MUTATION ANALYSIS OF WOLFRAM SYNDROME PATIENTS AND FUNCTIONAL STUDY OF THE WOLFRAMIN PROTEIN

By

SAMANTHA PRINCE

A thesis submitted to the University of Birmingham

For the degree of

DOCTOR OF PHILOSOPHY

Department of Medical and Molecular Genetics

School of Clinical and Experimental Medicine

University of Birmingham

January 2012

ABSTRACT

Mutations of the *WFS1* gene are responsible for most cases of Wolfram syndrome (WS), a rare, recessively inherited neurodegenerative disorder characterised by juvenile-onset non-autoimmune diabetes mellitus and optic atrophy. Variants of *WFS1* are also associated with non-syndromic hearing loss and type-2 diabetes.

Understanding the function of the *WFS1* protein-product wolframin, would enable developments in targeted therapy for WS patients and important insights to its possible contribution to type-2 diabetes pathogenesis. This study was aimed at expanding the spectrum of WS-associated genetic mutations and clinical data, and investigating the molecular mechanisms responsible for phenotypic variation associated with *WFS1*-mutation.

The mutational and phenotypic spectrum of WS is broadened by our report of novel *WFS1* mutations and a case of *WFS2*-associated WS. New perspectives on the molecular mechanisms linking mutation to disease manifestation are also taken by characterisation of representative *WFS1* mutations specific to phenotype, identification of potentially novel *WFS1* interacting partners, and the first evidence linking *WFS1* with the exocrine pancreas. Our data suggests that some *WFS1*-mutations may allow residual protein function and these findings lay the groundwork for future functional investigation of mutated wolframin to explore this hypothesis further.

This thesis is dedicated to the memory of

Mark Anthony Pallas

12th November 1970 - 28th January 2010

Thank you for sharing your wonderful father with us x

ACKNOWLEDGEMENTS

I would like to acknowledge my supervisor Professor Tim Barrett without whom this work would not have been possible. Many thanks for all the patience and support.

Special thanks also go to Dr Malgosia Zatyka, who has been an invaluable mentor in the laboratory throughout my time in Birmingham. She is a meticulous, hard-working person and a fine example of the kind of scientist I'd like to be.

The rest of the Medical and Molecular Genetics laboratory members are also acknowledged not only for their technical expertise and help, but also for their friendship, humour and oftentimes emotional support. A particular mention must go to: Dr Christopher Ricketts, Dr Mark Morris, Dr Dewi Astuti, Dr Christopher Bruce, Dr Ania Straatman-Iwanowska and Dr Stephen Jacques.

Outside of the laboratory I have had the tremendous support and camaraderie of my pals in the Molecular Neuroscience Group with whom I've shared some memorable times and some great laughs. I'm happy to say I have gained some life-long friends.

I would like to extend my eternal gratitude to my family and partner Robert. I am lucky enough to have the most supportive and loving parents and an inspirational step-father who has always had faith in me. I would not have gotten through the last year without Robert to hold my hand and my little angels Maggie and Brooke.

Finally, I would like to acknowledge the Medical Research Council for providing funding for this work

CONTENTS

CHAPTER 1

INTRODUCTION

CHAPTER 2

MATERIALS AND METHODS

CHAPTER 3

MUTATION ANALYSIS OF WOLFRAM SYNDROME PATIENTS

CHAPTER 4

CHARACTERISATION AND FUNCTIONAL ANALYSIS OF SELECTED *WFS1* VARIANTS

CHAPTER 5

IDENTIFYING *WFS1* MOLECULAR PARTNERS

CHAPTER 6

GENERAL DISCUSSION

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

ABBREVIATIONS

A	Adenine
AD	activating domain
APS	ammonium persulphate
ASK1	apoptosis signal-regulating kinase
ATF4	activating transcription factor 4
ATF6α	activating transcription factor 6α
ATP	adenosine triphosphate
BAX	Bcl2-associated X protein
Bcl2	B-cell lymphoma 2
BiFC	bimolecular fluorescence complementation
BiP	immunoglobulin binding protein/GRP78
bp	base pairs
BSA	bovine serum albumin
C	Cytosine
CaM	calmodulin
CdFYP	chaperone-dependent yellow fluorescent protein
cDNA	complementary deoxyribonucleic acid
cfu	colony forming units
CHOP	C/EBP homologous protein
cm	centimetre
cm^3	cubic centimetres
CX	cyclohexamide
CO_2	carbon dioxide
DAPI	4'-6-Diamidino-2-phenylindole
DBD	DNA-binding domain
ddH_2O	double-distilled dihydrogen monoxide/water
DIDMOAD	diabetes insipidus, diabetes mellitus, optic atrophy and deafness
DMEM	Dulbecco's modified eagle media
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphate
ECACC	European Collection of Cell Cultures
EDTA	ethylenediaminetetraacetic acid
eIF2	eukaryotic initiation factor 2
ER	endoplasmic reticulum
ERAD	ER-associated protein degradation
ERIS	endoplasmic reticulum intermembrane small protein
ERO1α	endoplasmic reticulum oxidoreductin 1

EtBr	ethidium bromide
eIF2	eukaryotic initiation factor 2
FACS	fluorescence-activated cell sorting
FRET	fluorescence resonance energy transfer
g	grams
G	guanine
GADD	growth arrest and DNA-damage-inducible
GDM	gestational diabetes mellitus
GWAS	genome wide association studies
HRD1	HMG-CoA reductase degradation protein 1
HRP	horseradish peroxidise
HSP90	heat shock protein 90
H_2O	dihydrogen monoxide/water
ICC	Immunocytochemistry
Ig	immunoglobulin
IHC	Immunohistochemistry
IMAGE	Integrated Molecular Analysis of Genomes and their Expression
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRE1α	inositol requiring enzyme 1α
JNK	c-Jun N-terminal kinase
kb	kilobase
kDa	kilo dalton
LB	Luria-Bertani
LCM	laser-capture microdissection
MAPK	mitogen-activated protein kinase
MCS	multiple cloning sites
ml	millilitre
mm	millimetre
mM	millimolar
MODY	maturity onset diabetes of the young
ng	nanograms
NO	nitric oxide
$P58^{IPK}$	P58 inhibitory protein kinase
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDI	protein disulphide isomerise
PERK	PKR-like ER kinase
pM	Picomolar
PPI	protein-protein interactions
QDO	quadruple drop-out
RIDD	regulated IRE-1 dependent decay
RNA	ribonucleic acid
rpm	revolutions per minute
SDM	spray dried barley malt extract
SDS-PAGE	sodium dodecyl sulfate polyacrylamide electrophoresis

SERCA	sarco-endoplasmic reticulum Ca^{2+}-ATPase
SNP	single nucleotide polymorphism
SOC	super optimal broth
SOD	superoxide dismutase
T	thymine
TBE	tris/borate/EDTA
TEMED	tetramethylethylenediamine
Tm	melting temperature
TRAF2	TNF-receptor-associated 2
U	units
UAS	upstream activating sequences
UPR	unfolded protein response
UV	ultraviolet
WF	wolframin
WFS1	Wolfram syndrome 1 gene
wfs1	Wolfram syndrome 1 gene mouse homolog
WFS2	Wolfram syndrome 2 gene
WS	Wolfram syndrome
XBP1	X-box-binding protein 1
YPD	yeast-extract, peptone, dextrose
µl	microlitre

CHAPTER 1

INTRODUCTION

1.0.0 Diabetes mellitus

Diabetes mellitus is a metabolic disorder, characterised by intolerance to glucose. The hormone insulin is required to maintain blood glucose homeostasis, and diabetes mellitus can be due to defects in insulin secretion or insulin action.

Acute hyperglycaemic episodes are a medical emergency. The consequences of excessive blood glucose include an osmotic diuresis, which may cause severe dehydration and electrolyte imbalance. In addition, the lack of glucose entry into cells means that they have to rely on alternative energy sources derived from fatty acid metabolism. This is a relatively inefficient process and leads to the production of toxic by-products, ketones. Without treatment a state of diabetic 'ketoacidosis' can result in loss of consciousness, renal failure and ultimately death.

Chronic hyperglycaemia is associated with long-term micro and macro vascular complications (Fowler, 2008). These include; nephropathy, retinopathy and peripheral/autonomic neuropathies, that can lead to permanent organ damage or failure and cardiovascular disease, ending in disability and premature mortality.

1.0.1 Prevalence and burden of diabetes mellitus

Diabetes is the fifth largest cause of death worldwide, affecting an estimated 285 million people; a figure that is projected to double by 2030 (Diabetes UK, 2010). The vast increase in the number of people with diabetes is particularly pronounced in developing countries, and is correlated with population growth, urbanisation, aging and increasing levels of obesity and physical inactivity (Wild, 2004).

In the UK it is estimated that 1 in 20 people have diabetes, many of whom are undiagnosed. All age groups are at risk, but prevalence is more than two times higher in deprived areas. People with diabetes are twice as likely to be admitted to hospital, experience prolonged hospital stays, and many incur social services costs for nursing care and home help. As a consequence, the annual cost of diabetes related health care is approximately 10% of the NHS yearly budget (~£9 billion), which is close to £1 million an hour (Diabetes UK, 2010).

1.1.0 Aetiology of diabetes mellitus

The aetiology of diabetes mellitus is not fully understood, but attributed to a combination of genetic and environmental factors. Since several pathogenic processes are involved in the development of impaired glucose regulation and diabetes mellitus, a system of etiological classification exists (fig. 1.0), which has been recently revised by the American Diabetes Association (ADA, 2010).

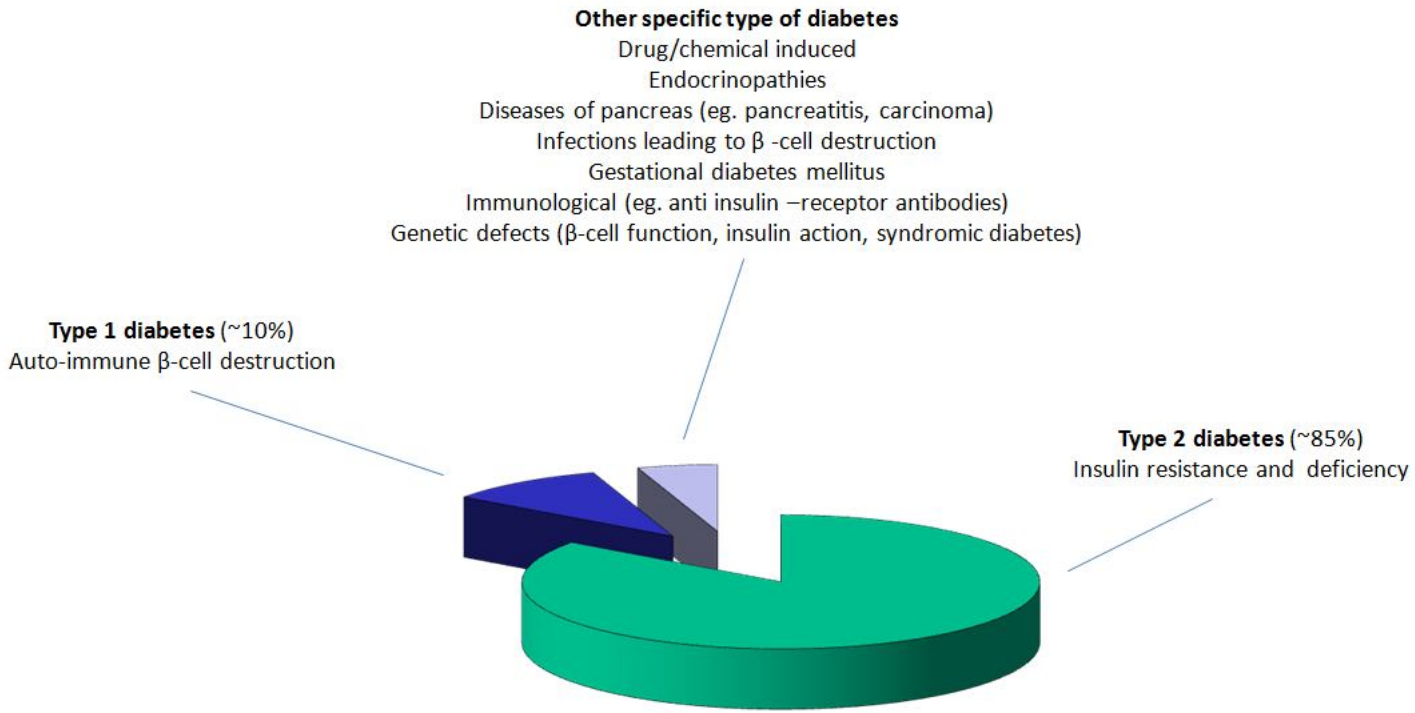


Figure 1.0. **Etiological classification of diabetes**.

