Science **316**(5829): 1336-1341.

Zhang, C., et al. (2001). "Activation of JNK and transcriptional repressor ATF3/LRF1 through the IRE1/TRAF2 pathway is implicated in human vascular endothelial cell death by homocysteine." Biochem Biophys Res Commun **289**(3): 718-724.

Zhang, P., et al. (2002). "The PERK eukaryotic initiation factor 2 alpha kinase is required for the development of the skeletal system, postnatal growth, and the function and viability of the pancreas." Mol Cell Biol **22**(11): 3864-3874.

Zhou, G., et al. (2001). "Role of AMP-activated protein kinase in mechanism of metformin action." The Journal of clinical investigation **108**(8): 1167-1174.

Zhu, X. L., et al. (1989). "Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process." Nature **339**(6224): 483-484.

Zimmet, P. (1982). "Type 2 (non-insulin-dependent) diabetes--an epidemiological overview." Diabetologia **22**(6): 399-411.

Zinszner, H., et al. (1998). "CHOP is implicated in programmed cell death in response to impaired function of the endoplasmic reticulum." Genes Dev **12**(7): 982-995.

Zuk, O., et al. (2014). "Searching for missing heritability: Designing rare variant association studies." Proceedings of the National Academy of Sciences **111**(4): E455.

Appendices

ARTICLES
https://doi.org/10.1038/s41588-018-0173-1
nature genetics

Type 2 diabetes risk alleles in *PAM* impact insulin release from human pancreatic β-cells

Soren K. Thomsen[1,7,9], Anne Raimondo[1,8,9], Benoit Hastoy[1,9], Shahana Sengupta[1,2], Xiao-Qing Dai[3], Austin Bautista[3], Jenny Censin[4,5], Anthony J. Payne[5], Mahesh M. Umapathysivam[1], Aliya F. Spigelman[3], Amy Barrett[1], Christopher J. Groves[1], Nicola L. Beer[1], Jocelyn E. Manning Fox[3], Mark I. McCarthy[1,5,6], Anne Clark[1], Anubha Mahajan[5], Patrik Rorsman[1,6], Patrick E. MacDonald[3] and Anna L. Gloyn[1,5,6*]

The molecular mechanisms underpinning susceptibility loci for type 2 diabetes (T2D) remain poorly understood. Coding variants in peptidylglycine α-amidating monooxygenase (*PAM*) are associated with both T2D risk and insulinogenic index. Here, we demonstrate that the T2D risk alleles impact negatively on overall PAM activity via defects in expression and catalytic function. PAM deficiency results in reduced insulin content and altered dynamics of insulin secretion in a human β-cell model and primary islets from cadaveric donors. Thus, our results demonstrate a role for PAM in β-cell function, and establish molecular mechanisms for T2D risk alleles at this locus.

Human genetics has the potential to identify novel mechanisms for disease, which in turn present opportunities for clinical translation. Over the past 10 years, genome-wide association studies (GWAS) have identified over 400 genomic regions robustly associated with type 2 diabetes (T2D) risk, and have highlighted a central role for pancreatic β-cell dysfunction in T2D pathogenesis[1–4]. However, translating GWAS signals into molecular mechanisms for disease has been slow, primarily because of uncertainty over the transcripts through which non-coding association signals operate[5]. Recently, studies focused on identifying coding variant association signals have identified non-synonymous alleles in peptidylglycine α-amidating monooxygenase (*PAM*) independently that are associated with a risk of T2D (rs78408340, NM_000919.3, c.1616C > G, odds ratio (OR) = 1.47, minor allele frequency (MAF) = 7.3×10^{-3} in Europeans; rs35658696, NM_000919.3, c.1688A > G, OR = 1.23, MAF = 0.045), providing a direct link to the effector transcript[3,6,7]. Although the low-frequency variant at rs35658696 was in linkage disequilibrium with a coding variant in a nearby gene (*PPIP5K2*; rs36046591), recent fine-mapping efforts have confirmed rs35658696 as the causal variant[4]. Both T2D risk alleles are also associated with a reduced insulinogenic index (rs78408340, β = −8.42; rs35658696, β = −1.96)—a measure of glucose-stimulated insulin secretion—suggesting that their effects are mediated via altered β-cell function[6,8,9].

PAM encodes PAM, an enzyme in neuroendocrine cells that modifies peptides with a C-terminal glycine residue to create peptide amides[10–12]. Amidation can dramatically increase the biological potency of a peptide relative to its unmodified glycine-extended conjugate[13]. PAM is localized to the Golgi complex, where it is packaged with other endocrine proteins into nascent granules[10,14]. The functional enzyme exists in both integral membrane and luminal forms; the latter is co-secreted with the endocrine peptide(s)[10,11,15].

Despite reports of PAM expression in pancreatic islets, a functional role in β-cells has not yet been described[16]. Insulin itself is not a PAM substrate, so the association of *PAM* variants with an insulinogenic index must be mediated via other peptide(s) or mechanisms. We hypothesized that T2D-associated *PAM* missense alleles reduce PAM function, affecting the amidation of peptides that is critical for insulin secretion. We demonstrate that both diabetes risk alleles negatively impact on PAM expression and/or activity, and elucidate an endogenous role for PAM in insulin granule packaging and release from β-cells. We also show that PAM amidates the granule packaging factor chromogranin A (CgA), and identify this neuroendocrine peptide as a likely downstream effector for PAM in β-cells. Our results are consistent with the direction and magnitude of effects for T2D-associated risk alleles in *PAM*, and establish molecular mechanisms for their impact on disease susceptibility.

Results

T2D-associated *PAM* alleles cause PAM loss-of-function. PAM is a bifunctional enzyme, possessing two contiguous catalytic domains: peptidylglycine α-hydroxylating monooxygenase (PHM) and peptidyl α-hydroxyglycine alpha-amidating lyase (PAL)[11]. Both T2D risk variants in *PAM* encode mutations located within the PAL domain (rs78408340, p.Ser539Trp; rs35658696, p.Asp563Gly) and are predicted by in silico tools (SIFT (Sorting Intolerant From Tolerant), PolyPhen2 (Polymorphism Phenotyping v2)) to be damaging, suggesting that they could affect enzymatic activity. To test this, we produced recombinant luminal (non-integral

[1]Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, Oxford, UK. [2]MRC Harwell Institute, Harwell Campus, Oxfordshire, UK. [3]Department of Pharmacology and Alberta Diabetes Institute, University of Alberta, Edmonton, Alberta, Canada. [4]Big Data Institute at the Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK. [5]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK. [6]Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK. [7]Present address: Vertex Pharmaceuticals Europe Ltd, Milton Park, Abingdon, UK. [8]Present address: National Health and Medical Research Council, Canberra, Australia. [9]These authors contributed equally: Soren K. Thomsen, Anne Raimondo and Benoit Hastoy. *e-mail: anna.gloyn@drl.ox.ac.uk

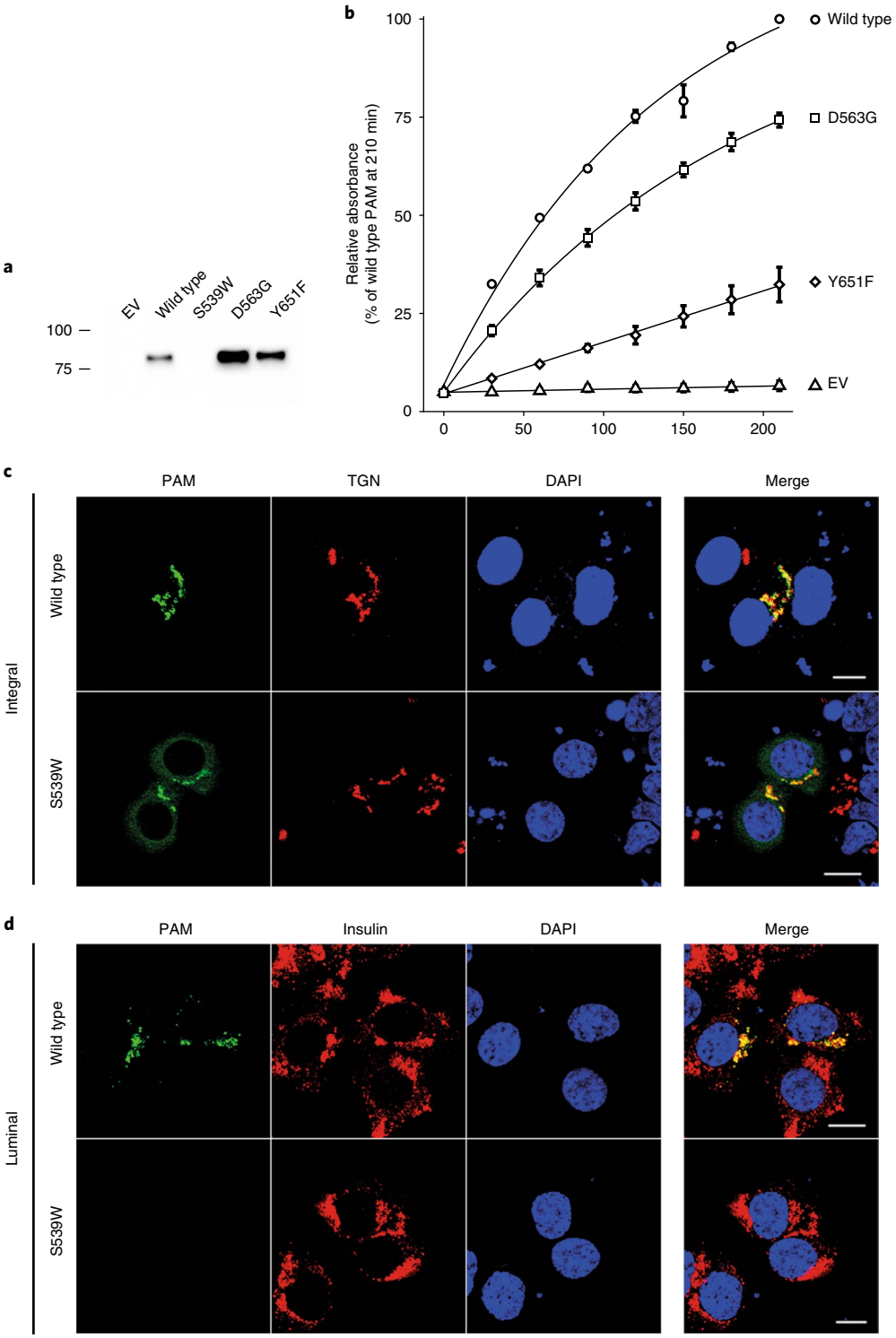

Fig. 1 | Analysis of wild-type and variant PAM function and expression. a, Western blot analysis of recombinant PAM protein production in supernatant from HEK293 stable cells. Size markers indicate protein mass in kilodaltons. **b**, Amidating activity of wild-type PAM (circles), p.Asp563Gly PAM (squares), p.Tyr651Phe PAM (diamonds) or empty vector (EV) (triangles) in vitro. The graph shows the means of four independent experiments; the error bars are the standard error of the mean. **c,d**, EndoC-βH1 cells transfected with expression vectors for integral membrane (**c**) or luminal (**d**) wild-type or variant PAM, and labeled for PAM (green), the trans-Golgi network (TGN) (red in **c**) or insulin (red in **d**). DAPI (blue) was used as a nuclear marker. Scale bar, 2 µm. Results in **a**,**c** and **d** are representative of three independent experiments.

membrane) PAM protein for in vitro amidation assays using human embryonic kidney 293 (HEK293) cells as a suitable human expression system. In line with previous observations, PAM was constitutively released into the supernatant (Fig. 1a)[17]. Wild-type PAM and p.Asp563Gly PAM were both robustly produced. Additionally, a catalytically inactive mutant protein, p.Tyr651Phe PAM, was also produced, which was used as a control[18]. Interestingly, we could not detect p.Ser539Trp PAM expression

(Fig. 1a). This was observed across three independently derived cell lines, and was not due to cellular retention of p.Ser539Trp PAM (data not shown).