Adapted from American Diabetes Association (2010).

1.1.1 Type 1 diabetes mellitus

Up to 10% of people affected by diabetes have type 1 diabetes mellitus (ADA, 2010), which results from autoimmune destruction of insulin secreting pancreatic β-cells. This occurs at variable rates, eventually leading to absolute insulin deficiency. Onset is generally in childhood, although latent autoimmune diabetes is known to present in approximately 10% of adults first diagnosed with type 2 diabetes mellitus (Palmer, 2005). Patients rely on insulin replacement therapy for survival, but some individuals (usually late-onset) retain limited β-cell function, sufficient to prevent ketoacidosis for a variable duration (Hosszufalusi, 2003).

The pathophysiology of type 1 diabetes is related to a permissive genetic background and environmental factors that are still poorly defined. It is thought that genetic factors contribute 30-70% of disease-factors (Pocoit, 2010). Half of the genetic risk of type 1 diabetes is conferred by Human Leukocyte Antigen (HLA) genes found on chromosome 6p (Steck, 2011) and the remainder of susceptibility genes are found in over 50 regions spread throughout the genome with smaller combined effects (Espino-Paisan, 2011).

1.1.2 Type 2 diabetes mellitus

Type 2 diabetes mellitus accounts for the vast majority of diabetes patients (~85%), and involves the combined effects of peripheral insulin resistance and impaired insulin secretion. Onset, usually in adulthood, tends to be a gradual development of

hyperglycaemia that may be asymptomatic in the early stages. At least initially, patients survive without pharmacologic treatment and can manage their blood glucose levels with diet and exercise modifications. The insulin resistance can also be improved with weight loss and/or pharmacological treatment, but insulin levels are rarely restored to normal values (ADA, 2010). Data collected by The Department of Health and Human Services suggest that ~80% of adult diabetics in the United States require some form of medication for diabetes management (CDC, 2010).

Causes of type 2 diabetes are complex and multifactorial. The majority of type 2 diabetics are obese, or carry a disproportionately large percentage of abdominal body fat, which contributes to insulin resistance in liver and muscle tissue. Studies have shown that visceral (abdominal) fat cells secrete excess amounts of free fatty acids and other biochemicals that interfere with normal glucose metabolism (Gastaldelli, 2008). Risk of developing this type of diabetes is also increased with age, lack of exercise, and also in people with hypertension or dyslipidemia. Individuals from particular ethnic subgroups are affected with higher frequencies than others (Zimmet, 2001), and there is also evidence confirming genetic predisposition (Voight, 2010).

1.1.3 Other types of diabetes

There are many categories of diabetes that make up the remaining 5% or so of cases. These include; drug and chemical induced diabetes, endocrinopathies, pancreatic

diseases such as pancreatitis and pancreatic carcinoma, infections associated with β-cell destruction, and gestational diabetes mellitus (GDM).

Rare cases of immunological disease can cause diabetes when antibodies against the insulin receptor either: competitively inhibit the action of insulin at target tissue, resulting in hyperglycaemia, or act as an insulin agonist leading to hypoglycaemia. Several conditions, of varying severity, are also derived from mutations in the insulin receptor protein, such as Rabson-Mendenhall syndrome and Leprechaunism (ADA, 2004).

A large number of other genetic defects may contribute to diabetogenesis including chromosomal abnormalities, and single and multiple gene mutations. For example, maturity-onset diabetes of the young (MODY), characterised by non auto-immune juvenile onset hyperglycaemia, is related to autosomal dominant mutations in one of at least six different genetic loci, which cause various defects of the insulin secretion pathway (Fajans, 2010). There are also genetic syndromes associated with diabetes. These include Wolfram syndrome, Prader-Willi syndrome, diabetes and deafness, and thiamine-responsive megaloblastic anemia, all differing in pathophysiology and etiological background (ADA, 2004).

1.2.0 Diabetes prevention and treatment

There is currently no cure for common diabetes although a range of drug and insulin replacement therapies are offered as a management strategy for many patients,

each with associated side-effects and individual limitations. An understanding of β-cell pathways is crucial for the development of better treatments.

Transplanting isolated islet cells, a procedure still in its infancy, is a current focal area of diabetes research efforts, and advances in genetic engineering could address the anticipated problem of 'donor' islet cell shortage in the future. Recent *in vivo* studies have also provided promising evidence that insulin secreting islet β-cells have the potential to regenerate in adults after near-total loss, by spontaneous reprogramming of endocrine α-cells (Thorel, 2010), or indeed may be induced by a combination of specific transcription factors to reprogramme from exocrine cells to β-cells (Zhou, 2008). These findings highlight the scope for regenerative treatments in patients with β-cell depletion.

A greater understanding of the genetic basis and molecular pathogenesis of diabetes would enable development of superior therapeutics and ultimately future preventative treatments. Modern technologies advanced since the completion of the Human Genome Project, not only enable us to examine disease from a single gene perspective, but to perform comparative assays on large-scale data sets relatively efficiently. In the field of diabetes research, these types of studies are particularly useful for identifying common genetic variants that may contribute to disease susceptibility. To date, collaborative genome wide association studies (GWAS) have enabled the identification of more than 35 genetic loci associated with type 2 diabetes in European populations (Voight, 2010). Single nucleotide polymorphism (SNP) chips allow analysis of over two thirds of common variation in the human

genome. The genes/regions so far identified (fig. 1.1) represent only a modest portion (~10%) of the familial risk underlying type 2 diabetes, and many more SNPs and copy number variants affecting susceptibility have yet to be discovered. After fine-mapping regions of interest, the next step is to explore the relationships between genetic variants and disease pathogenesis. It is anticipated that this kind of research could, in the future, be applied in a clinical setting to allow more accurate prognoses based on individual patient genotypes, where risk alleles can be considered for 'tailor-made' therapies (Aneesh, 2009; Xie, 2005).

Study of rare monogenic disorders and exceptional familial forms of disease traits has provided much of our current knowledge about complex diseases. The underlying molecular basis of over 1700 Mendelian disorders have been characterised (Peltonen, 2006) and these findings have helped to define critical pathways involved in many common 'multifactorial' diseases. In the case of diabetes research, examining monogenic disorders like MODY and Wolfram syndrome, not only directly benefits affected patients by way of specific therapeutic targeting, but also provides invaluable insights to the physiology of the β-cell, and this knowledge can be applied to understanding disease pathways in more prevalent types of diabetes. Since common variants of the Wolfram syndrome-causative gene (*WFS1*) correlate with type 2 diabetes risk (fig. 1.1), studying this rare disease may also have relevance to the pathogenesis of type 2 diabetes.

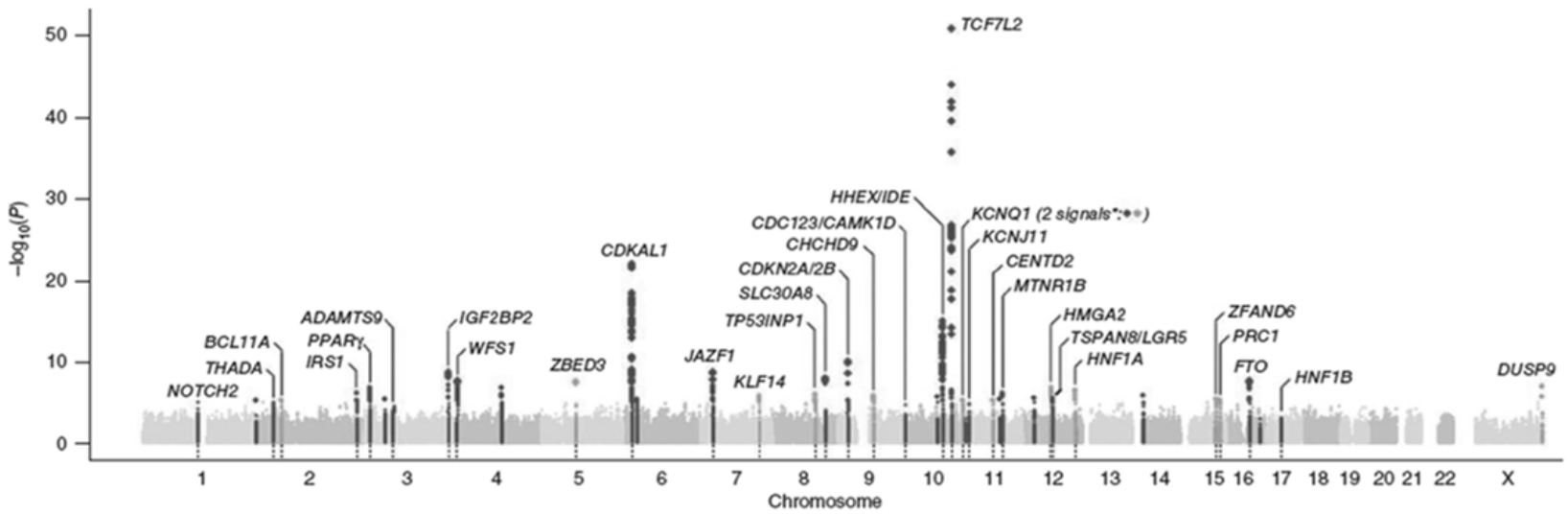


Figure 1.1. Genetic loci associated with type 2 diabetes predisposition. Manhattan plots summarizing meta-analysis results from a number of genome-wide association studies. Gene names are denoted on the basis of their proximity to the index SNP and are not presumed to indicate causality.

Adapted from Voight, B.F., *et al* (2010), *Nature Genetics*, **42**(7): 579-590.

1.3.0 Wolfram syndrome

Wolfram syndrome (WS), first described in 1938 (Wolfram, 1938), is a rare, autosomal recessive, neurodegenerative disease characterised primarily by juvenile onset diabetes mellitus and optic atrophy. A typical age of diabetes mellitus onset is 5.4±3.8 years (Rohayem, 2011), and the pathophysiology is attributed to impaired β-cell function and selective non-autoimmune β-cell death. A spectrum of associated complications also commonly present in patients throughout progression of disease, including diabetes insipidus, sensorineural deafness, renal and urinary tract abnormalities and neurological manifestations in the later stages (fig. 1.2) (Smith, 2004). In relation to these features, the syndrome has also been described as DIDMOAD (diabetes insipidus, diabetes mellitus, optic atrophy, and deafness).

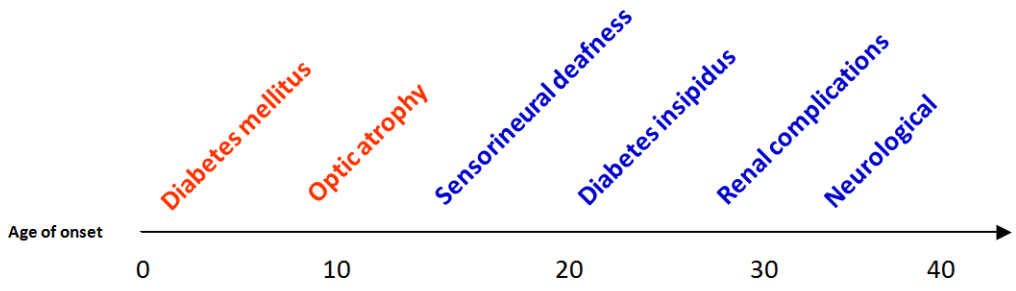


Figure 1.2. Natural history of Wolfram syndrome (WS). WS is characterised by juvenile onset diabetes mellitus and optic atrophy. All or some of a range of other clinical features may present as disease progresses with typical ages of onset indicated. The average life expectancy for a WS patient is less than 40 years with death commonly resulting from central respiratory failure due to brain stem atrophy.