We subsequently developed a cell-free kinetic assay capable of measuring PAM amidating activity via spectrophotometric detection of converted glyoxylate, a by-product of the amidation reaction[19,20]. Matching each reaction for PAM input, we observed reduced amidating activity for p.Asp563Gly (D563G) PAM ($P=1.0\times10^{-5}$) and p.Tyr651Phe (Y651F) PAM ($P=4.1\times10^{-6}$) (Fig. 1b). In agreement with its lack of expression, supernatant from p.Ser539Trp (S539W) PAM-transfected cells was inactive in this assay (Supplementary Fig. 1a). Further analysis showed no significant difference in substrate affinity between wild-type PAM and p.Asp563Gly PAM (substrate concentration which gives half maximal velocity of the enzymatic reaction [K_m] 0.95 mmol.l^{-1} versus 1.02 mmol.l^{-1}, $P=0.44$), suggesting that the p.Asp563Gly substitution affects the number of substrate molecules converted over time (K_{cat})(Supplementary Fig. 1b). These results demonstrate that the T2D-associated missense alleles in *PAM* decrease PAM function via a combination of defective expression and/or reduced catalysis.

PAM localizes to the β-cell secretory pathway. Having established the direction of effect for T2D risk alleles in *PAM*, we next explored a role for PAM in physiologically relevant tissues. Transcript expression profiling detected *PAM* in multiple tissue types, with the highest expression in human islets followed by the heart and salivary glands (Supplementary Fig. 2a). In mouse tissues, *Pam* expression was highest in the pituitary gland (not included in the human panel) followed by the heart and islets (Supplementary Fig. 2a). These results are consistent with the published association between *PAM* risk alleles and reduced measures of insulin secretion, and suggest a direct role for PAM in β-cell function. We confirmed broadly similar expression patterns in the publicly available Genotype Tissue Expression (GTEx) database, although data for pancreatic islets are not available from this project[21]. Published RNA sequencing data also showed *PAM* to be equally expressed in enriched fractions of β-cells and non-β-cells (Supplementary Fig. 2a)[22,23]. We verified this expression pattern at the protein level by performing immunofluorescence and electron microscopy of primary human islets, which showed co-localization of endogenous PAM in both insulin and glucagon vesicles (Fig. 2a,b). We also confirmed PAM localization to the insulin secretory pathway of the glucose-responsive human β-cell model EndoC-βH1 (Fig. 2c)[24,25]. These observations are in agreement with early studies that established PAM expression and activity in endocrine cells of human and murine islets[26–29].

We then performed detailed analysis of the subcellular localization of wild-type and variant PAM for both integral and luminal forms of the protein in EndoC-βH1 cells. Integral membrane wild-type PAM localized to the trans-Golgi network (TGN), where secretory granules originate, whereas luminal wild-type PAM displayed a punctate staining pattern and co-localized with endogenous insulin (Fig. 1c,d). Similar results were observed for p.Asp563Gly PAM and p.Tyr651Phe PAM (Supplementary Fig. 2b). As in HEK293 cells, luminal p.Ser539Trp PAM expression was undetectable (Fig. 1d); however, integral membrane p.Ser539Trp PAM displayed an aberrant staining pattern that localized to the endoplasmic reticulum as well as the TGN (Fig. 1c and Supplementary Fig. 2b). These results confirm that PAM localizes to components of the β-cell secretory pathway, in agreement with studies in other neuroendocrine cell types[10,14,30–32]. The p.Ser539Trp substitution may interfere with PAM protein folding, so that its translocation from the endoplasmic reticulum to the TGN is prevented or delayed.

Impact of risk variants on PAM expression. Given that luminal isoforms of PAM are known to be co-secreted with neuroendocrine peptides, we next wondered whether the impact of the p.Ser539Trp substitution would be reflected in lower circulating PAM levels in the bloodstream. Analyzing publicly available data from a human plasma proteome study by Sun et al.[33], we confirmed a significant decrease in the concentration of PAM for rs78408340 risk allele carriers (β = −1.53 per risk allele, 95% confidence interval (CI) (−1.84, −1.22), $P=5.6\times10^{-22}$, in standard deviation units of PAM levels), consistent with our in vitro observations in both HEK293 and EndoC-βH1 cells.

Interestingly, we also found that the T2D risk allele at rs35658696 was associated with reduced PAM protein levels in plasma (β = −0.97 per risk allele, 95% CI (−1.09, −0.85), $P=1.5\times10^{-67}$). To determine whether this effect could be mediated via regulatory effects on *PAM* gene expression, we analyzed RNA sequencing and genotype data from the largest available cohort of islets, which consisted of a combined collection of islets from Edmonton, Canada and Oxford, UK[34]. Out of a total of 118 donors, 11 individuals were identified as heterozygous for rs35658696. After correcting for potential mapping biases using WASP[35], allele-specific expression revealed a significant imbalance ($P=1.76\times10^{-4}$). The overall difference was relatively modest (53.2% versus 46.8%), but the imbalance was biased toward higher expression of the reference allele (Supplementary Fig. 3). This is directionally consistent with the negative functional impact of the minor allele on PAM amidation, highlighting a potential exacerbation of the reduction in PAM activity in risk allele carriers.

Finally, to identify a possible mechanism underlying this apparent difference in regulatory activity at rs35658696, we searched for genomic annotations overlapping the risk variant in reported

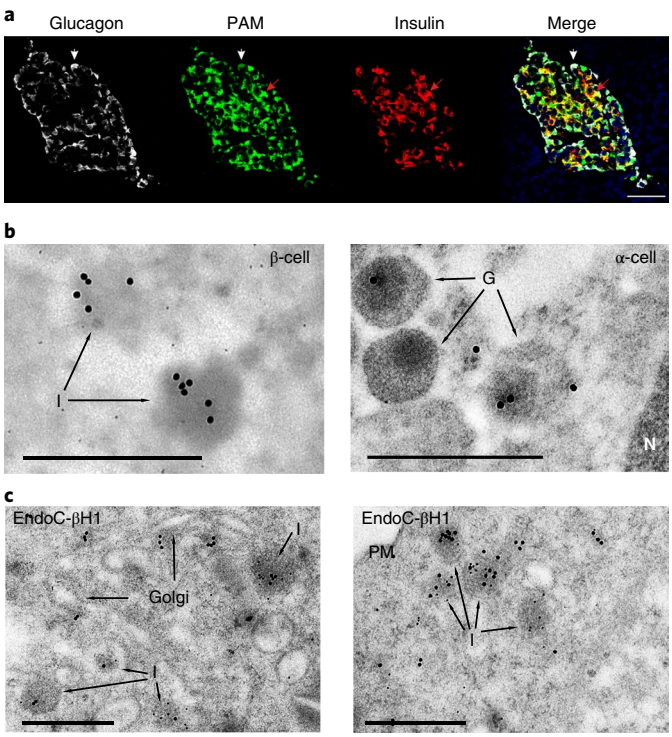

Fig. 2 | PAM expression in human pancreatic islets and EndoC-βH1 cells. **a**, Immunolocalization of endogenous PAM in a histological section of human pancreas; glucagon (white), PAM (green) and insulin (red). Red and white arrows indicate co-localization of PAM with insulin and glucagon, respectively; scale bar, 50 µm. **b**, Electron micrograph of PAM localization in human β- and α-cells. PAM immunogold-labeled (15 nm gold) was localized to insulin vesicles (I) and to glucagon vesicles (G). Scale bar, 500 nm. **c**, PAM localization in EndoC-βH1 cells. Insulin (10 nm), PAM (15 nm). PM, plasma membrane. N, nucleus. Scale bar, 500 nm.

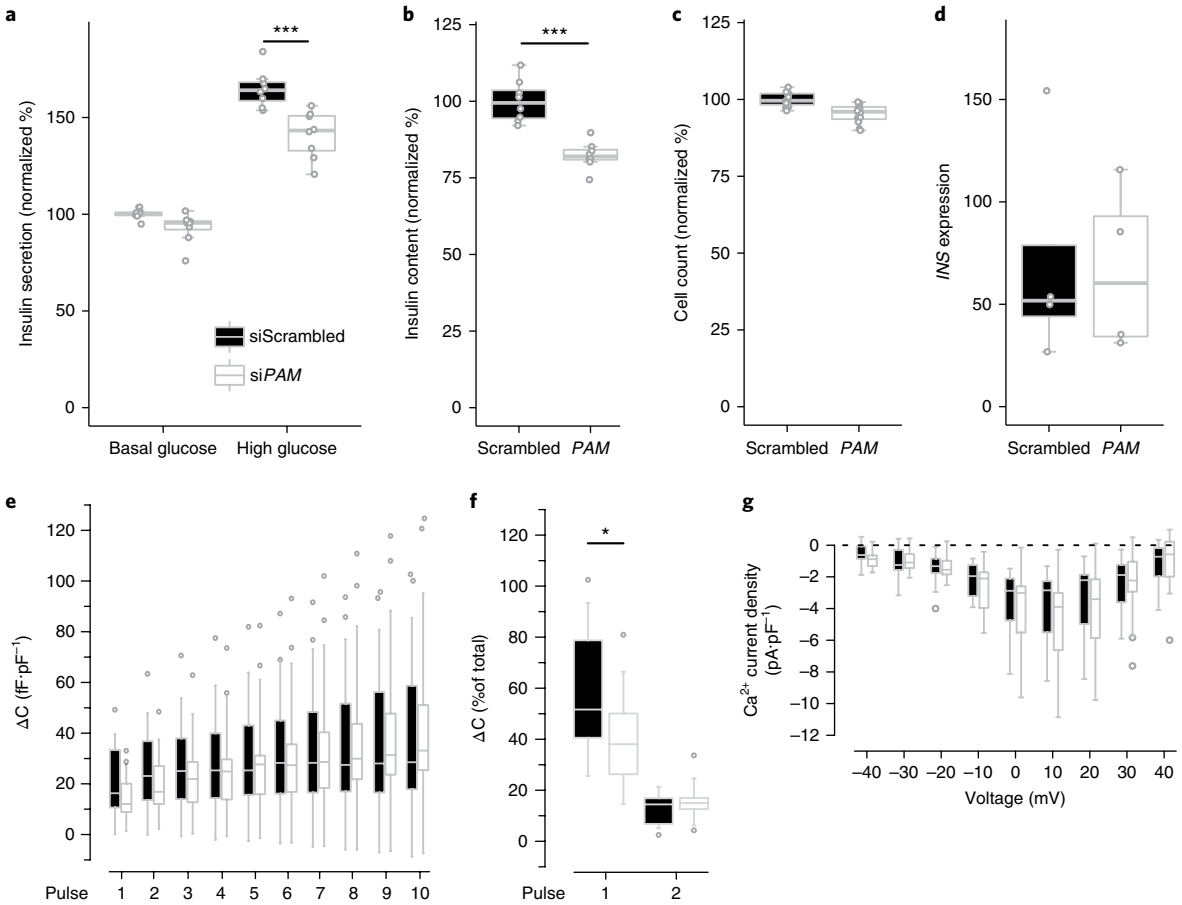

Fig. 3 | Effects of endogenous PAM on β-cell function. a–g, EndoC-βH1 cells were transfected with scrambled or *PAM* siRNA and then measured for insulin secretion ($n=8$ biologically independent samples) (**a**), cellular insulin content ($n=8$) (**b**), cell numbers ($n=16$) (**c**), insulin gene (*INS*) expression ($n=4$ independent experiments) (**d**), granule exocytosis measured as depolarization-evoked cumulative increase in membrane capacitance ($n=15$ scramble or $n=16$ *PAM* siRNA transfected cells) (**e**,**f**), and calcium current density ($n=15$ cells per treatment) (**g**). *P* values * < 0.05, ** < 0.01 and *** < 0.001 for two-tailed Student's *t*-tests where a single independent variable was being tested (**b–d**), or for two-way analysis of covariance (ANCOVA) followed by Tukey's HSD post-hoc test where two experimental variables were being tested (**a**,**e–g**). All box plots display the interquartile range (IQR) with medians, whiskers indicating 1.5 × the IQR, and outliers (for $n>10$) or all data points (for $n<10$).

chromatin states for human pancreatic islets[36]. This was extended to consider all variants in linkage disequilibrium ($r^2>0.8$) with the lead SNP. The predicted causal variant (rs35658696) overlapped annotations indicative of gene enhancer activity; two variants in strong linkage disequilibrium ($r^2>0.97$) with the causal variant overlapped weak enhancer annotations (Supplementary Table 1). These observations provide a framework for understanding the effect of the T2D risk allele on *PAM* transcript levels, although a more detailed mechanistic interrogation is outside the scope of this study.

Taken together, these results confirm the association of the rs78408340 (p.Ser539Trp) risk allele with reduced PAM protein levels and point to a regulatory component of the causal mechanism underlying the rs35658696 (p.Asp563Gly) risk allele. The latter effect may be relatively smaller than the functional impact on catalytic activity. Nevertheless, the effects are directionally aligned, thus compounding the negative effect of the risk allele through a dual mechanism of deficiencies in both catalytic function and expression.

PAM regulates glucose-responsive insulin secretion. In light of the negative impact of both T2D risk alleles on overall PAM function, we next determined the effects of reduced PAM levels on β-cell activity. Small interfering RNA (siRNA)-mediated knockdown of *PAM* in EndoC-βH1 cells caused a 79% reduction in *PAM* transcript levels (Supplementary Fig. 4a). This was associated with significantly decreased insulin secretion in high glucose (−14.4%; $P=5.6\times10^{-5}$), but not in basal glucose (−6.9%; $P=0.39$) (Fig. 3a). Furthermore, *PAM* knockdown cells displayed a significantly blunted response to glucose (ratio of high to basal secretion, −7.7%; $P=0.03$), which is indicative of a greater impact on secretion under stimulated conditions. Cellular insulin content was also reduced in *PAM* knockdown cells (−17.6%; $P=4.9\times10^{-5}$) despite normal viability and insulin (*INS*) gene expression (Fig. 3b–d). As a result, secretion measures could be normalized by adjusting for insulin content (that is, expressing secretion as a percentage of content), with no significant defects for small interfering *PAM* (si*PAM*) cells (Supplementary Fig. 5a–d). We also noted that content-adjusted rates of insulin secretion were not further impacted by stimulation with either depolarizing KCl or the 3′, 5′-cyclic adenosine monophosphate (cAMP)-elevating agent forskolin (Supplementary Fig. 6a). This indicates that the reductions in insulin secretion following *PAM* knockdown are not caused by a defect in calcium entry or total granule exocytosis.

We next performed quantitative electron microscopy using immunogold labeling of PAM and insulin. The results of this analysis confirmed successful knockdown of PAM at the protein level, with a 47% reduction in the number of particles per vesicle (ppv)

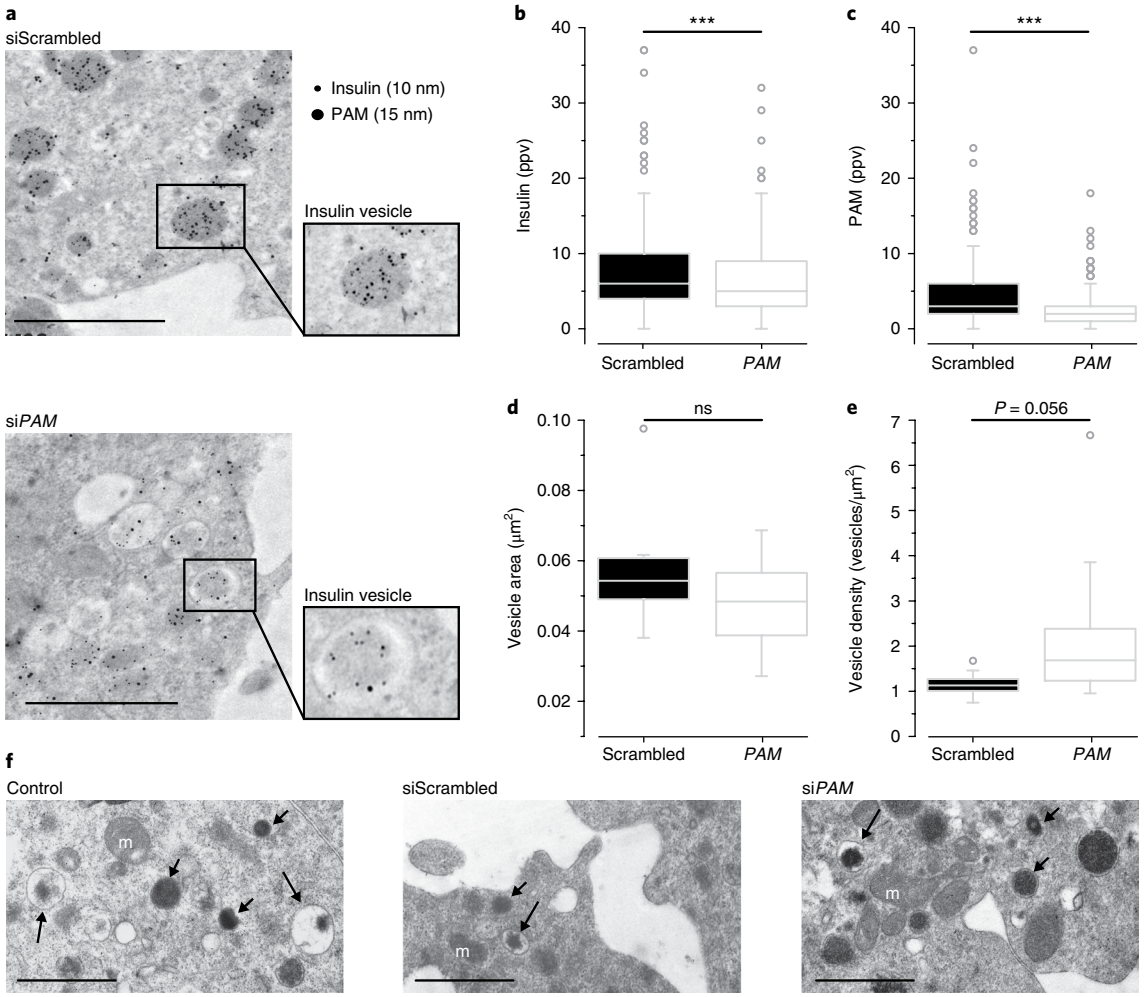

Fig. 4 | Analysis of β-cell ultrastructural features following PAM silencing. a, Electron micrographs from scrambled (siScrambled, upper panel) and *PAM* siRNA-transfected (si*PAM*, lower panel) EndoC-βH1 cells immunogold-labeled for insulin (10 nm) and PAM (15 nm). Scale bar, 2 μm. **b**, Number of immunogold insulin particles per vesicle (ppv). **c**, Number of immunogold PAM ppv. **d**, Vesicle area. **e**, Vesicle density (vesicle per μm^2). Results shown are the average of $n=235$ (scrambled) or $n=290$ (*PAM*) vesicles, except for **e**, where $n=10$ cells per treatment. **f**, Electron micrographs from EndoC-βH1 cells either non-transfected or siRNA-transfected, as indicated. Immature and mature insulin vesicles are highlighted by short and long arrows respectively. m, mitochondrion. *P* values * < 0.05, ** < 0.01, *** < 0.001 for two-tailed Student's *t*-tests. All box plots display the interquartile range (IQR) with medians, whiskers indicating 1.5 × the IQR, and outliers.

(Fig. 4a,c). si*PAM*-treated cells showed a 23% decrease in immunogold-labeled insulin per vesicle compared with control cells (6.26 ± 0.28 versus 8.1 ± 0.41 ppv; $P=1.77\times10^{-4}$) (Fig. 4a,b), thus corroborating the pooled content measures described earlier. There were no significant changes in cross-sectional vesicle area ($0.048\pm0.003\,\mu m^2$ versus $0.056\pm0.005\,\mu m^2$; $P=0.18$) or vesicle density (2.30 ± 0.55 vesicles/μm^2 versus 1.16 ± 0.09 vesicles/μm^2; $P=0.056$) in *PAM* knockdown cells (Fig. 4d,e). No qualitative differences in insulin crystallization were observed (Fig. 4f).

To probe further the effect of *PAM* silencing on insulin secretion, we performed high-resolution, single-cell capacitance measurements. Following stimulation of individual cells with a train of ten depolarizations, an increase in plasma membrane surface area due to vesicle exocytosis could be detected at each step. The cumulative increase in exocytosis was comparable between *PAM* knockdown and control cells (43.9 ± 8.6 fF pF^{-1} versus 38.0 ± 8.1 fF pF^{-1}; $P=0.64$) (Fig. 3e), and there was no difference in calcium entry (peak current 4.5 ± 0.7 pA pF^{-1} versus 3.6 ± 0.6 pA pF^{-1} for *PAM* knockdown versus control cells; $P=0.27$) (Fig. 3g). Moreover, the exocytotic response increased similarly with the calcium current amplitude in both control and *PAM* knockdown cells (Supplementary Fig. 6b), corroborating the normal calcium sensitivity inferred from secretion assays.

In contrast, the kinetics of vesicle release were significantly different; in controls cells, nearly 60% of the exocytotic response was elicited by the first pulse, compared with <40% in *PAM* knockdown cells ($P=0.013$) (Fig. 3f). The increase in cell capacitance during the first two depolarizations provides an estimate for the size of the readily releasable pool (RRP) of secretory granules, whereas subsequent pulses represent mobilization of granules from the more intracellular reserve pool[37]. Thus, our results indicate that PAM is an endogenous regulator of the exocytotic response in pancreatic β-cells, with a specific effect on the availability of the RRP.

Effects of the rs35658696 risk variant on primary islet function. Having established a role for PAM in a human β-cell model, we sought to corroborate our results by examining the effects of the low-frequency *PAM* variant (rs35658696) in primary islets. First, we analyzed insulin secretion and content measures for intact human islets isolated from cadaveric donors. In the Edmonton collection of functionally characterized islets, we identified 16 individuals heterozygous for the T2D risk allele ("carriers") and

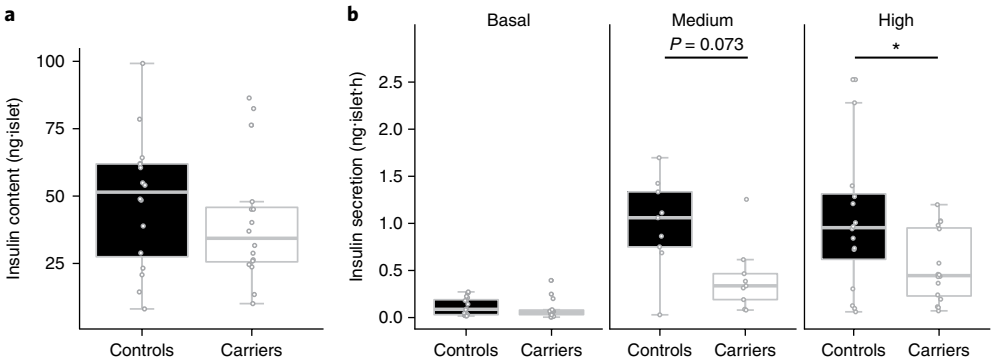

Fig. 5 | Effects of rs35658696 genotype status on insulin secretion measures from intact human islets. a,b, Primary islets isolated from individuals heterozygous for the rs35658696 risk variant (Carriers) and matched individuals homozygous for the reference allele (Controls) were analyzed for total insulin content (**a**; $n=16$ donors in each group) and insulin secretion (**b**) in basal (1 mM; $n=16$), medium (10 mM; $n=9$) and high glucose (16.7 mM; $n=16$). *P* values * < 0.05, ** < 0.01, *** < 0.001 for two-way analysis of covariance (ANCOVA) followed by Tukey's HSD post-hoc test. All box plots display the interquartile range (IQR) with medians, whiskers indicating 1.5 × the IQR, and individual data points.

16 age-, sex-, and body mass index (BMI)-matched donors homozygous for the reference allele ("controls"). All individuals were analyzed for insulin content, as well as secretion under basal (1 mM) and high (16.7 mM) glucose in sequential incubations. In a subset of nine individuals, insulin secretion was also assessed for an intermediate glucose level (10 mM).

We detected no significant change in insulin content (−16.3%; $P=0.37$), but we detected a highly significant reduction in overall insulin secretion for risk carriers compared with controls ($P=7.81\times10^{-4}$ by analysis of covariance (ANCOVA)) (Fig. 5). The effect on secretion showed a positive interaction with glucose levels ($P=0.041$), highlighting a greater difference between risk carriers and controls under stimulated conditions. Thus, testing for pairwise significant differences, we identified no effect of carrier status in basal glucose (−23.8%; $P=1.00$), but a trend toward a reduction in intermediate glucose (−58%; $P=0.073$) and significantly decreased secretion in high glucose (−49%; $P=0.016$). Importantly, these defects in raw insulin secretion could be fully rescued by normalizing to insulin content ($P=0.88$ by ANCOVA), with no significant pairwise differences ($P=1.00$ for basal; $P=1.00$ for intermediate; and $P=0.99$ for high glucose) (Supplementary Fig. 7). Overall, our results establish a defect in intact islets from risk allele carriers highly concordant with that observed following *PAM* knockdown in the EndoC-βH1 model.