The prevalence of WS in the UK was estimated at one in 770,000 with a carrier frequency of 1 in 354, by a nationwide study conducted in 1995 (Barrett, 1995), although this figure is likely to have risen over the last fifteen years as more cases of WS have been reported. Regions such as North Africa and the Middle East, where consanguineous marriages are more common, are likely to have a higher prevalence of WS, and possibly a larger number of unreported cases. The average life expectancy for a WS patient is less than 40 years, with death often resulting from central respiratory failure due to brain stem atrophy. For those patients in 'developing' countries, the life expectancy is probably shorter still, as a consequence of poorer health care.

1.3.1 Genetics of Wolfram syndrome

Since WS shares common phenotypic characteristics with diseases of oxidative phosphorylation such as Leber's hereditary optic neuropathy, and maternally inherited diabetes and deafness, it was thought that disease pathogenesis could be related to defective mitochondrial function. In support of this theory, there have been deletions in mitochondrial DNA found in some WS patients (Rötig, 1993; Barrientos, 1996a; Barrientos 1996b, Gómez-Zaera, 2001, Mezghani, 2011), although there is no evidence to suggest that WS is related to systemic abnormality of respiratory chain function (Jackson, 1994).

The majority of WS cases are attributed to mutations in the *WFS1* gene, located at chromosome 4p16, which was mapped to the region firstly by linkage analysis (Polymeropoulos, 1994; Collier, 1996), then later identified by positional cloning in 1998 (Inoue, 1998; Strom, 1998). More recently, a second locus (4q24) containing a gene designated *WFS2*, was found to be associated with a small number of affected Jordanian families (Amr, 2007). These patients presented with defective platelet aggregation with collagen, and bleeding tendency in addition to typical WS clinical features. Some individuals also had peptic ulcer disease and gastrointestinal bleeding but none presented with diabetes insipidus.

1.3.2 *WFS1*

The *WSF1* gene contains 8 exons, the first of which is non-coding, and a wide spectrum of mutations have been identified throughout the gene sequence (Ensembl, 2011). No significant mutational hotspots have been determined although the majority have been found in exon 8, which is the largest, representing >70% of the total transcript (Ensembl, 2011). Previous mutation analysis studies have been unable to identify clear genotype-phenotype correlations in WS cohorts, although genetic heterogeneity and small numbers of patients have made analyses difficult. It has been postulated that nonsense or insertion/deletion mutations resulting in protein truncation may induce a more severe phenotype due to complete loss of function (Smith, 2004; Cano, 2007). Possible correlations between location of mutations and neurological manifestations have also recently been made (Chaussenot, 2011).

Some heterozygous *WFS1* variants are associated with disease states distinct from WS. In addition to the above-mentioned links to type 2 diabetes, some *WFS1*-variants may be associated with psychiatric illness (Furlong, 1999; Torres, 2001; Swift, 2005; Zalsman, 2009) and others with non-syndromic hearing loss (Young, 2001; Bespalova, 2001). To date, more than 30 *WFS1* mutations are associated with cases of autosomal dominant, low frequency deafness, also known as DFNA6 (Hildebrand, 2008).

1.3.3 Wolframin

The *WFS1* gene codes for an 890 amino acid protein named wolframin (WF). Analysis of *WFS1* mRNA demonstrates ubiquitous expression, including brain, lungs, heart, kidneys and pancreas, correlating with many of the organs affected during disease progression (Strom 1998; Inoue 1998).

In the pancreas, immunohistochemical staining of mouse tissue has shown that β-cells are primary sites of wolframin expression (Ueda 2005), corresponding with WS pancreas pathology, where β-cells are selectively lost from patient islets.

Studies in rat brain revealed expression in selected neurons, particularly prominent in regions such as; hippocampus, amygdala, olfactory tubercle and allocortex (Takeda, 2001), which are components of, or closely related to, the limbic system. These findings would suggest a possible role for wolframin in emotional and behavioural control. This is consistent with the psychiatric abnormalities seen in

some WS patients. Other brain regions of *WFS1* expression include the cerebellum and brainstem, and affected patients may have severe symptoms related to these areas (cerebellar ataxia, central respiratory sleep apnoeas). In a study of the normal mouse visual system, *WFS1* expression was shown to be widely distributed; present in occipital cortex, retina and optic nerve (Kawano, 2005), although notably only moderate expression in the optic nerve, which is most susceptible to atrophy in WS.

Immunofluorescence and cell fractionation studies have demonstrated the localisation of wolframin primarily to the endoplasmic reticulum (ER) membrane (Takeda, 2001), which it spans with nine putative transmembrane domains. The amino (N) terminal portion is positioned on the cytosolic side of the membrane, and the carboxy (C) terminus lies within the ER lumen (fig. 1.3). More recent immunofluorescent and electron microscopy evidence has also shown localisation of wolframin to secretory granules in pancreatic β-cells (Hatanaka, 2011).

The wolframin protein has a molecular mass of ~100kDa, but has been shown experimentally to organise into higher molecular weight complexes of ~200 and 400kDa (Hoffman, 2003) in both transfected and endogenously expressing cell lines. These could represent combination with other, as yet unknown proteins, but more likely homo-dimeric and tetrameric assemblies.

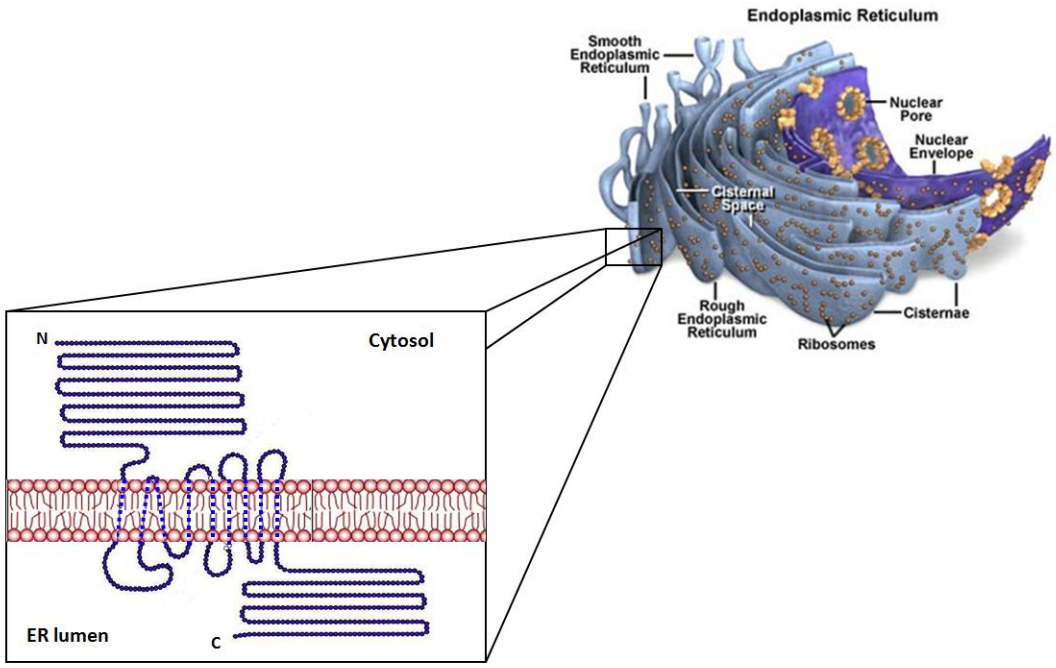


Figure 1.3. Hypothetical structure of the wolframin protein. Wolframin primarily localises to the endoplasmic reticulum (ER), spanning the membrane with nine putative transmembrane domains; Amino acids 1-310 comprise the cytosolic amino-terminal (N) domain and 653-890 the ER lumenal carboxy-domain (C).

Image adapted from Davidson, 2010.

Pulse-chase experiments suggest that mature wolframin protein is relatively stable, with a half-life estimated at ~2 days (Hoffman 2003). N-linked glycosylation has been the only post-translational modification demonstrated, and following treatment with tunicamycin (glycosylation inhibitor) appears to confer stability and/or allows proper wolframin protein folding (Hoffman 2003; Yamaguchi 2004).

1.4.0 Wolframin function

The sub-cellular localisation and potential dimerisation of the wolframin protein points towards functional roles as a protein chaperone and/or in membrane trafficking. Indeed, functional studies have implicated wolframin in calcium

homeostasis, insulin biosynthesis and secretion (fig. 1.4), and endoplasmic reticulum stress pathways.

1.4.1 Wolframin, insulin secretion and Ca^{2+} homeostasis

Mouse models lacking functional wolframin have been generated in order to examine the pathophysiology of WS. Over time, mice with disrupted *wfs1* (*WFS1* homolog) were reported to develop either glucose intolerance or overt diabetes depending on their genetic background (Ishihara, 2004, Riggs, 2005). Following oral glucose tolerance tests, the mutant mice demonstrated significantly higher blood glucose levels compared to control mice, indicating that disruption of the *wfs1* gene induces impaired glucose homeostasis and is sufficient to cause a diabetes mellitus phenotype on a permissive genetic background. This observation is thought to be linked predominantly with insulin secretion defects rather than insulin resistance, as hormone responses were markedly blunted in *wfs*1 deficient animals.

The impaired insulin secretion in *wfs1* deficient mice was shown to be correlated with both a reduction in stimulus-secretion coupling, and an overall reduction in β-cell mass and insulin content. Compared with wild-type mouse cells, glucose-stimulated insulin secretion has been shown to be reduced by ~25% in islets from mutant mice together with a 36% reduction in cytosolic Ca^{2+} flux (Ishihara, 2004). Normal insulin secretion could be restored following *wfs1* re-expression using a recombinant adenoviral vector, and significantly enhanced (~50%) by over-expression of *wfs1.* Only half of the total number of islet cells could be extracted

from mutant mouse pancreata compared to wild-type, and the loss of β-cells is considered a consequence of apoptosis, due to glucose toxicity and endoplasmic reticulum stress (section 1.5.0).

The role *WFS1* plays in insulin secretion could be related to calcium homeostasis. Ectopic expression of *WFS1* in *Xenopus* oocytes, demonstrated that *WFS1* has a modulatory effect on intracellular Ca^{2+} levels (Osman, 2003) and combined study of *WFS1* over-expressing and knock-down HEK293 cells showed that wolframin positively regulates the Ca^{2+} filling state of the ER (Takei, 2006). Experimental evidence generated from *wfs1*-/- islets and depleted β-cells also shows that decreased glucose-stimulated insulin secretion is partially due to impaired cytosolic calcium rise and a defective ATP rise in response to glucose (Ishihara, 2004; Zatyka, 2008b).

More recently calmodulin (CaM), an important second messenger of intracellular calcium signalling, was described as a molecular partner of wolframin (Yurimoto, 2009). It has been postulated that the wolframin protein may form an ion channel, similar to other ER membrane-associated proteins, commonly found in tetrameric complexes, although it remains unclear as to whether wolframin acts directly as a channel or functions only in a regulatory capacity.