Next, we analyzed exocytosis in single β-cells from dispersed human islets, with capacitance measurements following a train of ten depolarizations in either 1 or 10 mM glucose. Seven individuals heterozygous for the T2D risk allele were compared to seven matched controls (Fig. 6a). As with the insulin secretion measures from intact islets, we observed a statistically significant interaction between glucose level and genotype status ($P=2.83\times10^{-4}$ by ANCOVA), with a greater negative effect of the risk allele under stimulated conditions ($P=1.22\times10^{-3}$) than in basal glucose ($P=0.48$). To specifically test for differences in the RRP, we calculated the total increase in capacitance during the first two depolarizations (Fig. 6b). Although no differences were detected in low glucose, a significant decrease in capacitance was seen in risk allele carriers at 10 mM glucose (−26.5%; $P=0.04$), which is indicative of reduced exocytosis from the RRP. Our results are therefore consistent with an effect of the rs35658696 genotype on the kinetics of granule release that manifests predominantly under stimulated conditions.

Chromogranin A is a substrate and plausible downstream effector of PAM in β-cells. To elucidate the PAM-regulated pathway in β-cells, we finally sought to identify possible target proteins that could act as downstream effectors of PAM activity. While any C-terminal glycine-extended peptide could be a substrate for PAM, we reasoned that a key downstream mediator of the observed effects in β-cells would need to be co-expressed with PAM along the secretory pathway. Therefore, we searched a published catalog of the islet secretome for potential targets of PAM amidation[38]. This highlighted two proteins, chromogranin A (CgA, encoded by *CHGA*) and islet amyloid polypeptide (IAPP, encoded by *IAPP*), both of which are highly expressed in β-cells and have plausible links with PAM function. CgA is a granular packaging protein that self-aggregates to condense large numbers of neuroendocrine peptides into nascent secretory granules[39–41]. It can also be cleaved to produce smaller peptides with additional biological roles[42]. The role of IAPP in β-cells is less well understood, although it is known to be stored in secretory granules and co-released with insulin[43].

CHGA silencing (Supplementary Fig. 4b) resulted in a phenotype similar to that observed for *PAM*, with significant reductions in both insulin content and secretion (Fig. 7a–c and Supplementary Fig. 5e–h). Using an antibody specific to the glycine-extended (non-amidated) form of CgA (CgA-Gly), we also observed an accumulation of CgA-Gly in EndoC-βH1 cells treated with *PAM* siRNA by western blot analysis (Fig. 7d,e and Supplementary Fig. 8 for full blots). Treatment of cells with 4-phenyl-3-butenoic acid (4P3BA), an irreversible inhibitor of PAM catalytic activity, produced similar results, showing that this effect is not due to off-target effects. In both cases, the ratio of non-amidated CgA to total CgA was significantly increased following perturbation of PAM (+226%, $P=0.016$ for *PAM* knockdown; +282%, $P=0.043$ for PAM inhibition) (Fig. 7d)[44,45]. Thus, CgA is a plausible PAM substrate in β-cells and a strong candidate for mediating the effects of PAM on β-cell function. In contrast, silencing of *IAPP* in EndoC-βH1 cells had no effect on insulin content or secretion (Supplementary Fig. 4c–f).

Finally, we attempted to detect patterns of CgA amidation in primary human islets by immunostaining pancreatic sections from a subset of six risk allele carriers and six matched controls (a subset of the donors analyzed by electrophysiology). By co-staining for insulin and either CgA-Gly or total CgA, we successfully identified CgA in the endocrine component of the pancreas (Supplementary Fig. 9a). Using a mask to focus exclusively on insulin-positive β-cells, we then quantified average intensities and derived a ratio of non-amidated to total CgA. The results were highly variable, with more than a 15-fold discrepancy between the lowest (6.89%) and highest (109%) ratios. Thus, no overall difference was detected between carriers and controls ($P=0.90$) (Supplementary Fig. 9b); larger studies may be required to detect any amidation defects in risk carriers.

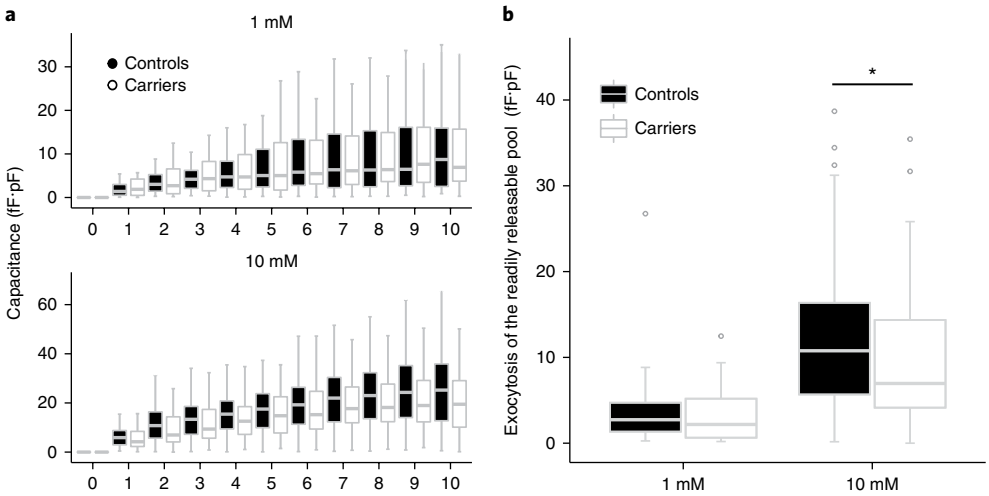

Fig. 6 | Effects of rs35658696 genotype status on exocytosis measurements in dispersed human islets. Capacitance measurements for primary β-cells from seven individuals heterozygous for the rs35658696 risk variant (Carriers; $n = 45$ and 41 cells analyzed in 1 and 10 mM glucose, respectively) and seven individuals homozygous for the reference allele (Controls; $n = 42$ and 40 cells). **a**, Cumulative capacitance measurements after each of ten depolarizations. **b**, Cumulative increase in capacitance for the first two depolarizations. *P* values * < 0.05, ** < 0.01 and *** < 0.001 for two-way analysis of covariance (ANCOVA) followed by Tukey's HSD post-hoc test. All box plots display the interquartile range (IQR) with medians; the whiskers indicate 1.5 × the IQR.

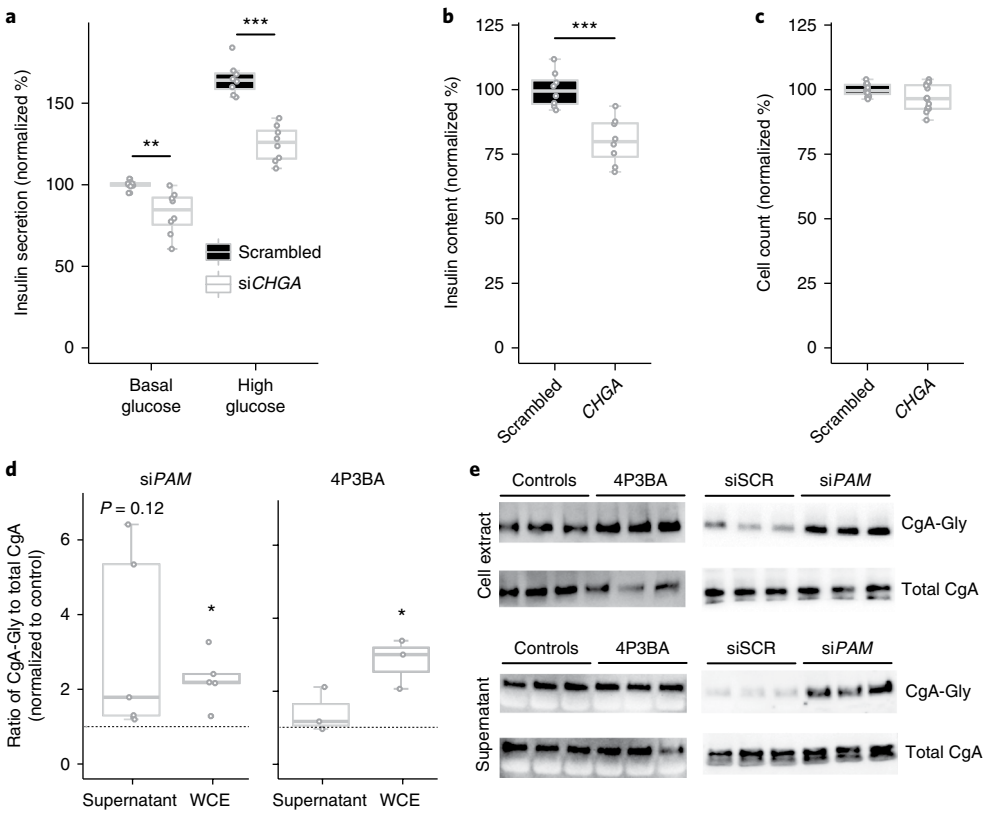

Fig. 7 | Effects of endogenous CgA on β-cell function. EndoC-βH1 cells were transfected with scrambled (siSCR) or *CHGA* siRNA then measured for **a** insulin secretion ($n = 8$ independent biological samples), **b** cellular insulin content ($n = 8$) and **c** cell numbers ($n = 16$). **d**–**e**, EndoC-βH1 cells were transfected with either scrambled or *PAM* siRNA ($n = 5$ independent experiments), or treated with either vehicle (Control) or 4P3BA ($n = 3$ independent experiments), then analyzed by western blot using antibodies specific for non-amidated CgA (CgA-Gly) and total CgA. Control levels are indicated in **d** by the dashed grey line. Results in **e** are representative of $n = 5$ independent experiments for *PAM* silencing, and $n = 3$ independent experiments for PAM inhibition. Corresponding unedited blots are provided in Supplementary Fig. 8. **b,c**, *P* values * <0.05, ** <0.01, *** <0.001 for paired, two-tailed Student's *t*-tests where a single independent variable was being tested. **a,d**, Two-way analysis of covariance (ANCOVA) followed by Tukey's HSD post-hoc test where two experimental variables were being tested. All box plots display the interquartile range (IQR) with medians, whiskers indicating 1.5 × the IQR and individual data points. WCE, whole-cell extract.

Discussion

Recent efforts to catalog coding variants associated with T2D risk have identified two independent association signals in the *PAM* gene[3,6,8]. The coding risk alleles are associated with a decreased insulinogenic index, a physiological measure of insulin secretion in response to a glucose challenge, indicating a likely endogenous role for PAM in β-cell function. Here, we provide evidence that both T2D risk alleles reduce overall PAM activity. While the risk allele at rs78408340 causes defective PAM expression, our results demonstrate that the risk allele at rs35658696 acts through a dual mechanism of reduced catalytic function and, albeit to a lesser extent, decreased expression.

We have further shown that reduced PAM activity causes β-cell dysfunction through multiple effects on insulin availability and the dynamics of granule exocytosis. Importantly, these defects preferentially affect insulin secretion under stimulated conditions, leading to a blunted glucose response. Taken together, our findings support a molecular mechanism whereby T2D risk alleles in *PAM* decrease PAM function, directly impacting on the ability of pancreatic β-cells to mobilize insulin in response to glucose.

This novel mechanism provides an explanation for the relatively larger effect size of rs78408340 (p.Ser539Trp) compared with rs35658696 (p.Asp563Gly), as reported in two independent GWAS for T2D risk and insulinogenic index[6,8]. Mislocalization or misexpression of p.Ser539Trp PAM would presumably compromise function more severely than the combined reduction in amidation activity and expression observed for p.Asp563Gly PAM. This is particularly true for cell types in which the luminal form predominates over the integral form of PAM, since misexpressed p.Ser539Trp PAM may retain some residual function. The precise cellular dynamics that yield luminal PAM in β-cells have not been elucidated; however, it appears generally to be more predominant in neuroendocrine cell types[17]. This raises the intriguing possibility of tissue-specific differences in PAM activity for rs78408340 risk allele carriers.

To explore the molecular mechanism underlying β-cell dysfunction in *PAM* risk allele carriers, we performed detailed characterizations in primary islets and the human β-cell line EndoC-βH1. The observed defects were broadly similar, with significant reductions in insulin secretion driven by decreased insulin content. Further, the size of the RRP was diminished in both primary β and EndoC-βH1 cells. However, total exocytosis remained unaffected, which is consistent with the lack of effect on content-normalized insulin secretion. Taken together, these results highlight a defect in insulin mobilization mediated through granule biogenesis and trafficking.

Interestingly, the phenotype caused by reduced PAM activity was consistently more pronounced under stimulated conditions. In EndoC-βH1 cells, insulin secretion was not affected in low glucose and the stimulation index (ratio of high to basal secretion) was significantly reduced. In primary β-cells, defects in exocytosis were apparent only in 10 mM glucose, with normal levels in basal glucose. Moreover, insulin secretion from intact islets was significantly reduced under stimulated conditions only and with a significant interaction between genotype status and glucose levels. *PAM* loss-of-function in β-cells thus manifests preferentially under conditions resembling the fed state in vivo, in line with the association of the T2D risk alleles with a reduced insulinogenic index.

Our molecular studies in EndoC-βH1 cells also confirmed that CgA is a substrate for PAM amidation in β-cells[45]. CgA is a granular packaging protein known to influence vesicle composition in other neuroendocrine cell types by facilitating the condensation of secreted proteins within nascent secretory vesicles[39–41,46–48]. Previous studies using global *Chga* knockout mouse models have attempted to investigate a role for CgA in pancreatic islets. However, they have been confounded by compensatory increases in other granin proteins, including chromogranin B[49,50]. In mice, knockdown of chromogranin B leads to strong inhibition of insulin secretion by mechanisms that have not been resolved[51]. Our results suggest a model whereby amidation of CgA by PAM enhances its ability to aggregate and package insulin efficiently in granules. In support of this hypothesis, amidation of numerous biological peptides has been shown to promote protein structure and self-aggregation due to improved electrostatic interactions[52–54]. C-terminal cleavage of CgA also increases protein stability in pituitary cells[55,56]. More detailed studies in primary islets is required to confirm any defective amidation of CgA in risk allele carriers.

Although our tissue expression analyses showed that *PAM* is highly expressed in pancreatic islets, the results also revealed *PAM* to be widely expressed across a range of other cell types. It is therefore possible that *PAM* could indirectly modulate β-cell function via amidated signals released from other tissues[23,57]. Glucagon-like peptide 1 (GLP-1) is an amidated peptide released from intestinal L cells that amplifies insulin secretion in response to a meal[58]. Moreover, GLP-1 (7–36) amide has an increased half-life compared to GLP-1 (7–37) in vitro, suggesting that amidation could lead to elevated GLP-1 levels in the circulation[59]. Therefore, PAM may have endocrine effects on β-cell function via decreased GLP-1 amidation and/or secretion, with possible implications for genotype-specific responses to incretin-based therapies.

Within the β-cell, it remains likely that additional PAM substrates may exist, some of which could further modulate insulin mobilization. We also cannot exclude the possibility of pleiotropic effects at the level of the T2D risk alleles; nevertheless, the presence of two independent coding variants in *PAM* strongly favor a primary causal mechanism mediated through *PAM*. This is further supported by our experimental findings, which establish direct impacts on PAM function and expression.

In conclusion, we have shown that the amidating enzyme PAM is a critical component of the regulated secretory pathway in β-cells, consistent with the observed effects of *PAM* missense variants on T2D risk and the insulinogenic index. Our results establish molecular mechanisms for T2D-associated *PAM* risk alleles and show multiple effects of reduced PAM function on insulin granule packaging and release. Our study illustrates the importance of coding variants implicated in disease risk for the identification of candidate causal genes. Further work is required to explore the possible contribution of other tissues to the effect of *PAM* risk alleles, particularly those that have implications for incretin-based therapies.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41588-018-0173-1.

Received: 21 September 2017; Accepted: 6 June 2018;
Published online: 27 July 2018

References

1. Dimas, A. S. et al. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes* **63**, 2158–2171 (2014).
2. Morris, A. P. et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981–990 (2012).
3. Fuchsberger, C. et al. The genetic architecture of type 2 diabetes. *Nature* **536**, 41–47 (2016).
4. Mahajan, A. et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nat. Genet.* **50**, 559–571 (2018).
5. Thomsen, S. K. & Gloyn, A. L. Human genetics as a model for target validation: finding new therapies for diabetes. *Diabetologia* **60**, 960–970 (2017).
6. Steinthorsdottir, V. et al. Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. *Nat. Genet.* **46**, 294–298 (2014).

7. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
8. Huyghe, J. R. et al. Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. *Nat. Genet.* **45**, 197–201 (2013).
9. Phillips, D. I., Clark, P. M., Hales, C. N. & Osmond, C. Understanding oral glucose tolerance: comparison of glucose or insulin measurements during the oral glucose tolerance test with specific measurements of insulin resistance and insulin secretion. *Diabet. Med.* **11**, 286–292 (1994).
10. Milgram, S. L., Mains, R. E. & Eipper, B. A. Identification of routing determinants in the cytosolic domain of a secretory granule-associated integral membrane protein. *J. Biol. Chem.* **271**, 17526–17535 (1996).
11. Oyarce, A. M. & Eipper, B. A. Neurosecretory vesicles contain soluble and membrane-associated monofunctional and bifunctional peptidylglycine alpha-amidating monooxygenase proteins. *J. Neurochem.* **60**, 1105–1114 (1993).
12. Eipper, B. A., Milgram, S. L., Husten, E. J., Yun, H. Y. & Mains, R. E. Peptidylglycine alpha-amidating monooxygenase: a multifunctional protein with catalytic, processing, and routing domains. *Protein Sci.* **2**, 489–497 (1993).
13. Merkler, D. J. C-terminal amidated peptides: production by the *in vitro* enzymatic amidation of glycine-extended peptides and the importance of the amide to bioactivity. *Enzym. Microb. Technol.* **16**, 450–456 (1994).
14. Milgram, S. L., Kho, S. T., Martin, G. V., Mains, R. E. & Eipper, B. A. Localization of integral membrane peptidylglycine alpha-amidating monooxygenase in neuroendocrine cells. *J. Cell Sci.* **110**, 695–706 (1997).
15. Eipper, B. A. et al. Alternative splicing and endoproteolytic processing generate tissue-specific forms of pituitary peptidylglycine alpha-amidating monooxygenase (PAM). *J. Biol. Chem.* **267**, 4008–4015 (1992).
16. Garmendia, O., Rodriguez, M. P., Burrell, M. A. & Villaro, A. C. Immunocytochemical finding of the amidating enzymes in mouse pancreatic A-, B-, and D-cells: a comparison with human and rat. *J. Histochem. Cytochem.* **50**, 1401–1416 (2002).
17. Tausk, F. A., Milgram, S. L., Mains, R. E. & Eipper, B. A. Expression of a peptide processing enzyme in cultured cells: truncation mutants reveal a routing domain. *Mol. Endocrinol.* **6**, 2185–2196 (1992).
18. De, M., Bell, J., Blackburn, N. J., Mains, R. E. & Eipper, B. A. Role for an essential tyrosine in peptide amidation. *J. Biol. Chem.* **281**, 20873–20882 (2006).
19. Maeda-Nakai, E. & Ichiyama, A. A spectrophotometric method for the determination of glycolate in urine and plasma with glycolate oxidase. *J. Biochem.* **127**, 279–287 (2000).
20. Carpenter, S. E. & Merkler, D. J. An enzyme-coupled assay for glyoxylic acid. *Anal. Biochem.* **323**, 242–246 (2003).
21. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
22. Nica, A. C. et al. Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. *Genome Res.* **23**, 1554–1562 (2013).
23. Blodgett, D. M. et al. Novel observations from next-generation RNA sequencing of highly purified human adult and fetal islet cell subsets. *Diabetes* **64**, 3172–3181 (2015).
24. Ravassard, P. et al. A genetically engineered human pancreatic β cell line exhibiting glucose-inducible insulin secretion. *J. Clin. Invest.* **121**, 3589–3597 (2011).
25. Tsonkova, V. G. et al. The EndoC-βH1 cell line is a valid model of human beta cells and applicable for screenings to identify novel drug target candidates. *Mol. Metab.* **8**, 144–157 (2018).
26. Scharfmann, R. et al. Persistence of peptidylglycine alpha-amidating monooxygenase activity and elevated thyrotropin-releasing hormone concentrations in fetal rat islets in culture. *Endocrinology* **123**, 1329–1334 (1988).
27. Zhou, A. & Thorn, N. A. Evidence for presence of peptide alpha-amidating activity in pancreatic islets from newborn rats. *Biochem. J.* **267**, 253–256 (1990).
28. Maltese, J. Y. et al. Ontogenetic expression of peptidyl-glycine alpha-amidating monooxygenase mRNA in the rat pancreas. *Biochem. Biophys. Res. Commun.* **158**, 244–250 (1989).
29. Martinez, A. et al. Immunocytochemical localization of peptidylglycine alpha-amidating monooxygenase enzymes (PAM) in human endocrine pancreas. *J. Histochem. Cytochem.* **41**, 375–380 (1993).
30. Gurgul-Convey, E., Kaminski, M. T. & Lenzen, S. Physiological characterization of the human EndoC-βH1 β-cell line. *Biochem. Biophys. Res. Commun.* **464**, 13–19 (2015).
31. El Meskini, R., Mains, R. E. & Eipper, B. A. Cell type-specific metabolism of peptidylglycine alpha-amidating monooxygenase in anterior pituitary. *Endocrinology* **141**, 3020–3034 (2000).
32. Andersson, L. E. et al. Characterization of stimulus-secretion coupling in the human pancreatic EndoC-βH1 beta cell line. *PLoS One* **10**, e0120879 (2015).
33. Sun, B.B. et al. Consequences of natural perturbations in the human plasma proteome. *bioRxiv* https://doi.org/10.1101/134551 (2017).
34. van de Bunt, M. et al. Transcript expression data from human islets links regulatory signals from genome-wide association studies for type 2 diabetes and glycemic traits to their downstream effectors. *PLoS Genet.* **11**, e1005694 (2015).
35. van de Geijn, B., McVicker, G., Gilad, Y. & Pritchard, J. K. WASP: allele-specific software for robust molecular quantitative trait locus discovery. *Nat. Methods* **12**, 1061–1063 (2015).
36. Thurner, M. et al. Integration of human pancreatic islet genomic data refines regulatory mechanisms at Type 2 Diabetes susceptibility loci. *Elife* **7**, e31977 (2018).
37. Gillis, K. D., Mossner, R. & Neher, E. Protein kinase C enhances exocytosis from chromaffin cells by increasing the size of the readily releasable pool of secretory granules. *Neuron* **16**, 1209–1220 (1996).
38. Yang, Y. H. et al. Paracrine signalling loops in adult human and mouse pancreatic islets: netrins modulate beta cell apoptosis signalling via dependence receptors. *Diabetologia* **54**, 828–842 (2011).
39. Kim, T., Tao-Cheng, J. H., Eiden, L. E. & Loh, Y. P. Chromogranin A, an "on/off" switch controlling dense-core secretory granule biogenesis. *Cell* **106**, 499–509 (2001).
40. Colomer, V., Kicska, G. A. & Rindler, M. J. Secretory granule content proteins and the luminal domains of granule membrane proteins aggregate *in vitro* at mildly acidic pH. *J. Biol. Chem.* **271**, 48–55 (1996).
41. Yoo, S. H. & Albanesi, J. P. Ca^{2+}-induced conformational change and aggregation of chromogranin A. *J. Biol. Chem.* **265**, 14414–14421 (1990).
42. Helle, K. B., Corti, A., Metz-Boutigue, M. H. & Tota, B. The endocrine role for chromogranin A: a prohormone for peptides with regulatory properties. *Cell. Mol. Life Sci.* **64**, 2863–2886 (2007).
43. Rorsman, P. & Braun, M. Regulation of insulin secretion in human pancreatic islets. *Annu. Rev. Physiol.* **75**, 155–179 (2013).
44. Katopodis, A. G. & May, S. W. Novel substrates and inhibitors of peptidylglycine alpha-amidating monooxygenase. *Biochemistry* **29**, 4541–4548 (1990).
45. Simpson, P. D. et al. Striking oxygen sensitivity of the peptidylglycine alpha-amidating monooxygenase (PAM) in neuroendocrine cells. *J. Biol. Chem.* **290**, 24891–24901 (2015).
46. Yoo, S. H. & Lewis, M. S. Dimerization and tetramerization properties of the C-terminal region of chromogranin A: a thermodynamic analysis. *Biochemistry* **32**, 8816–8822 (1993).
47. Mosley, C. A. et al. Biogenesis of the secretory granule: chromogranin A coiled-coil structure results in unusual physical properties and suggests a mechanism for granule core condensation. *Biochemistry* **46**, 10999–11012 (2007).
48. Bandyopadhyay, G. K. & Mahata, S. K. Chromogranin A regulation of obesity and peripheral insulin sensitivity. *Front. Endocrinol. (Lausanne)* **8**, 20 (2017).
49. Wollam, J. et al. Chromogranin A regulates vesicle storage and mitochondrial dynamics to influence insulin secretion. *Cell Tissue Res.* **368**, 487–501 (2017).
50. Bartolomucci, A. et al. The extended granin family: structure, function, and biomedical implications. *Endocr. Rev.* **32**, 755–797 (2011).
51. Obermüller, S. et al. Defective secretion of islet hormones in chromogranin-B deficient mice. *PLoS One* **5**, e8936 (2010).
52. Nanga, R. P., Brender, J. R., Vivekanandan, S. & Ramamoorthy, A. Structure and membrane orientation of IAPP in its natively amidated form at physiological pH in a membrane environment. *Biochim. Biophys. Acta* **1808**, 2337–2342 (2011).
53. Shalev, D. E., Mor, A. & Kustanovich, I. Structural consequences of carboxyamidation of dermaseptin S3. *Biochemistry* **41**, 7312–7317 (2002).
54. Sforça, M. L. et al. How C-terminal carboxyamidation alters the biological activity of peptides from the venom of the eumenine solitary wasp. *Biochemistry* **43**, 5608–5617 (2004).
55. Kim, T. & Loh, Y. P. Protease nexin-1 promotes secretory granule biogenesis by preventing granule protein degradation. *Mol. Biol. Cell* **17**, 789–798 (2006).
56. Koshimizu, H., Cawley, N. X., Kim, T., Yergey, A. L. & Loh, Y. P. Serpinin: a novel chromogranin A-derived, secreted peptide up-regulates protease nexin-1 expression and granule biogenesis in endocrine cells. *Mol. Endocrinol.* **25**, 732–744 (2011).
57. Benner, C. et al. The transcriptional landscape of mouse beta cells compared to human beta cells reveals notable species differences in long non-coding RNA and protein-coding gene expression. *BMC Genom.* **15**, 620 (2014).
58. MacDonald, P. E. et al. The multiple actions of GLP-1 on the process of glucose-stimulated insulin secretion. *Diabetes 51* **Suppl 3**, S434–S442 (2002).
59. Wettergren, A., Pridal, L., Wøjdemann, M. & Holst, J. J. Amidated and non-amidated glucagon-like peptide-1 (GLP-1): non-pancreatic effects (cephalic phase acid secretion) and stability in plasma in humans. *Regul. Pept.* **77**, 83–87 (1998).