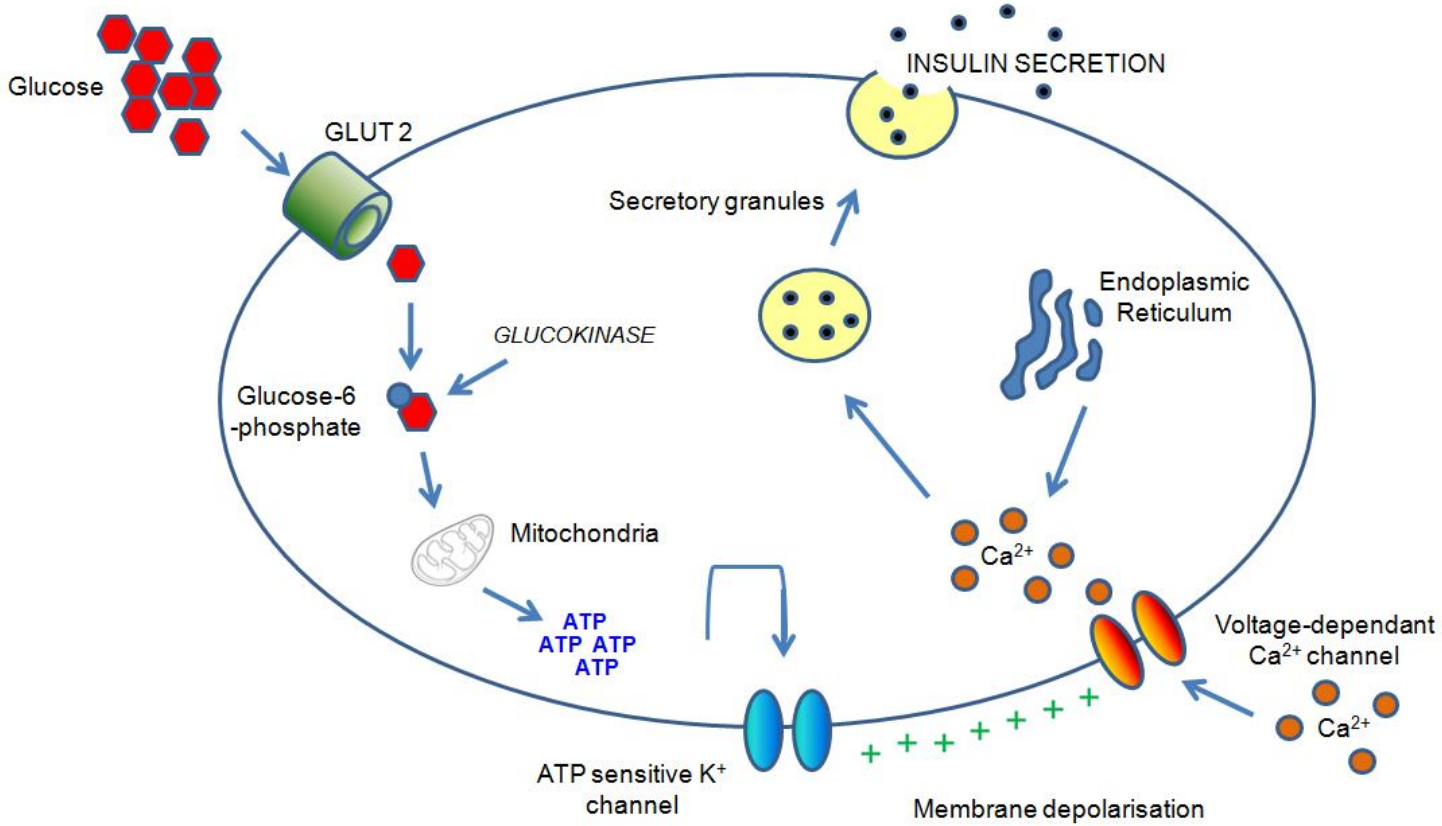


Figure 1.4. Glucose-stimulated insulin secretion. Insulin secretion in β-cells is triggered by rising blood glucose levels. Glucose enters cells via the GLUT2 transporter and the glycolytic phosphorylation of glucose causes a rise in the ATP:ADP ratio. ATP-dependent potassium channels (K^+) are activated by this change causing membrane depolarisation and inflow of Ca^{2+} ions through voltage-gated ion channels. Increases in cellular Ca^{2+} activate release of more Ca^{2+} from endoplasmic reticulum stores, which further raises the cytoplasmic concentration of Ca^{2+}, sufficient to induce release of insulin from secretory vesicles.

Adapted from: Ren, J., *et al* (2007).

1.4.2 Wolframin and cell proliferation

Alternative studies of wolframin depleted cells suggest that the loss of β-cells seen in WS patients could also be linked to reduced rates of cell proliferation. A study of a stably wolframin depleted rat β-cell line (BRIN-BD11), demonstrated dramatic reduction in cell proliferation rates and changes in morphology compared to normal controls (McBain, 2003). Analysis of cellular DNA profiles revealed no evidence of DNA fragmentation suggestive of apoptosis, but instead, elevated levels of diploid DNA and a reduction of the G2/M peak associated with DNA replication and cell division. These results suggest that reduced wolframin expression *in vivo* may negatively affect β-cell replacement during normal cell turn-over, leading to gradual depletion of β-cell mass over time. In concordance with these findings, expression profile data generated from patient and control fibroblasts (Philbrook, 2005) revealed a 2.1 fold up-regulation of fibulin-3 in patient cells, a gene previously shown to be associated with cellular senescence (Lecka-Czernik, 1995).

1.5.0 Endoplasmic reticulum stress

The endoplasmic reticulum (ER) is the subcellular organelle involved in the synthesis and folding of secretory, membrane, and some organelle proteins and lipids, termed ER 'clients'. Carefully balanced conditions including ATP, Ca^{2+} and redox state are required for optimal client protein folding and thus the ER is highly sensitive to environmental changes. Any perturbation that compromises the folding capacity of the ER is termed ER stress, and may result in accumulation and aggregation of

unfolded proteins. Several neurodegenerative diseases as well as diabetes and ischemia (Lindholm, 2006) have been associated with ER stress.

The triggers for ER stress range from acute cellular insults such as hypoxia or glucose deprivation, to chronic stressors like viral infections or genetic mutations that cause permanent changes in cell function. Some types of cells, including neuroendocrine cells are by nature, subject to persistently elevated levels of ER stress due to their specialism in protein synthesis and secretion, and high ER client load.

An increase of ER volume is observed in cells under ER stress together with activation of complex signalling cascades, termed the ER stress response or unfolded protein response (UPR). The outcome of the stress response can be two-fold, inducing both adaptive measures to maintain ER homeostasis, such as expansion of the ER size and upregulation of protective genes, and apoptosis in cases of prolonged stress or irreversible damage.

1.5.1 ER stress sensors

Three transmembrane receptors known to mediate the ER stress response are; PERK (PKR-like ER kinase), IRE1α (inositol requiring enzyme 1α), and ATF6α (activating transcription factor 6α). It is understood that the ER luminal domains of all three receptors are bound by an ER chaperone BiP (immunoglobulin binding protein/GRP78), which maintains them in an inactive state. This coupling is recognised as one of potentially several mechanisms controlling the activity of ER stress sensors (Oikawa, 2007). Following accumulation of unfolded proteins in the ER

lumen, BiP dissociates from the receptors, allowing activation of downstream targets (fig 1.5) that collectively constitute an initial pro-survival response. This response may 'switch' to pro-apoptosis in situations where stress cannot be resolved, although the precise mechanisms facilitating this 'switch' remain to be elucidated (Rutkowski, 2007).

The proposed pro-survival responses include reversible measures to immediately relieve the protein folding load, by inhibiting the translocation of many ER substrates (Kang, 2006), degradation of some ER-associated mRNAs to assist in membrane clearance, and general inhibition of ER processing pathways. More long-term coping strategies involve the upregulation of genes relating to protein folding and processing, thus increasing ER capacity for handling client proteins (Schröder, 2005). Adaptive responses to ER stress are also coupled with transcriptional activation of processes not specifically ER-dependent, such as regulation of energy production, amino acid metabolism, and cellular redox state (Rutkowski, 2007).

Following dissociation of BiP, PERK homodimerises and autophosphorylates. This 'active' PERK has been shown to attenuate general protein translation, via phosphorylation of eIF2 (eukaryotic initiation factor 2), whilst concomitantly increasing transcription of selected target genes such as *ATF4* (activating transcription factor 4). These genes link to downstream pathways involving growth arrest and DNA damage, apoptosis, redox balance, and amino acid synthesis (Harding, 2000).

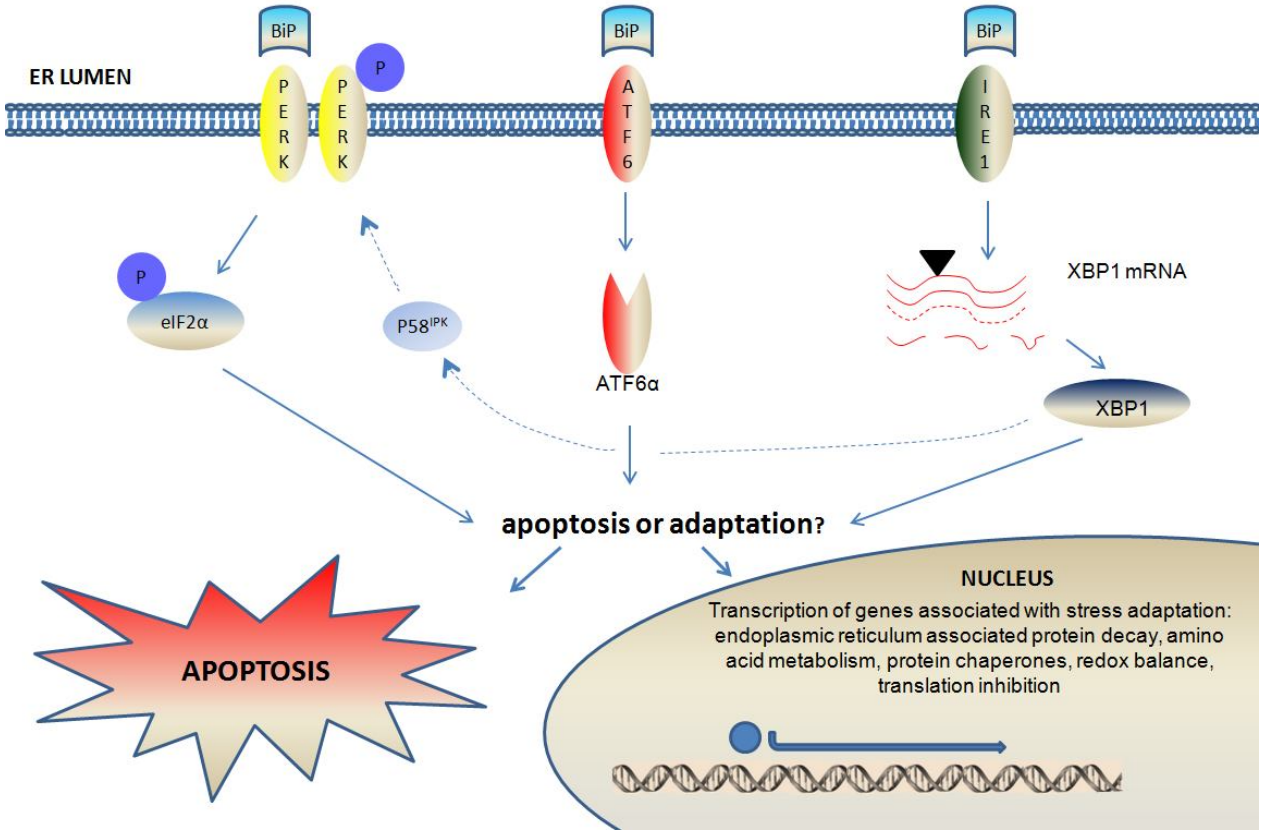


Figure 1.5. Endoplasmic Reticulum (ER) stress pathways. ER transmembrane stress sensors PKR-like ER kinase (PERK), activating transcription factor 6α (ATF6α) and inositol requiring enzyme 1α (IRE1α) activate various arms of the unfolded protein response (UPR). The outcome of this response may be adaptive enabling the cell to more efficiently cope with a high client protein load, or apoptosis, when stress cannot be resolved. The mechanisms causing a switch between survival and death decisions are not yet fully understood (Rutkowski, 2007).