Acknowledgements

We acknowledge sharing of data from the GoT2D and T2D-GENES consortia before publication. We thank J. Galvanovskis (University of Oxford) for microscopy assistance, J. Lyon (Alberta Diabetes Institute IsletCore) for his work on human islet isolations, and J. Buteau and Y. Wang (both Alberta Diabetes Institute) for their assistance with imaging human pancreatic sections. We also thank the Human Organ Procurement and Exchange Program (Edmonton) and the Trillium Gift of Life Network (Toronto) and other organ

procurement agencies for their efforts in obtaining human pancreata for research. A.L.G. is a Wellcome Trust Senior Fellow in Basic Biomedical Science. P.E.M. holds a 2016–2017 Killam Annual Professorship. M.I.M. is a Wellcome Senior Investigator. S.K.T. is a Radcliffe Department of Medicine Scholar. S.S. is funded by the Medical Research council (MC_ST_15019 [2015 DTG/DTA]). J.C. is funded by the Oxford – Medical Research Council Doctoral Training Partnership and the Nuffield Department of Clinical Medicine. This work was funded by the Wellcome Trust (095101 (A.L.G.), 200837 (A.L.G.), 098381 (M.I.M.), 106130 (A.L.G., M.I.M.), 203141 (M.I.M.), 090531 (P.R.)), Medical Research Council (MR/L020149/1) (M.I.M., A.L.G., P.R.), European Union Horizon 2020 Programme (T2D Systems) (A.L.G.), and National Institutes of Health (U01-DK105535; U01-DK085545) (M.I.M., A.L.G.). Human islet isolation and phenotyping was supported by funding from the Alberta Diabetes Foundation (P.E.M.F.). The research was funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (A.L.G., M.I.M., P.R.). The views expressed are those of the author(s) and not necessarily those of the NHS, NIHR or the Department of Health.

Author contributions

S.K.T., A.R., B.H., and A.L.G. conceived the study. A.R., B.H., S.S., M.M.U., A. Barrett, C.J.G., N.L.B., P.R., and A.L.G. performed kinetic and cellular characterization of T2D-associated alleles. S.K.T., A.R., B.H., S.S., A.C., P.R., and A.L.G. performed the characterization of PAM knockdown in β-cells. S.K.T., X.-Q.D., A. Bautista, A.F.S., J.E.M.F., P.E.M., and A.L.G. performed characterization of primary human islets. A.J.P., M.I.M., and A.L.G. performed islet transcriptomics. J.C., M.I.M., and A.M. performed plasma proteomics. A.R., S.K.T., and A.L.G. wrote the manuscript. All authors approved the final draft of the manuscript.

Competing interests

S.K.T. is now an employee of Vertex Pharmaceuticals, and N.L.B. is now an employee of Novo Nordisk, although all experimental work was carried out under employment at the University of Oxford. A.L.G. has received research funding and honoraria from Novo Nordisk. M.I.M. serves on advisory panels for Pfizer, Novo Nordisk and Zoe Global; has received honoraria from Pfizer, Novo Nordisk and Eli Lilly; has stock options in Zoe Global; and has received research funding from AbbVie, Astra Zeneca, Boehringer Ingelheim, Eli Lilly, Janssen, Merck, Novo Nordisk, Pfizer, Roche, Sanofi Aventis, Servier and Takeda.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-018-0173-1.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to A.L.G.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Ethics statement. All studies were approved by the University of Oxford's Tropical Research Ethics Committee (OxTREC reference 2–15) or the Oxfordshire Regional Ethics Committee B (REC reference 09/H0605/2). Human pancreatic tissue was obtained from subjects at postmortem examination according to local and national ethics permissions. Human pancreatic islets were isolated, with ethical approval and clinical consent, at the Diabetes Research and Wellness Foundation Human Islet Isolation Facility (Oxford Centre for Diabetes, Endocrinology and Metabolism, Oxford, UK). All animal experiments were conducted in accordance with the UK Animals (Scientific Procedures) Act 1986 and the University of Oxford ethical guidelines, and were approved by the local ethical committees.

Cloning of PAM expression plasmids. We purchased a pCMV6 expression vector encoding integral membrane *Homo sapiens PAM* (NM_138821.2) with an in-frame C-terminal Myc-Flag tag (PAM-Myc-Flag) from Origene. Luminal functional PAM-Myc-Flag encompassing the first 710 amino acids was generated via PCR using the primers luminal (F) and luminal (R) (Supplementary Table 2). Puromycin-resistant PAM-Myc-Flag expression vectors were generated via BglII/XmaI digestion, blunt ending with T4 DNA polymerase and subcloning into HpaI-digested pIRES puro 2 (Takara Bio USA). Protein alterations were introduced via site-directed mutagenesis using the Stratagene QuikChange II kit (Agilent Biotechnologies), according to manufacturer's instructions, with the following primers: SDM_S539W_F and SDM_S539W_R for generating p.Ser539Trp PAM; SDM_D563G_F and SDM_D563G_R for generating p.Asp563Gly PAM; and SDM_Y651F_F and SDM_Y651F_R for generating p.Tyr651Phe PAM (Supplementary Table 2). All plasmids were verified by sequencing.

Cell culture and transfection. Primary human islets were freshly isolated from deceased donors at the Oxford Centre for Islet Transplantation in Oxford, UK or at the Alberta Diabetes Institute IsletCore as previously described[60,61]. Islets were then cultured in CMRL medium (Thermo Fisher Scientific) at 37 °C or kept in UW cold storage solution (Viaspan) at 4 °C for short-term storage before being processed. HEK293 cells were routinely passaged in DMEM 6429 (Sigma-Aldrich) supplemented with 10% fetal calf serum, 50 international units ml^{-1} penicillin and 50 µg ml^{-1} streptomycin. EndoC-βH1 cells were routinely passaged in growth medium supplemented with 5.5 mM glucose[24]. All cells were maintained at 37 °C and 5% CO_2.

To generate PAM stable cell lines, HEK293 cells were transfected using FuGENE 6 (Promega Corporation), according to the manufacturer's instructions. Cells were routinely passaged and maintained them in 1 µg ml^{-1} puromycin. For fluorescence microscopy, EndoC-βH1 cells were transfected in chamber slides (BD Biosciences) using FuGENE 6. For siRNA experiments, EndoC-βH1 cells were seeded in Opti-MEM reduced serum-free medium (Thermo Fisher Scientific) containing 25 nM ON-TARGET *plus* SMARTpool siRNA complexes (Dharmacon) preformed with Lipofectamine RNAiMAX transfection reagent (Thermo Fisher Scientific), according to the manufacturer's instructions. Growth medium was replaced 24 h later and cells were incubated for a further 48 h. For PAM inhibitor experiments, seeded cells were treated with 500 µM 4P3BA or dimethylsulfoxide vehicle (Sigma-Aldrich) for 72 h.

Recombinant PAM protein production. HEK293 stable cell lines were seeded into Nunc TripleFlasks (Thermo Fisher Scientific). The medium was replaced with DMEM 1145, supplemented with 1 mM sodium pyruvate and incubated for 3–5 days. Supernatant containing luminal recombinant PAM protein was harvested directly from the flasks, filtered into a sterile storage bottle (Corning GmbH) and pH corrected to 5.5 with HCl. The filtered supernatant was concentrated using a Centricon Plus-70 Centrifugal Filter Device (Millipore Limited) with a 30 kDa molecular weight cutoff, according to the manufacturer's instructions.

Quantitation of recombinant PAM. For the relative quantitation of recombinant PAM proteins, 3 × 1 µl aliquots of concentrated supernatant were separated on a 4–20% Criterion TGX Stain-Free Precast Gel (Bio-Rad Laboratories, Inc) and transferred to polyvinylidene difluoride using a Trans-blot Turbo Transfer Pack (Bio-Rad Laboratories), according to the manufacturer's instructions. Membranes were incubated overnight in 4A6 anti-Myc tag antibody (Millipore Limited) at 4 °C and developed using the Clarity Western ECL Blotting Substrate (Bio-Rad Laboratories), according to manufacturer's instructions. Proteins were visualized on a ChemiDoc MP (Bio-Rad Laboratories). Lanes and bands were drawn in the Image Lab version 5.0 for Windows software; the average band volume (intensity) was quantified for each triplicate of samples. The amount of recombinant variant PAM protein per µl was then calculated relative to wild-type PAM.

Kinetic analysis of PAM amidating activity. We used the spectrophotometric detection of chemically converted glyoxylate to measure the amidating activity of PAM[20]. An assay mix of 100 mM MES pH 5.5, 2 µM copper sulfate, 10 mM L-ascorbic acid, 0.1 mg ml^{-1} bovine liver catalase and 20 mM hippuric acid (Sigma-Aldrich) was split into glass vials. Equimolar amounts of recombinant PAM were added to each vial and incubated in a 37 °C water bath. Every 30 min, samples were removed to a conical 96-well plate (4titude Limited) containing 40 mM EDTA pH 8.0. The glyoxylate produced was converted to glyoxylate phenylhydrazone by adding 0.1% phenylhydrazine-HCl (Sigma-Aldrich) to the plate, incubating it at room temperature for 15 min and saturating it with concentrated (37%) HCl. The contents of the entire plate were then aspirated and dispensed into a flat-bottomed 96-well plate (Corning GmbH) containing 0.2% potassium ferricyanide using an i-Pipette Automated Pipettor (Apricot Designs). We measured the absorbance of 1,5-diphenylformazan at 515 nm using a VersaMax Microplate Reader (Molecular Devices, LLC).

Kinetic analysis of variant PAM substrate affinity. For the K_m calculations, a 12-point, twofold serial dilution of hippuric acid in which the maximum concentration was 20 mM was prepared in the assay mix (100 mM MES pH 5.5, 2 µM $CuSO_4$, 10 mM L-ascorbic acid, 0.1 mg ml^{-1} bovine liver catalase). Recombinant PAM or empty vector control supernatant was then added. The amount of p.Asp563Gly PAM added to the assay was increased so that it displayed an equal V_{max} to wild-type PAM. The volume of empty vector supernatant matched the volume of wild-type PAM used in that experiment. Vials were incubated in a 37 °C water bath for 1 h. Samples were then removed into a conical 96-well plate containing 40 mM EDTA pH 8.0.

Gene expression analyses. Gene expression was measured with TaqMan real-time PCR assays. RNA was purified from TRIzol (Thermo Fisher Scientific) homogenates or sourced from commercially available RNA tissue panels: Human Total RNA Master Panel II (Takara Bio USA); FirstChoice Human Total RNA Survey Panel (Thermo Fisher Scientific); and mouse Total RNA Master Panel (Thermo Fisher Scientific). Pooled human adult pancreas RNA from three donors, and pooled human adult islet RNA from three donors, was sourced from the Oxford Islet Biobank[34]. Pooled mouse islet RNA was sourced from four 12–28-week-old Naval Medical Research Institute mice. Hypothalamic and pituitary RNA were commercially sourced from AMS Biotechnology Ltd. Complementary DNA was generated using the GoScript Reverse Transcription System (Promega Corporation), according to the manufacturer's instructions, and analyzed with the 7900HT Fast Real-Time PCR System (Thermo Fisher Scientific). C_t values were normalized to the combined average of the housekeeping genes peptidylprolyl isomerase A (*PPIA*) and TATA-box binding protein (*TBP*). *PAM* expression in sorted β-cells and non-β-cells was analyzed based on previously published RNA sequencing data[22].

Allele-specific expression was performed using publicly available RNA sequencing data from the previously described islets of 118 individuals[34]. In brief, 40 samples were freshly isolated at the Oxford Centre for Islet Transplantation as described[60]. The other samples, from Edmonton, Canada, were either extracted from the long-term cryopreserved biobank and thawed as described[62] ($n = 65$) or freshly isolated from donor pancreas ($n = 13$). Poly-A selected libraries were prepared from total RNA at the Oxford Genomics Centre, multiplexed (3 samples per lane) and clustered; 100 nucleotide paired-end reads were sequenced using Illumina TruSeq v3 chemistry on the Illumina HiSeq 2000 Sequencing System (Illumina, Inc). DNA was extracted from the spleen or the exocrine fraction of the islet for most donors, or from the islets when no other tissue was available. All samples were genotyped on the Illumina Omni2.5 Exome-8 version 1.3 BeadChip array.

In total, 11 out of 118 individuals were heterozygous at rs35658696 and were thus used for allele-specific expression analysis. Reads were mapped with the WASP pipeline[35] combined with the Spliced Transcripts Alignment to a Reference (STAR) aligner[63] and based on human genome assembly GRCh37.75[64]. The WASP pipeline removes reads with evidence of mapping bias. Final allele counts were then calculated with an allele counter[65] and a binomial test for allelic imbalance was performed in R against a null allelic proportion of 0.5.

Associations for rs35658696 (p.Asp563Gly) and rs78408340 (p.Ser539Trp) with PAM protein plasma levels in humans were looked up in the supplementary material to the article by Sun et al.[33]. Association analyses in their study had been performed within a linear regression framework under an additive model, adjusting for age, sex, time between sampling and processing, and three principal components. Effect directions were then aligned to correspond to the T2D risk alleles.

Immunofluorescent labeling and microscopy. Human pancreatic specimens fixed in 10% formalin and embedded in wax were used to localize PAM in islets. Sections of 4 µm were de-waxed, rehydrated and heated to 95 °C in 10 mmol l^{-1} Tris + 1 mmol l^{-1} EDTA pH 6.0 for 10 min for antigen retrieval. Tissue sections were incubated overnight in anti-PAM antibody (ab109175; Abcam), guinea pig anti-insulin antibody (in-house antibody) and mouse anti-glucagon antibody (G2654; Sigma-Aldrich). PAM immunoreactivity was enhanced using a tyramide signal amplification step (Thermo Fisher Scientific) and visualized with Alexa Fluor 488 (Thermo Fisher Scientific). Insulin and glucagon were visualized with Alexa Fluor 633 (Thermo Fisher Scientific) and anti-mouse tetramethylrhodamine (Sigma-Aldrich), respectively.

EndoC-βH1 cells transfected with *PAM* expression vectors were fixed in 4% paraformaldehyde, permeabilized in 0.1% (v/v) Triton X-100 and blocked in 10% (w/v) bovine serum albumin. Primary antibodies (4A6 anti-Myc, Belfast anti-insulin, TGN46 anti-trans Golgi (Sigma-Aldrich) and H-70 anti-calnexin (sc-11397; Santa Cruz Biotechnology, Inc)) were incubated overnight at 4 °C.

Slides were then incubated in Alexa Fluor-conjugated secondary antibodies (Thermo Fisher Scientific) and mounted in VECTASHIELD Mounting Medium with DAPI (Vector Laboratories Ltd). Cells were visualized with a Zeiss LSM 510 META confocal laser scanning module arranged on an Axiovert 200 microscope and a Plan-Apochromat 63×/1.4 oil immersion objective (Carl Zeiss). An argon laser was used to excite Alexa Fluor 488 at λ = 488 nm; helium–neon lasers 543 and 633 were used to excite Alexa Fluor 543 at λ = 543 nm and Alexa Fluor 633 at λ = 633nm, respectively. The signals were collected between 640 and 694 nm for Alexa Fluor 633, between 512 and 533 nm for Alexa Fluor 488, and between 587 and 619 nm for Alexa Fluor 543. The main dichroic beam splitter (HFT 488/543/633) was combined with a secondary one, either the NFT490 or NFT545 or NFT635vis. DAPI was excited in two-photon mode using the 740-nm line of an infrared light Chameleon laser combined with an HFT KP 650 filter (Carl Zeiss).

Gene silencing in EndoC-βH1. To perform gene silencing by RNA interference, a master mix of Opti-MEM reduced serum-free medium (Thermo Fisher Scientific) was pre-mixed with 0.4% RNAiMAX transfection reagent (Thermo Fisher Scientific) and 50 nM siRNA (commercially available SMARTpools from Dharmacon against *PAM*, *CHGA* or *IAPP* as indicated). The solution was then incubated at room temperature for 15 min to form siRNA–lipid complexes before 50 µl of the master mix was added to each well on a 96-well black OPTI-plate (Corning GmbH). Cells resuspended in standard culturing media were added in a 1:1 volume until a final density of 75,000 cells cm^{-2} and siRNA concentrations of 25 nM were reached. After 24 h, the transfection media were replaced with complete culturing media. After a further 2 days, cells were taken forward for functional assays or gene expression analyses as required.

Insulin secretion assays in EndoC-βH1. Seventy-two hours after gene silencing, cells prepared as described earlier were starved overnight in standard cell culture medium supplemented with glucose to a final concentration of 2.8 mM. Following overnight incubation, cells were further incubated for 1 h in culture medium with 0 mM glucose. The medium was then replaced with the appropriate conditions to initiate secretion in static incubation. After 1 h, supernatants were harvested in a staggered fashion, leaving a residual volume of supernatant (~30 µl in 96-well format) to minimize disturbances of the adherent cells. Supernatant samples were then centrifuged at 300 *g* in a pre-chilled centrifuge for 5 min and a small volume of the supernatant was transferred to a new plate.

Meanwhile, adherent cells were processed for cell counting based on the CyQUANT Direct Cell Proliferation Kit (Thermo Fisher Scientific), according to the manufacturer's recommendations. Briefly, residual medium was replaced with medium supplemented with 0.2% direct nucleic acid dye and 1% background suppressor dye. Cells were then incubated for 1 h at 37 °C in the dark and fluorescence was measured on the EnSpire Multimode Plate Reader (PerkinElmer) using fluorescein isothiocyanate filter settings (ex480/em535 nm). Cell counts were expressed as relative fluorescence units (RFU) normalized to the earliest time point. Following the CyQUANT assays, cells were washed once in standard media, before adding ice-cold acid-ethanol buffer (1.5% concentrated HCl, 75% ethanol and 23.5% deionized water) to extract the intracellular insulin content. Plates were then covered with parafilm to prevent evaporation and stored at −20 °C for short-term storage or −80 °C for long-term storage.

To analyze the insulin samples, the Insulin (human) AlphaLISA Detection Kit (PerkinElmer) was used to measure the concentrations of both supernatants and cellular contents. Samples were diluted in 1 × AlphaLISA immunoassay buffer (1:10 for supernatants and 1:200 for contents) and then analyzed in white 1/2 AreaPlate-96 (PerkinElmer) on the EnSpire using the AlphaScreen protocol (PerkinElmer), according to the manufacturer's recommendations. Normalized secretion values were calculated on a per-well basis as the ratio of total secreted insulin to total intracellular insulin content (%), or as the ratio of total secreted insulin to cell counts (pg/RFU), which were averaged across replicates. Corrected values were then normalized to basal insulin secretion rates for negative control on a per-experiment basis to eliminate variability in baseline secretion rates across independent experiments.

Insulin secretion assays in primary human islets. Insulin secretion assays were performed on intact human islets isolated from 16 risk allele carriers and 16 controls, matched for age ($P = 0.88$), sex, BMI ($P = 0.88$) and time of isolation. The experiments were performed at 37 °C in Krebs–Ringer buffer (KRB) (NaCl 115 mM; KCl 5 mM; $NaHCO_3$ 24 mM; $CaCl_2$ 2.5 mM; $MgCl_2$ 1 mM; HEPES 10 mM; 0.1% bovine serum albumin, pH 7.4) with glucose concentrations as indicated. For each donor, triplicate groups of 15 handpicked human islets were pre-incubated for 2 h with 1 mM glucose KRB. Islets were subsequently incubated for 1 h in 1 mM glucose KRB, after which supernatants were collected and replaced with 10 mM or 16.7 mM glucose KRB. After another hour, supernatants were again collected, and total insulin content extracted from the islet pellet using acid-ethanol. Samples were stored at −20 °C and assayed for insulin via electrochemiluminescence (Meso Scale Discovery, Rockville, MD, USA or ALPCO, Salem, NH, USA). Insulin secretion measurements at 1 mM and 16.7 mM glucose were performed for all individuals ($n = 16$ per group); measurements at 10 mM glucose were performed for a subset hereof ($n = 9$ per group). Averages were calculated after excluding any outliers using Grubb's test, and either displayed as raw values or normalized to insulin content on a per-individual basis.

Membrane capacitance assays in EndoC-βH1. Electrophysiological measurements were performed using an HEKA EPC10 Patch Clamp Amplifier (HEKA Elektronik Dr. Schulze GmbH) at 32 °C using the standard whole-cell configuration as previously described[66]. The pipette solution (intracellular medium) contained 125 mM Cs-glutamate, 20 mM CsCl, 15 mM NaCl, 1 mM $MgCl_2$, 5 mM HEPES, 0.05 mM EGTA, 0.1 mM cAMP and 3 mM Mg adenosine triphosphate (pH 7.15 with CsOH). The extracellular solution contained 1 mM glucose, 118 mM NaCl, 5.6 mM KCl, 1.2 mM $MgCl_2$, 5 mM HEPES and 2.6 mM $CaCl_2$ (pH 7.4 with NaOH). Tetraethylammonium chloride (20 mM) was added to block residual potassium currents. Calcium sensitivity of the vesicles was inferred from the relationship between calcium current density and exocytosis. The calcium current was measured at the first pulse, after inactivation of the sodium current (5 ms).

Analysis of exocytosis in primary human β-cells. Human pancreatic islets were isolated and prepared at the Alberta Diabetes Institute IsletCore as previously described; these had already been genotyped[34,61]. Single-cell capacitance responses were monitored as described earlier for EndoC-βH1 cells (including 0.1 mM cAMP), with an additional pre-incubation period of 1 h at 1 mM glucose. Following recordings, cells were positively identified by insulin immunostaining. The kinetics of exocytosis were examined at 1 and 10 mM glucose for seven risk allele carriers and seven controls, matched as groups for age ($P = 0.52$), sex and BMI ($P = 0.68$). The total number of cells per individual were [3; 5; 6; 6; 7; 8; 10] (low glucose, controls), [2; 3; 6; 7; 7; 7; 9] (high glucose, controls), [3; 4; 4; 5; 6; 8; 12] (low glucose, carriers), and [4; 4; 4; 6; 6; 8; 8] (high glucose, carriers). Mean capacitance was calculated per condition for each depolarization step.