IRE1α has downstream endonuclease activity, which results in cleavage of a 26-nucleotide intron from XBP1 (X-box-binding protein 1) mRNA. The XBP1 splice variant (sXBP1) translates to a stable, active transcription factor, which targets several genes encoding folding chaperones such as PDI (protein disulphide isomerise) and BiP/Grp78 (also a Ca^{2+} binding chaperone), and others involved in ER-associated protein degradation (ERAD). Another protein activated by sXBP1 is $P58^{IPK}$ (P58 inhibitory protein kinase), which binds and inhibits PERK. This has been proposed as a negative feedback loop that may signify the termination of the ER stress response, since upregulation of $P58^{IPK}$ occurs several hours after PERK activation and translation-block (Yan, 2002). Recent reports show that other mRNA substrates are targeted for degradation by IRE1α (Hollien, 2009) through the regulated IRE-1 dependent decay (RIDD) signalling pathway, an additional measure to alleviate the ER burden.

The third sensor ATF6, is reduced following initiation of ER stress and subsequently translocated to the golgi apparatus, where it is cleaved and activated by site-1 and site-2 proteases (Chen, 2002). Active ATF6 is a basic leucine zipper (bZIP) transcription factor. Like XBP1, ATF6 stimulates expression of protein folding and ERAD-associated genes, thus promoting clearance of proteins from the ER lumen (Kokame, 2001). ATF6 also has the ability to regulate XBP1 mRNA expression and interact directly with the XBP1 protein. This cross-talk may have evolved due to redundancies in the stress response pathways, or may in fact represent adaptive features that enhance the efficiency and feedback of the system (Wu, 2007).

1.5.2 ER stress and apoptosis

ER stress mediated apoptosis is thought to occur via a number of pathways; some linking to general apoptotic signalling, others seemingly more specific to ER stress. All three arms of the ER stress response have been linked with proapoptotic signalling, so in addition to adaptive gene upregulation, the ER stress sensors are able to initiate the cell death machinery.

A well characterised end-point of programmed cell death is the activation of effector caspases. These proteins are cysteine proteases, present within cells as inactive zymogens, which are cleaved to active enzymes following the induction of apoptosis. During programmed cell death caspases are responsible for cleaving critical cellular substrates such as chromosomal DNA, and the sequential dismantling of structural nucleic proteins, typical of apoptosis.

Many pathways may converge upstream of this stage, involving sub-sets of initiator caspases, which have yet to be wholly ascertained in terms of ER stress signalling. Some studies of caspase-12 -/- mouse embryonic fibroblasts have demonstrated a partial resistance to ER stress induced apoptosis (Szegezdi, 2003), suggesting this may play a principal role in death decisions. When translating these findings to human systems, it has been proposed by some that caspase-4 could perform a similar function, although recent work on knock-out models (Saleh, 2006), produced contradictory results, leaving the role of caspase-12 in ER stress mediated apoptosis a topic of current debate.

Other works using over-expression and knockout mouse models have identified potent upstream components of the pro-apoptotic response such as CHOP (C/EBP homologous protein), ASK1 (apoptosis signal-regulating kinase), and members of the Bcl2 (B-cell lymphoma 2) protein family, indicating signal relay between the ER and mitochondria (Rutkowski, 2007).

CHOP activation under conditions of ER stress is linked to all three of the stress sensors, although upregulation of CHOP protein expression is essentially carried out via PERK induced ATF4. The importance of CHOP signalling has been confirmed in knock-out mouse models (CHOP-/-), which show a partial resistance to ER stress induced apoptosis (Oyadomari, 2002). Genes involved in cell death pathways are promoted by CHOP expression and numerous downstream targets have been identified including; BCL2, GADD (growth arrest and DNA-damage-inducible) 134 and 38, and ERO1α (endoplasmic reticulum oxidoreductin 1).

Notably, active IRE-1 can bind to TRAF2 (TNF-receptor-associated 2), which in turn is thought to recruit ASK1 (apoptosis-signal-regulating 1). This complex is able to communicate with downstream kinases such as p38 MAPK (mitogen-activated protein kinase) and JNK (c-Jun N-terminal kinase), which are known to initiate apoptosis (Urano, 2000). Further evidence of pathway convergence comes from the post-translational regulation of CHOP by p38 MAPK, linking the predominantly PERK activated CHOP upregulation with IRE1 induced p38. It is thought that this mechanism may serve to perpetuate *CHOP*-mediated apoptotic cascades in the late

stages of severe or prolonged ER stress, since the IRE1 arm of the UPR is considered the last to be activated.

Despite evidence linking the various UPR pathways, a fundamental question that still remains unanswered is whether these multiple stress response pathways are all equally necessary and sufficient for making the transition from prosurvival to proapoptosis, or is there a key event that is paramount in executing cell death decisions? Also unclear, is to what extent the switch from adaptation to programmed cell death involves down-regulation of protective genes, and upregulation of suicide genes, and when does a cell under ER stress reach a point of no return?

1.5.3 ER stress and diabetes

Links between ER stress and diabetes pathology have been directly demonstrated in studies of human monogenic and type 2 diabetes with supporting evidence from a number of rodent diabetes models. Some tentative connections have also been made between immune-mediated type 1 diabetes and the UPR in sealing the fate of apoptotic β-cells.

Autopsy data from type 2 diabetic patients reveals the decline in insulin secretion can be attributed to a decrease in β-cell mass following apoptosis (Butler, 2003). Histological staining of human pancreas sections taken from type 2 diabetics, illustrated markedly higher levels of ER stress markers, such as CHOP and $P58^{IPK}$ in

islet cells (Laybutt, 2007), as well as a two-fold increase in ER size when compared to normal controls (Marchetti, 2007), indicating UPR pathways are likely to be engaged in the execution of apoptosis. The ER stress response may also explain the association between obesity and type 2 diabetes; Free fatty acids trigger ER stress signalling, cell dysfunction and death in rodent β-cells (Kharroubi, 2004; Karaskov, 2006; Cnop, 2007), suggesting lipotoxicity is an important factor in the health and survival of β-cells *in vivo*.

Involvement of ER stress in human type 1 diabetes has yet to be proven, although studies in animal models have shown that cytokine induced β-cell apoptosis does in fact link in with UPR signalling pathways. In rodent models of type 1 diabetes, islet β-cell destruction is known to be coupled with the infiltration of T cells (O'Brien 1997, Jörns 2005), and also with the expression of cytokines (IL-1β, interferon-γ, TNF-α) and inducible nitric oxide synthase (iNOS) (Jörns 2005, Uno, 2007), which may be early mediators of cell death. *In vitro*, IL-1 β or TNF-α combined with IFN-γ can induce nitric oxide (NO) production, severe functional repression and death of β-cells (Southern 1990, Eizirik 1992), and several well established UPR signalling molecules such as XBP1 and CHOP are upregulated by these cytokines (Pirot 2007, Cardozo 2001, Kutlu 2003, Cardozo 2005). NO action may also serve to deplete ER Ca^{2+} stores through inhibition of SERCA (sarco-endoplasmic reticulum Ca^{2+}-ATPase) expression (Cardozo 2005) and interaction with SNAP (synaptosomal-associated protein), a chemical donor of NO (Oyadomari 2001). This Ca^{2+} depletion likely compromises the protein folding capacity of ER resident chaperones, triggering the UPR.

1.5.4 Wolfram syndrome and ER stress

A lack of *wfs1* in mice is sufficient to cause the diabetes mellitus phenotype (Riggs, 2005) and accumulating evidence indicates that the pathogenesis of Wolfram syndrome involves chronic ER stress, which triggers apoptosis due to loss of wolframin function. Compared to wild-type cells, elevated levels of ER stress markers such as BiP and CHOP have been detected in cells derived from rodent *wfs1* knock-out (*wfs1-/-*) models (Riggs, 2005), and transcriptional up-regulation of *WFS1/wfs1* in both human and mouse cells is triggered by stress inducing agents such as thapsigargin, dithiothreitol and tunicamycin (Yamaguchi, 2004; Ueda, 2005; Fonseca, 2005). Under the electron microscope, further signs of ER stress in *wfs1-/-* cells are indicated by dramatic abnormalities in morphology, including reduced secretory granules and dilated ER (Riggs, 2005).

It seems that *WFS1* may be necessary for cell survival, particularly under conditions of ER stress. *Wfs1* deficient mouse insulinoma (MIN6) cells grown *in vitro* under standard conditions exhibit ~10% increase in apoptosis and display a 2.5-fold increase in cleaved caspase-3, a marker of apoptosis (Riggs, 2005). Furthermore, a 6-fold increase in *wfs1* protein expression can be detected in cells derived from Akita mice (Ueda, 2005), which are a model of MODY. These mice carry a C96Y mutation in the *Ins2* gene, and have progressive loss of β-cells, shown to be a result, at least in part, of ER stress-induced apoptosis.

Links have been uncovered between wolframin and all three of the known ER stress-sensing proteins. Firstly, in both IRE1α-/- and PERK-/- cells, the induction of wolframin protein expression normally triggered by ER stress is attenuated (Fonseca, 2005). These findings suggest that wolframin expression under stress conditions is regulated by IRE1α and PERK.

Furthermore, transcriptional activity of ATF6α is abated by *WFS1* expression, suggesting wolframin has a role in the negative regulation of a feedback loop during ER stress signalling (Fonseca, 2010). Previously shown to be anchored to the ER membrane by BiP, ATF6α is cleaved and released from the ER to upregulate target genes in the nucleus under ER stress conditions. An even stronger suppression of this activity by wolframin was recently demonstrated with a promoter reporter assay (Fonseca, 2010). Further to these findings, a direct protein-protein interaction was observed under non-stress conditions between ATF6α and wolframin in pancreatic and neuronal cell lines, which dissociated in a time-dependent manner following treatment with stress inducing agents (thapsigargin, dithiothreitol).

Experiments carried out in pancreatic MIN6 cells stably expressing shRNA against *WFS1*, lead to the conclusion that ATF6α is regulated by wolframin at the post-translational level (Fonseca, 2010). ATF6α mRNA levels remained unchanged in wolframin depleted MIN6 cells, whereas a dose-dependent increase in ATF6α protein and downstream target genes (*p58IPK*, *BiP*) were observed, which could be reversed upon reintroduction of *WFS1* with a lentiviral vector. The two other known

UPR master regulators, IRE1 and PERK, appeared unaffected by the reduction of *WFS1*.

WFS1 enhances ubiquitination of ATF6α and seems to have a role in proteasomal degradation of ATF6α, which can be rescued following proteosome inhibition with MG132 (Fonseca, 2010). Glycerol gradient fractionation of ER-isolated lysates resulted in comigration of the proteosome, ATF6α and wolframin, suggesting the possibility of complex formation between these three proteins. Furthermore, HRD1 (ER-resident E3 ligase) was also found to form a complex with these proteins and in *wfs1-/-* islets and patient cells has been shown to be upregulated by *WFS1* expression. The interaction between HRD1 and ATF6α was disrupted by ER stress conditions. Collectively, these results suggest that wolframin stabilises and enhances HRD1, which functions as an E3 ligase for ATF6α. ATF6α hyperactivation, likely to occur in Wolfram syndrome patients is probably involved in β-cell apoptosis, due to dysregulated ER stress signalling.

Contrary reports have attributed beneficial effects of ATF6α upregulation in the protection of neurones and cardiomyocytes from ER stress/ischemia-mediated apoptosis (Martindale, 2006; Kudo, 2008), and also to attenuate diet-induced obesity and insulin resistance (Ye, 2010). Additional evidence comes from ATF6α -/- mice, and knock-out hepatocytes, both of which are more susceptible to the effects of ER stress compared to their wild-type counterparts (Wu, 2007). It seems likely that the opposing roles for ATF6α in both cell survival and death under conditions of stress are related to the extent of the ER stress and play a key part in the switch from

adaptive to apoptotic decisions. Polymorphisms and haplotypes of ATF6α are associated with type-2 diabetes and impaired glucose regulation (Meex, 2007), further highlighting the importance and likely involvement of this gene in WS pathology, and overall development of diabetes.

1.6.0 *WFS2*

Genetic heterogeneity was first demonstrated in Wolfram syndrome by El-shanti and colleagues in 2000 (El-shanti, 2000) who showed linkage to chromosome 4q22-24 in four consanguineous Jordanian Bedouin families (16 patients). Further microsatellite marker analysis within the defined critical region, direct sequencing and SSCP analysis revealed a novel gene, designated *WFS2* (*ZCD2*), which carried a single homozygous nucleotide change (E37Q) in all affected individuals (Amr, 2007).

WFS2 is a 3 exon gene, coding for a small, 135 amino acid, transmembrane protein known as ERIS (endoplasmic reticulum intermembrane small protein), which is expressed in a variety of tissue-types. Immunocytochemical analyses demonstrated co-localisation of ERIS and calnexin, an ER marker, but no direct interaction could be detected when co-immunoprecipitated with wolframin (Amr, 2007).

WFS2, otherwise known as *CISD2* or *Miner1* belongs to an iron-sulphur domain gene family, which includes *CISD1* (*ZCD1*, *mitoNEET*) and *CISD3* (*Miner 2*) (Wiley, 2007). *WFS2* shares a unique highly conserved CDGSH consensus sequence with both

Miner2 and *mitoNEET*, which is thought to be important for iron-binding. The mitoNEET protein (*CISD1*) resides on the outer mitochondrial membrane and is a target protein for pioglitazone, an insulin sensitiser drug used to treat type 2 diabetes (Colca, 2004). It is thought that mitoNEET functions in some capacity in redox reactions and electron transport due to the reduction of oxidative capacity observed in mouse mitoNEET-null heart tissue (Wiley, 2007). Although the function of *WFS2* is not clearly defined, iron-containing proteins are known to frequently function as redox-sensitive molecules and dysregulation of iron metabolism is associated with diabetes (Fernandez-Real, 2004).

WFS2 deficiency appears to confer susceptibility to apoptosis, as demonstrated by the presence of pro-apoptotic markers CHOP and BAX (Bcl2-associated X protein) in knock-down rat pancreatic β-cells (Amr, 2010). Unlike findings from *WFS1* knock-down models, no activation of the UPR was detectable in *WFS2* depleted (INS-1) cells. Translational upregulation of antioxidant enzymes SOD1 and SOD2 (superoxide dismutase 1 and 2) together with a significant increase in global tyrosine nitration were detected however, suggesting the presence of oxidative stress. It is unclear whether this is a cause or consequence of *WFS2*-mediated pathogenesis.

Earlier murine knockout (*WFS2-/-*) studies indicated that *WFS2* is necessary for mitochondrial integrity, and reported co-localisation with mitochondrial markers (Chen, 2009). Deficiency of *WFS2* in mice caused autophagic cell death accompanied by mitochondrial breakdown and dysfunction, which are precedents to nerve and muscle degenerative pathways, classic phenotypic features of premature aging.

Notably, the same region of chromosome 4q in which *WFS2* is located was linked to longevity in humans (Puca, 2001). These observations support the previously disputed notion that WS is at least in part associated with mitochondrial defects, although the phenotypic effects of *WFS2*-/- in these model mice are somewhat distinct from that of WS in humans, so proposed links remain debatable.

1.7.0 Understanding Wolfram syndrome pathogenesis and the role of *WFS1* in type 2 diabetes

It is known that the majority of Wolfram syndrome patients carry mutations in *WFS1*, although how genetic alterations disrupt the function of the wolframin protein is unclear. Previous study of *WFS1* mutants and *WFS1*-null models has led to the belief that WS is caused by loss of *WFS1* function, which leads to ER-stress mediated apoptosis of pancreatic β-cells and other susceptible cell-types.

An interesting feature of WS is the variability of phenotype, which is as yet unexplained. Some individuals present with the full spectrum of associated clinical features, whereas others appear only to develop a select few. A number of individuals also appear to have a milder form of WS, with relatively late-onset of disease characteristics (Van de Ouweland, 2003; Cano, 2007; Gasparin, 2009).

Accepting unresolvable ER stress as the primary cause of disease pathogenesis, the question arises as to why clinical manifestations of disease present with delayed

onset in some individuals? Perhaps it requires a number of years for levels of chronic cellular stress to reach a critical threshold where the 'switch' to apoptosis is activated, beginning the process of irreversible tissue degeneration. Conversely, apoptotic pathways may indeed be activated from earlier stages of life, but go undetected until sufficient tissue damage prompts clinical detection. It is certainly possible that background genetics of WS patients may play a fundamental role in the severity of disease progression. Combined actions of other undefined proteins involved in wolframin functional pathways including the UPR, could play a significant role in the efficiency of the system as a whole. A poorer prognosis or more severe phenotype may then depend on the presence of *WFS1* mutation together with one, or a number of other risk alleles.

The mutation-type carried by an affected individual could also determine the relative severity of disease. Other studies postulate that mutation causes loss of wolframin function as a result of protein aggregation or instability leading to degradation of wolframin (Hofmann, 2003; 2006).

HYPOTHESIS

In light of the phenotypic variability in WS and the additional non-syndromic disorders linked to *WFS1*-variants, we hypothesise that a number of alternative wolframin-dependent pathways could be affected distinctly by different types of mutation, and this may reflect the spectrum of phenotypic features observed in WS patients and *WFS1*-related disorders.

The work presented in this thesis is aimed at exploring the hypothesis that some wolframin variants retain partial function. The basis for this proposal is drawn from a number of WS attributes, namely the heterogeneous complement of clinical features associated with *WFS1* mutations. The report of 'milder' phenotypes suggests that some wolframin variants have partial function. We would like to ascertain whether this may be the case, perhaps opening the door for new therapeutic intervention in these instances. Furthermore we hope to understand more about how the diabetes risk allele (*WFS1*-H611) contributes to disease pathogenesis, by establishing the foundations for functional experiments.

By characterising representative wolframin variants we intend to test aspects of wolframin function such as protein-interaction, ER stress response, insulin secretion and Ca^{2+} homeostasis and uncover any contrast in function, which might explain the phenotypic variation. It is anticipated that those wolframin variants relating to mildly affected individuals may display at least some residual function. Our aim is to expand current functional data about missense variants from the perspective of genotype-phenotype relations. To achieve this we are primarily asking:

Do some WFS1 mutations allow partial function of wolframin?

We will begin to address this question by firstly expanding the current *WFS1* mutational spectrum, to provide the largest single UK Wolfram syndrome patient cohort to date. The groundwork for functional study of *WFS1* genotype and phenotype relationships is then laid by characterisation of selected *WFS1*-variants

corresponding to phenotypes of variable severity. Finally, a number of potential novel wolframin molecular partners from the pancreas are described, together with the first evidence of *WFS1* expression in exocrine tissue, providing further insights about wolframin function.

CHAPTER 2

MATERIALS AND METHODS

2.0.0 Suppliers of laboratory materials

Appendix 1A-B lists supplier details of laboratory equipment and reagents. Unless otherwise stated, general tissue culture plastics were obtained from Corning Life Sciences (Appleton Woods).

2.0.1 Patient DNA samples and clinical data

The Wolfram patient DNA was obtained from Birmingham Womens Hospital diagnostics department, courtesy of Prof. T.G. Barrett. Stored samples were collected from patients referred to the West Midlands Regional Genetics Laboratory over a period from the early 1990s to 2010. Healthy control DNA was available from the Birmingham University, Medical and Molecular Genetics laboratory.

Patient clinical data was provided by Prof. T.G Barrett, Diabetes Unit, Birmingham Children's Hospital, UK.

2.0.2 Cell lines

The mammalian cell lines used for this work, unless stated, were available from the Birmingham University, Medical and Molecular Genetics laboratory.

2.1.0 DNA amplification

The polymerase chain reaction (PCR) is used to amplify specific target regions of DNA from starting material by several orders of magnitude. This technique enables generation of quantities necessary for applications such as sequencing, cloning and gene expression assays.

Starting DNA is mixed with an excess of deoxynucleoside triphosphates (dNTPs), which provide the building blocks for synthesis of new DNA strands. A heat-stable DNA polymerase in a suitable buffering solution provides the enzymatic activity for assembly of new complementary 'daughter' DNA strands from primers specific to both sense and anti-sense strands of the parent DNA.

Reactions, typically carried out on a thermal cycling machine, involve 20-40 repeat cycles of 3 discrete temperature intervals. The first *denaturing* step is achieved at temperatures between 94˚C-98˚C and serves to separate the DNA template into single

strands by disruption of the hydrogen bonds between complementary bases. The second *annealing* step allows primers to bind to the DNA templates and is achieved at temperatures specifically just below the melting point of the primers, usually in the region of 50˚C-68˚C. At this stage the polymerase is able to bind to the primer-template complex and initiate synthesis of new DNA strands. A final *elongation* step at a temperature optimal for polymerase activity (72˚C-80˚C) allows dNTPs to be sequentially added to the template by condensing the 5' phosphate group of the free dNTPs with the 3' hydroxyl group at the end of the new DNA strand. In optimal conditions exponential amplification of the DNA fragment occurs. Once the repeat cycles have completed, an additional elongation step of several minutes usually concludes the reaction to ensure any remaining single-stranded DNA is fully extended.

2.1.1 Genomic DNA PCR

Components

Standard 30µl reactions consisting of 50ng DNA and 20pmol of appropriate forward and reverse primers (see appendix 2), were mixed with reagents obtained from Invitrogen including: 0.25mM dNTP mixture, 2.0mM magnesium chloride, 0.5 units *Taq* DNA polymerase and x1 PCR reaction buffer. Q-solution (x1), a novel additive developed by Qiagen, was added to reactions where appropriate. This reagent helps to increase the efficiency of reactions containing G-C rich templates, which are often particularly difficult to amplify.

The remaining reaction volume (up to 30 µl) was made up with ddH_2O and mixed thoroughly by pipetting.

Conditions

Initial reaction conditions were set at 95˚C for 5 minutes for preliminary denaturation, followed by 35 cycles of 95˚C for 30 seconds, a specified annealing temperature (see *optimisation* below) for 30 seconds, and 72˚C for 1 minute to allow for elongation. The programme was completed with a final extension time of 10 minutes at 72˚C.

Optimisation

Primer annealing temperatures were initially set according to GC content using the formula (64.9+0.41 {GC%}-600/n), where n is equal to the number of bases in the primer. Extension time allowed 1 minute for every 1kb of DNA to be amplified, with a minimum of 30 seconds. Optimisation of PCR was achieved with variation of annealing temperature, and magnesium chloride concentration. Touchdown PCR programmes were also used to successfully amplify DNA fragments from particularly difficult primer sets. These programmes include additional cycles of annealing and elongation at decreasingly higher than optimal temperatures to minimize non-specific primer annealing.

Controls

To test for the presence of possible contaminating amplicons, negative controls containing all reaction components except template DNA (substituted with ddH_20) were included.

Positive external control samples, known to amplify in given conditions, were also used to verify negative amplification results where appropriate. These controls test the efficiency of primer sets, reaction components and cycling conditions.

Primers

Primers were designed manually or using the Primer 3 design software (Rozen, 2000) and obtained, unless stated otherwise, from AltaBioscience. Primer sets ideally spanned a maximum of 400-600 base pairs (bp) of DNA with the 5' end at least 50 bp upstream of the region of interest. Oligonucleotides ranged between 20-30 bases in length, with ~50-60% G+C base composition, to provide for practical annealing temperatures. Inverted repeat sequences and self-complementary sequences were avoided to minimize the synthesis of primer dimers and 2^o structures such as hairpins. Runs of any single base or dinucleotide repeats were also avoided to reduce mispriming. Where possible G and C rich 3' ends were selected, to enhance the annealing end for extension.

A comprehensive list of primer sequences can be found in appendix 2.

2.1.2 Reverse-transcription PCR

Reverse transcription PCR enables synthesis of single-stranded complementary DNA (cDNA) from total or poly (A)+ isolated RNA, and these types of assay are a useful method for characterizing and confirming gene expression patterns. The reverse transcription reaction is carried out using a retroviral enzyme with the capability of producing cDNA copies from RNA templates. The resulting cDNA may then be stored at -20°C or used as a template in a subsequent traditional PCR amplification (section 2.1.1) with specific primers as desired.