Electron microscopy. Primary islets or EndoC-βH1 cells were fixed in 2.5% glutaraldehyde, post-fixed in 2% uranyl acetate, dehydrated in graded methanol and embedded in London Resin Gold (Agar Scientific Ltd). Ultrathin sections cut onto nickel grids were immunolabeled for PAM (ab109175; Abcam) followed by anti-rabbit biotin (Vector Laboratories Ltd) and streptavidin gold 15 nm (British Biocell International). For EndoC-βH1 cells, insulin was immunolabeled with an in-house antibody followed by anti-guinea pig gold 10 nm (British Biocell International). Sections were viewed on a JEOL JEM-1010 microscope (accelerating voltage 80 kV) with a digital camera (Gatan, Inc.). To estimate PAM concentration in secretory vesicles, quantification of immunogold labeling was performed by a blinded observer using the ImageJ software. The following parameters were determined: cross-sectional vesicle area; vesicle density (number of vesicles per cytoplasmic area); and insulin or PAM labeling per vesicle (number of immunogold particles per vesicle).

Western blot analysis of CgA amidation. To mimic the effect of gene silencing, cells were cultured for 24 h with the irreversible *PAM* inhibitor 4P3BA (Sigma-Aldrich) at a concentration of 500 µM[45]. Whole-cell extracts from siRNA or inhibitor-treated EndoC-βH1 cells were prepared as previously described[67]. Supernatant was harvested directly from cells and filtered with a 0.22-µm filter. Protein yields were determined with the Bio-Rad Protein Assay (Bio-Rad Laboratories), according to the manufacturer's instructions. For western blotting, 10 µg whole-cell extract and 0.5% total supernatant volume were analyzed with denaturing SDS–PAGE. Antibodies specific for CgA-Gly (ab52983; Abcam)[45] and total CgA (ab15160; Abcam) were used to calculate the ratio of amidated to non-amidated CgA using the Image Lab, version 5.0 for Windows (Bio-Rad Laboratories).

Immunohistochemistry of human pancreatic sections. Paraffin-embedded human pancreas biopsy sections were rehydrated, and citrate buffer antigen retrieval was performed. Consecutive 5-µm sections were stained for either total CgA (ab15160, 1:100; Abcam) or CgA-Gly (ab52983, 1:100; Abcam), plus insulin (1:500; Dako Canada) at 4 °C overnight. Species-specific secondary antibodies were Alexa Fluor 488 or 594 conjugates (Thermo Fisher Scientific) diluted 1:400 in 2% goat serum. Slides were cover-slipped with ProLong Gold Antifade Mountant with DAPI (Thermo Fisher Scientific). Images were acquired with an Olympus VS120 slide scanner (Olympus Canada Inc., Richmond Hill, Ontario, Canada), utilizing excitation filter wavelengths of 432 nm for DAPI, 515 nm for AlexaFluor 488 and 595 nm for AlexaFluor 594. Images were analyzed with the Olympus VS120-S6 analysis software (Olympus Corporation, Shinjuku, Tokyo, Japan. Olympus VS120-S6-ASW 2.9 [Build 13741] for Windows). and the ratio of mean total CgA and CgA-Gly signal intensity of paired insulin-positive regions was calculated.

Statistics. All statistical analyses were performed in R version 3.3.1 for Windows, except electron microscopy data and electrophysiological measurements in EndoC-βH1 cells, which were analyzed with the OriginPro software, version 8.5.1 for Windows (OriginLab Corporation). For kinetic data, rates were found to be dependent on substrate concentration; exponential models were fitted. Enzyme kinetics, electron microscopy, electrophysiological and cellular assay data were analyzed with two-tailed Student's *t*-tests where a single independent variable was being tested, and with two-way analysis of covariance (ANCOVA) followed by

Tukey's HSD post-hoc test in cases where two experimental variables were being tested (that is, glucose level and genotype status or knockdown). Western blot data were analyzed with paired, two-tailed Student's *t*-tests to eliminate differences in baseline intensity between experiments. Graphs show means of the indicated number of biological replicates, and error bars are s.e.m. Box plots display the interquartile range (IQR), whiskers indicating 1.5x the IQR, and outliers.

URLs. GTEx portal, https://www.gtexportal.org/home/gene/PAM.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. Previously published data sets analyzed during the current study can be accessed as follows: transcriptomic data from sorted human islets are available from the European Genome-phenome Archive under accession EGAS00001000442; tissue expression data for *PAM* from the GTEx consortium are available through the GTEx portal (see URLs); genotype data for primary human islets used in the analysis of exocytosis data are available from the European Genome-phenome Archive under accession EGAD00001001601; and secreted islet factors were analyzed based on data available from the supplementary information of PMID 21212933[68]. Genetic associations for plasma PAM expression levels are available in the supplementary material to the article by Sun et al.[33]. Data generated during this study are included in the published article (and its supplementary information files) or are available from the corresponding author on reasonable request.

References

60. Cross, S. E., Hughes, S. J., Clark, A., Gray, D. W. & Johnson, P. R. Collagenase does not persist in human islets following isolation. *Cell Transplant.* **21**, 2531–2535 (2012).
61. Lyon, J. et al. Research-focused isolation of human islets from donors with and without diabetes at the Alberta Diabetes Institute IsletCore. *Endocrinology* **157**, 560–569 (2016).
62. Manning Fox, J. E. et al. Human islet function following 20 years of cryogenic biobanking. *Diabetologia* **58**, 1503–1512 (2015).
63. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
64. McPherson, J. D. et al. A physical map of the human genome. *Nature* **409**, 934–941 (2001).
65. Castel, S. E., Levy-Moonshine, A., Mohammadi, P., Banks, E. & Lappalainen, T. Tools and best practices for data processing in allelic expression analysis. *Genome Biol.* **16**, 195 (2015).
66. Chandra, V. et al. RFX6 regulates insulin secretion by modulating Ca^{2+} homeostasis in human β cells. *Cell Rep.* **9**, 2206–2218 (2014).
67. Lees, M. J. & Whitelaw, M. L. Multiple roles of ligand in transforming the dioxin receptor to an active basic helix-loop-helix/PAS transcription factor complex with the nuclear protein Arnt. *Mol. Cell. Biol.* **19**, 5811–5822 (1999).
68. Yang, Y. H. et al. Paracrine signalling loops in adult human and mouse pancreatic islets: netrins modulate beta cell apoptosis signalling via dependence receptors. *Diabetologia* **54**, 828–842 (2011).

nature research

Corresponding author(s): Anna L Gloyn

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Functional and transcriptomic data from primary islets were analyzed based on the maximum number of risk allele carriers available in the combined collection of islets from study centers in Oxford, UK, and Edmonton, Canada. For cellular data, effect sizes could not be estimated prospectively (precluding power analysis), and standard scientific conventions were followed for these studies. Specifically, for insulin secretion and exocytosis experiments, three independent passages of cells were analyzed in biologically independent replicates. In each case, cells were only taken forward for detailed phenotyping if a satisfactory glucose responsiveness (fold induction > 1.5) could be ascertained. For the quantitative electron microscopy, a total of ten cells were counted per treatment. Tissue expression panels were analyzed in single replicates of pooled tissues (full information available from commercial suppliers, as stated in the methods). In the case of determining CgA amidation status, additional replicates were requested at the revision stage; the estimated effect size derived from the first set of experiments was used to inform the level of replication required in the follow-up analysis.

2. Data exclusions

Describe any data exclusions.

To remove significant outliers among technical replicates for primary islet data, Grubbs outlier test was systematically applied on a per-individual basis.

3. Replication

Describe whether the experimental findings were reliably reproduced.

For the human islet data, replication could not be attempted due to the difficulties in obtaining primary islets from cadaveric donors and the low frequency of the T2D risk alleles in PAM. Several aspects of the cellular data, including insulin secretion defects, have been replicated by parallel studies in the lab, and on-going work has successfully replicated the effects of the PAM variants in kinetic assays. Electron microscopy was independently quantified by two blinded observers, identifying the same overall trends. Western blotting following PAM knockdown or inhibition was successfully reproduced in at least three different passages of the cell line, in experiments performed by two independent researchers.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

PAM variant carriers were matched for age, gender, and BMI to available controls in the same cohort of primary human islets (n = 187). For cellular phenotyping, samples were randomly positioned on microplates to minimize any possibly systematic bias from technical artefacts.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

The post-hoc matching of islet samples was performed based on the three listed covariates and sample genotype, with the researcher being blinded to experimental data. All subjective quantifications of cellular data, including the scoring of electron micrographs, were performed by blinded observers.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Data were analysed using:
- Olympus VS120-S6 analysis software,
- OriginPro V8.5.1
- ImageJ 1.52b
- Image Lab 5.0
- R version 3.3.1 using the following standard packages: drc, ggplot2, mass, pwr, scales, stats.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

All unique materials used are available from the authors upon request or from standard commercial sources (as indicated in relevant sections of the online methods), except the cell line, EndoC-βH1, which is available from Univercell-Biosolutions (formerly EndoCells), subject to an MTA.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

All antibodies were validated by dilution test in the cell models used either for immunofluorescence (IF), electron microscopy (EM) or western blot (WB).
Primary Ab:
• rabbit anti-PAM (Abcam ab109175, clone EPR2643(2), IF and EM 1:100)
• guinea-pig anti-insulin (Belfast, in-house antibody, IF 1:500), ref doi:10.1038/ncomms5639
• mouse anti-glucagon (Sigma-Aldrich G2654, clone K79bB10, IF 1:500)
• mouse anti-Myc (Millipore, Clone 4A6, WB 1:10000, IF 1:1000)
• rabbit anti-TGN46 (Invitrogen, PA5-23068, IF 1:200)
• rabbit anti-calnexin H-70 (Santa Cruz Biotechnology sc-11397, IF 1:50)
• rabbit anti-CgAGly (Abcam ab52983, clone EP1031Y, IF and WB 1:100)
• rabbit polyclonal anti-total CgA (Abcam, ab15160, IF and WB 1:100)
• guinea pig anti-insulin (Dako, A0564, EM 1:500)

Secondary Ab:
• mouse alexa Fluor 488 (Life Technologies, A11029, IF 1:500)
• goat anti-guinea pig Alexa Fluor 488 (Life technologies, A11073, IF 1:500)
• guinea- pig Alexa Fluor 633 (Life Technologies, A21105, IF 1:250)
• goat anti-rabbit Alexa Fluor 594 (ThermoFisher Technologies, A11037, IF 1:400)
• goat anti-mouse TRITC (Sigma-Aldrich, T7782, lot# 023M4752V, IF 1:250)
• tyramide amplification (IF, thermofisher, TSA Kit T20912)
• anti-rabbit biotin (Vector Laboratories, BA-1000, EM 1:50)
• streptavidin gold 15nm (Biocell, EMSTP15, EM 1:20)
• anti-guinea pig gold 10nm (Biocell, EMGAG10, EM 1:20)

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

• The EndoC-βH1 cell line was acquired from Raphael Scharfmann (EndoCells, Paris) under an MTA.
• HEK293 has been purchased from sigma, ref: 85120602.

b. Describe the method of cell line authentication used.

The EndoC-βH1 has not been authenticated by formal STR profiling, but the cell line is routinely subjected to rigorous testing in a number of different ways, all of which have confirmed the identity as a bona fide human beta cell line. Relevant tests include microarray-based genotyping, RNA-seq, ATAC-seq, and extensive cellular phenotyping, including insulin secretion and content (see Thomsen et al, 2016, Diabetes, PMID: 27554474) and electrophysiology (see Hastoy et al, BiorXiv, doi: 10.1101/226282)

c. Report whether the cell lines were tested for mycoplasma contamination.

Cells were tested for mycoplasma contamination on a quarterly basis, and tested negative on all occasions.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

HEK-293 cells were used exclusively for expression and purification of recombinant PAM protein, and the identity of the cell line is, as such, not relevant to the interpretation of the kinetic data.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

Pancreatic islet RNA was extracted from female NMRI mice between 12 and 28 weeks old.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Islet donors analyzed for insulin secretion and content measurements had the following population characteristics (Carriers; Controls): n (16; 16), Age [y] (51.8; 52.6), BMI [kg/m2] (29.6; 29.3), Gender [m/f] (8/8; 8/8).

Islet donors analyzed for exocytosis measurements had the following population characteristics (Carriers; Controls): n (7; 7), Age [y] (49; 58.4), BMI [kg/m2] (29.3; 26.6), Gender [m/f] (5/2; 3/4).