Conditions

According to protocol, two methods of cDNA synthesis were used:

Promega Reverse Transcription system

Isolated RNA, extracted from cells using TRIzol (section 2.10.2), was prepared by initial incubation at 70°C for 10 minutes, followed by a brief spin in a micro-centrifuge and incubation on ice. Reactions were mixed using Reverse Transcription System reagents from Promega consisting of: 5mM $MgCl_2$, reverse transcription buffer (x1), 1mM dNTP mixture, 0.5 units recombinant RNasin ribonuclease inhibitor, 15 units AMV reverse transcriptase (high conc.), 0.5µg (1µl) random primers, 1µg total RNA and nuclease free H_20 to a final volume of 20µl.

Reactions were incubated at room temperature for 10 minutes, then 42°C for 15 minutes to allow primer annealing and enzymatic extension of new cDNA strands by reverse transcriptase. Samples were then heated at 95°C for 5 minutes to inactivate the AMV reverse transcriptase, and incubated at 4°C to prevent binding of enzymes to the cDNA.

Applied Biosystems High Capacity cDNA Reverse Transcription kit

A reverse transcription master mix was prepared on ice with: 2µl RT buffer (x2), 0.8µl 100mM dNTP mix, 2µl random primers, 1µl MultiScribe reverse transcriptase, 1µl RNase inhibitor and 3.2µl nuclease-free H_20. In a 96-well plate, 10µl of master mix was pipetted into each well and mixed with 10µl (350ng) of sample RNA. Reverse transcription was carried out on a thermal cycler, following the same principles as that described above, with slightly different optimal temperature intervals: 25°C for 10 minutes, 37°C for 120 minutes, 85°C for 5 minutes, then cooling at 4°C.

2.1.3 Real-time PCR/Quantitative PCR (Q-PCR)

Q-PCR is a method which allows simultaneous amplification, detection and quantification of targeted DNA molecules. It is often preceded by reverse transcription of RNA samples to enable measurement of mRNA levels in cells and tissues. The general principles of the reaction are similar to standard PCR (section 2.1.1) but include probes and fluorescent reporter dyes for specific or non-specific target sequences, which

permit quantification of copy numbers, either absolutely or relative to normalising genes.

Conditions and analysis

Real-time PCR was employed for the measurement of *WFS1* expression in different sample sets, normalized to the β-actin gene. Total RNA was harvested from cells with Trizol (section 2.10.2) and reverse transcribed to cDNA using the Applied Biosystems High Capacity cDNA Reverse Transcription kit (section 2.1.2).

The Taqman Gene Expression Assay from Applied Biosystems, which runs under universal cycling conditions, was the assay of choice as the system is target specific and pre-optimized for measurement of a wide range of genes, including *WFS1*. FAM dye-labelled probes for a specific region of the human *WFS1* and *ACTB* (β-actin) genes were selected (Applied Biosystems) and used in conjunction with Taqman universal mastermix.

Samples were prepared in a 7500 System compatible 96-well plate (Applied Biosystems) including: 5µl cDNA (8.75ng), 10µl Taqman universal mastermix, 1µl (250nM) combined forward and reverse primers (labelled), and ddH_20 to a final volume of 20µl. Q-PCR runs were then performed on an Applied Biosystems 7500 Real-time PCR machine with 7500 v.2.0 software using the comparative C_t (threshold cycle) method for relative

quantification of *WFS1* expression. This programme features an initial 10 minute hold step at 95°C, followed by x40 cycles of 95°C for 15 seconds and 60°C for 1 minute.

The C_t value for a given sample following Q-PCR is defined as the number of cycles required for the signal to exceed background levels (threshold). These values are inversely proportional to the amount of target nucleic acid in the sample. C_t values for all samples were normalised to an appropriate endogenous house-keeping gene (e.g. *ACTB*), then experimental samples compared to 'non-treated' controls. The change in gene expression between samples and controls is expressed as [delta][delta]C_t and was calculated according to the formula below:

$$[delta][delta]C_t = [delta]C_{t\ experiment} - [delta]C_{t\ control}$$

Where [delta]$C_{t\ experiment}$ is the C_t value for any sample normalised to the endogenous housekeeping gene and [delta]$C_{t\ control}$ is the C_t value for the control also normalised to the house-keeping gene. In this case the experimental samples refer to *WFS1*-variants and reference controls represent wild-type *WFS1*.

Assuming efficiency is 100%, we also calculated fold change between control and experimental samples (2^ [delta][delta]C_t = fold change):

$$\text{Ratio experimental/control (fold change)} = \frac{\text{fold change in target gene (}\textit{WFS1-mutant}\text{)}}{\text{fold change in control gene (}\textit{WFS1-wild-type}\text{)}}$$

2.2.0 Restriction Digestion

A range of restriction endonucleases were used to cut specific fragments of DNA for the purposes of isolation and identification. DNA was incubated with an appropriate amount of restriction enzyme, in its respective buffer as recommended by the supplier. Reactions typically incubated for 1-3 hours at the enzyme specific optimal temperature (~37°C). Double digests were performed with enzymes activated in the same buffer and incubation temperature.

Typical digestions included a unit of enzyme per microgram of starting DNA, where one enzyme unit (depending on the supplier) is defined as the amount of enzyme needed to completely digest one microgram of double-stranded DNA in one hour at the appropriate temperature. Completed restriction digests were separated into constituent fragments and visualized by agarose gel electrophoresis (section 2.3.0).

2.3.0 Agarose Gel Electrophoresis

Electrophoresis on agarose gel was used to detect successful PCR amplification and restriction digestion of correctly sized DNA fragments. When DNA (negatively charged) is moved by an electric current through an agarose matrix, products can be distinguished by position (band) as differing sized fragments move through the gel at variable rates towards the cathode. Gels were made with agarose heated to dissolve in

x1 TBE buffer solution (890mM Tris Borate pH8.3 and 2mM Na_2EDTA). Typical gel concentrations were 1-2%, depending on product size to be resolved. Approximately 20ng ethidium bromide (EtBr) was added on cooling of solution. EtBr fluoresces under ultraviolet light when intercalated with DNA, thus allowing visualization of products.

Once cooled (below 60°C), molten agarose was poured carefully into an electrophoresis tray and left to polymerize with one or several well-combs inserted (size dependent on samples). The combs were removed from solidified gel and trays placed into electrophoresis tanks, immersed in x1 TBE buffer solution. Samples (≥20ng) mixed in loading buffer (30% glycerol, ddH_20 and Orange G dye) were pipetted into empty wells and an appropriately sized DNA ladder (Invitrogen) was loaded alongside as a size marker. Gels were routinely run at 180 volts as required.

The results were visualized and imaged in a UV light box connected with a high-resolution camera and GelDoc software (Syngene).

2.3.1 Agarose Gel Extraction

The QIAquick gel extraction kit (Qiagen) was used to extract DNA from agarose gel following manufacturers' guidelines. DNA samples run through agarose gel were

visualized under a UV light source and bands of interest were excised using a sterile scalpel.

Three volumes of buffer QG were added to 1 volume of gel and incubated at 50°C for 10 minutes to dissolve. One gel volume of isopropanol was added to the mixture, which was then loaded onto a Qiagen spin column. The column was centrifuged for 1 minute at 13,000rpm and the flow-through discarded. The DNA (bound to column matrix) was then washed and centrifuged at 13,000rpm; firstly with 0.5ml of buffer QG to remove all traces of agarose, then 0.75ml of wash buffer PE. A further 1 minute spin ensured the removal of any remaining flow-through from the column. The DNA was eluted in 30µl of sterile H_2O following a final centrifugation step of 1-2minutes at 13,000rpm.

2.4.0 Sequencing

DNA Cleanup

PCR products were cleaned up in preparation for sequencing using a combination of hydrolytic enzymes in suitable buffer. Antarctic Phosphatase, used to remove excess dNTPs and primers, was combined with Exonuclease I, which degrades residual single-stranded primers and any extraneous single-stranded DNA from the PCR reaction.

Reactions were carried out on a thermal cycler in 10µl total volumes, comprising 5µl PCR product, 1µl Antarctic Phosphatase (20 units), 1µl Antarctic Phosphatase buffer,

0.25μl Exonuclease I (0.25 units), and 0.75μl sterile H_2O. Enzymes were activated at 37°C for 30 minutes then subsequently inactivated at 85°C for another 30 minutes.

Sequencing Reaction

Sequencing reactions were carried out on 96-well plates (Fisher) in a thermal cycler. Each sample was prepared in a 10μl total volume, comprising; 5μl DNA (≥5ng), 1μl Big Dye Terminator mix (Applied Biosystems), 2μl sequencing buffer (200mM Tris-HCl pH 9.0, 5mM $MgCl_2$), 1μl sterile H_2O, and 1μl primer (20pmol).

The samples were heated to 95°C for 25 seconds, cooled to 50°C for 25 seconds, and ramped up to 60°C over 28 cycles. Following the reaction, samples were kept covered at 4°C.

Ethanol precipitation

Ethanol precipitation was used to concentrate and de-salt DNA prior to sequencing. 100μl of absolute ethanol, plus 3.5μl sodium acetate (3M pH 5.2) EDTA (125mM pH 8.0) was added to samples taken directly from a sequencing reaction (see above). DNA was then pelleted in a centrifuge (4°C) at 4,000rpm for 30 minutes. Supernatant was poured from wells by inversion of the plate.

Two subsequent washes in 70% ethanol (200μl) were carried out as per the above conditions. Following the final spin, and supernatant removal, the plate was inverted

and centrifuged in a swinging bucket rotor at 400rpm for 30 seconds to remove excess ethanol. The plate was then left to air dry for up to 30 minutes.

Sequencing analysis

Samples were resuspended after air-drying in 10μl of Hi-Di Formamide, a ready to use stable injection solvent from Applied Biosystems that is suitable for use on ABI PRISM DNA analysis machines. The plate was denatured on a heat block at 95°C for 5 minutes then placed directly onto ice for at least 5 minutes to cool prior to sequencing.

DNA was sequenced on an automated ABI3730 sequencer and analysed using the ABI sequencing analysis software (Applied Biosystems).

2.5.0 Microsatellite marker analysis

Microsatellites, or simple sequence repeats (SSRs) are tandemly repeated genetic elements of 1-4 nucleotides, usually found in non-coding regions of DNA. The number of repeats in a given microsatellite is highly variable, which makes them very useful as genetic markers for application in mapping, linkage analysis and tracing inheritance patterns.

PCR amplification of SSRs of interest generates DNA fragments of differing lengths according to the haplotype of the individual. A fluorescent probe can then be mixed

with the products and loaded onto a sequencing machine with appropriate software (e.g. GENESCAN, ABI) that will analyze the fragment lengths for comparison.

In this study (section 3.3.2) microsatellite analysis was used to compare specific markers between families with *WFS2*-associated Wolfram syndrome to investigate the likelihood of common ancestry. A common haplotype previously described in *WFS2* families (Amr, 2000), was investigated in our affected individual with *WFS2* mutations.

PCR with selected markers

PCR reactions were carried out on a thermal cycling machine (section 2.1.1) using labeled primers (appendix 2) corresponding to the critical region (D4S3256, D4S1531, D4S1564; Amr, 2000) and DNA from: our affected individual, a normal control, and an affected individual from the earlier *WFS2* cohort (Amr, 2000) kindly provided by Dr Rita Shiang, Virginia Commonwealth University.

Total reaction volumes of 10µl comprised: 20ng of genomic DNA, fluorescently labeled forward primer (0.2mM), unlabelled reverse primer (0.2mM), dNTPs (0.2mM), *Taq* polymerase (0.5 units) in x1 *Taq* buffer (containing 1.5mM $MgCl_2$). After initial denaturation (3 min at 95°C), amplification was performed over 26 cycles (95°C for 30sec, 55°C for 30sec, 72°C for 30sec) followed by a final extension at 72°C for 10 minutes.

Concentrated PCR products (1μl), and 1:10, 1:20 and 1:50 dilutions in ddH_2O, were mixed with 0.5μl GeneScan Internal Lane Size Standard (Applied Biosystems) and 12μl deionized formamide and loaded onto 96 well plates. Empty wells were filled with equal volumes of ddH_2O (12.5μl). The samples were denatured for 5 minutes at 95°C on a heat block and then placed on ice to cool.

Products were analyzed on an automated ABI3730 sequencer, using the Genotyper v3.5/4 software (ABI).

2.6.0 Site-directed mutagenesis

The QuickChange Site-Directed Mutagenesis kit (Stratagene) was used to introduce site-specific point mutations into double-stranded plasmid DNA. Complementary primers containing a desired single base change were extended by *PfuTurbo* DNA polymerase during a thermal cycling reaction. The products were treated with *Dpn I* endonuclease, which specifically targets methylated and hemimethylated DNA. This allows targeted degradation of 'non-mutated' parental DNA templates since the majority of *E.coli* strains, including vectors used in this protocol, contain dam methylated DNA. The newly synthesized vector DNA incorporating the desired mutation could therefore be selected and transformed into XL-1 Blue supercompetent cells, which serve to repair the nicks in the plasmid and produce large numbers of copies.

Mutagenic primer design

The web-based QuikChange Primer Design Programme from Stratagene (table 2.3) was used to design primers for mutagenesis.

The oligonucleotide design parameters specify primer lengths ranging from 25-45 bases with the desired mutation in the middle. Both forward and reverse primers should contain the mutation with matched complementary sequences to opposite strands of the plasmid. The primer sequence should ideally terminate with one or more guanine or cytosine residues and have over 40% GC content. An optimum melting temperature (Tm) is ≥78°C, which was checked with the following formula;

$$Tm = 81.5 + 0.41\,(\%GC) - 675/N - \%\ \text{mismatch}$$

Where N is the length of oligonucleotide, and values for %GC and % mismatch are whole numbers.

Mutagenesis reaction

Sample reactions were set up with 10ng of dsDNA template, 5μl of 10x reaction buffer, 125ng of both forward and reverse primers (FPLC purified), 1μl dNTP mix (of a proprietary composition optimised by Stratagene), 3μl Quiksolution (Stratagene), 1μl (2.5 units) *PfuTurbo* DNA polymerase and ddH_2O to a final volume of 50μl.

A control reaction was included to test the efficiency of the polymerase and demonstrate the effectiveness of the QuikChange Mutagenesis kit. This comprised 10ng of pWhitescript 4.5kb control plasmid, 5μl of 10x reaction buffer, 125ng of both forward and reverse control primers, 1μl dNTP mix, 3μl Quiksolution, 1μl (2.5 units) *PfuTurbo* DNA polymerase and ddH_2O to a final volume of 50μl.

The pWhitescript plasmid contains a stop codon (TAA) in the position where a glutamine codon (CAA) would normally appear in the β–galactosidase gene of the pBluescript II SK(-) phagemid. The control primers create a point mutation to revert the T residue of the stop codon back to the C of glutamine. This enables successfully mutated and transformed clones to appear 'blue' when grown on selective media containing IPTG and X-Gal.

On a thermal cycler, reactions were run at 95°C for 1 minute, followed by 12 cycles of; 95°C for 50 seconds, 60°C for 50 seconds, and 68°C for 9 minutes.

Dpn I digestion

Following the temperature cycling, the products were cooled on ice for at least 2 minutes, before addition of 1μl (10U) *Dpn I* restriction enzyme. The solutions were mixed with a pipette and incubated at 37°C for 1 hour.

Transformation of XL1-Blue Supercompetent cells

XL1-Blue Supercompetent cells were thawed on ice and aliquoted to pre-chilled 1.5ml polypropylene tubes in 50μl volumes. The *Dpn I* treated DNA samples (1μl) were added to the cells, mixed gently with a pipette tip, and incubated on ice for 30 minutes. To test the transformation efficiency of the Supercompetent Cells, 1μl of pUC18 control plasmid (0.1ng) was added to a separate aliquot of cells as a positive control.

Following the incubation period, the tubes were heat-shocked in a 42°C water bath for 45 seconds, and placed immediately on ice for 2 minutes. Recovery medium was added to the cells (0.5ml SOC) and the transformed cells were placed into an orbital shaker for 1 hour at 37°C.

The cells were subsequently spread in 250μl volumes onto LB-ampicillin agar plates containing 80μg/ml X-Gal and 20mM IPTG (section 2.7), and incubated at 37°C for >16 hours.

2.7.0 Microbial cell culture and manipulation

Microbes were exploited for their ability to replicate quickly and reliably in the laboratory for the purposes of cloning and manipulation of plasmid DNA. Typical steps for handling *E. coli* strains are outlined in figure 2.0.

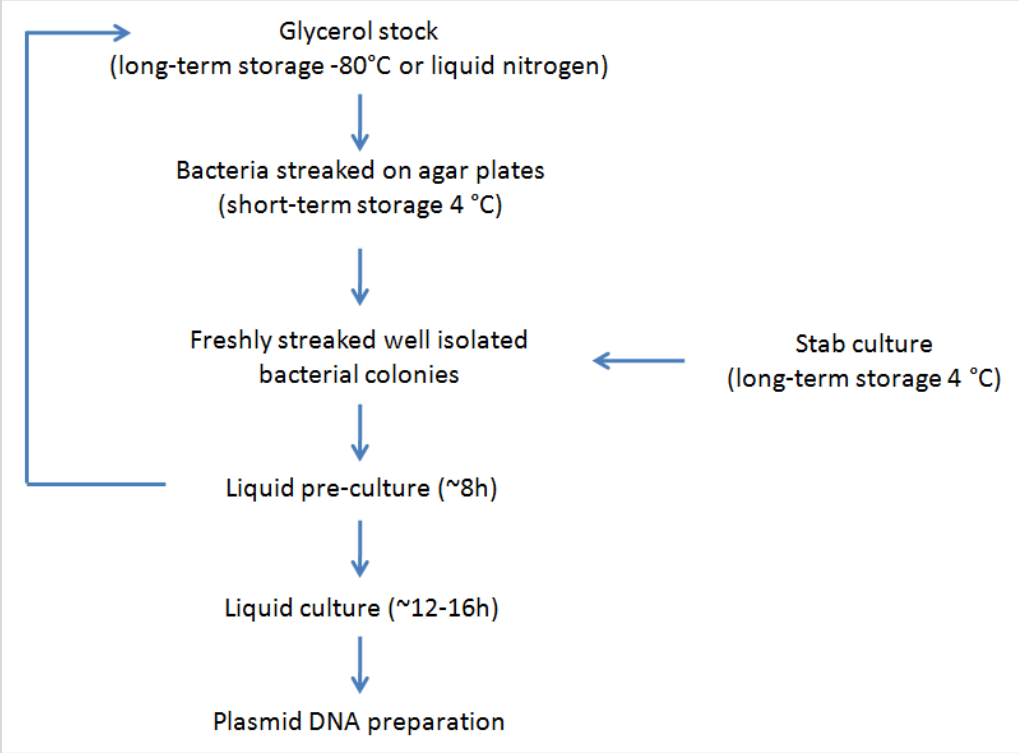


Figure 2.0. Essential steps for storage and handling of *E. coli* strains (adapted from QIAGEN Bench Guides)

2.7.1 Long-term storage and revival of bacteria

Although agar plates can be used for short-term storage of bacteria (4°C), glycerol stocks of various strains were prepared for long-term storage in liquid nitrogen (approximately -196°C). Screw cap cryovials (2ml) were filled with 0.85ml logarithmic-phase cell culture and 0.15ml sterile glycerol and vortexed to thoroughly mix. Cells were frozen to -80°C before transfer to liquid nitrogen. Stocks cells were revived by warming ampoules and inoculating selective growth plates (section 2.7.3).

2.7.2 Preparation of Luria-Bertani (LB) broth and agar plates

LB broth was prepared using 20g of LB broth powder (Sigma) per 100ml water required. The suspension was autoclaved and allowed to cool prior to use or storage (4°C).

LB agar plates were prepared using 35g of LB agar powder (Sigma) per 100ml water required. The suspension was autoclaved and allowed to cool below 50°C. In a laminar flow cabinet any necessary antibiotics or selection reagents were added and swirled gently to mix. Molten agar was then poured into sterile Petri dishes (10cm diameter) filling to approximately 1/8", taking care to avoid formation of air bubbles. Once air dried, the plates were ready for use or storage at 4°C. The plates were stacked and inverted for refrigeration to prevent condensation dripping onto the agar surface.

Selection additives

During cell culture, selection additives such as antibiotics are combined with the growth media as a method of isolating clones of interest. Depending on the genetic complement of the cells, the ability of microbes to grow on a given substrate varies and this can be manipulated in the laboratory by introducing plasmids containing specified genes, such as antibiotic resistance or other markers, which can be used to distinguish transformed cells.

Antibiotics

Antibiotics were added to broth or molten agar after cooling below 50°C at a typical concentration range of 50-100μg/ml.

X-Gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside)

X-Gal was used as a substrate for the purposes of blue/white screening, a technique used to indicate the presence of β-galactosidase expression, encoded by the *lacZ* gene. β-galactosidase cleaves X-Gal to produce an insoluble blue product, so incorporating *lacZ* onto cloning plasmids can be a useful tool for distinguishing clones of interest. For the purposes of cloning, successful ligations can be isolated by inserting fragments within a functional *LacZ* sequence so that all 'ligated' colonies grown on X-Gal media will be white in colour.

X- Gal was dissolved at 2mg/ml in dimethylformamide (DMF) and stored in the dark at -20°C. For incorporation into agar plates 2μl of X-Gal stock per ml of autoclaved molten agar (cooled below 60°C) was mixed before pouring to give a final concentration of 40μg/ml. Concentrations vary according to protocol.

IPTG (isopropyl beta-D-thiogalactopyranoside)

IPTG is a compound that triggers transcription of the *lac* operon and production of the β-galactosidase enzyme. For application in blue/white colony screening protocols IPTG was used as an additive to agar plates in combination with X-gal.

A 0.1M stock solution of IPTG in sterile H_2O was prepared, filtered and stored at -20°C. For incorporation into agar plates, 1µl IPTG stock per ml of autoclaved molten agar (cooled below 60°C) was mixed before pouring to give a final concentration of 0.1mM IPTG. Concentrations vary according to protocol.

2.7.3 General culture and selection of bacterial cells

Growth on agar plates and selection of clones

To obtain isolated clonal colonies: in a sterile environment, an inoculum of bacteria was taken from a glycerol stock, another plate, or from a stab culture and streaked with a loop onto the corner of an agar plate containing appropriate selection additives. A fresh loop was then passed through the first streak and streaked again across a fresh corner of the plate. The plate was incubated upside down at 37°C for 12-24 hours until colonies developed.

To select a single clone, a well isolated colony was picked gently from the surface of the plate with a sterile loop and seeded into 2-5mls of LB broth (with selection additives). Following an incubation period of ~16 hours (37°C), the cells were either harvested or added as a pre-culture to a larger volume of LB broth (100ml), which was grown for a further 16 hours (37°C) to propagate much larger quantities of cells.

Growth in liquid culture

Bacterial cell cultures were grown in Luria-Bertani (LB) broth at 37°C in an orbital shaker. Flasks that enabled liquid culture with a large surface to volume ratio were used to facilitate aeration and respiration; typically no more than 10-20% of the total volume of the vessel was filled with broth.

Following propagation of liquid culture cells were harvested. Centrifugation (≤6,000xg for 5-15 minutes) in suitable polypropylene tubes removed cells from suspension to form a dense pellet at the bottom of the tube. The supernatant was decanted and tubes were inverted for a few minutes to allow any excess media to run out. The remaining bacterial pellets were then ready for DNA extraction (section 2.7.5) or stored (-20°C).

2.7.4 Transformation of bacterial cells (cellular cloning)

Bacterial cells were transformed with plasmids of interest and subsequently grown into culture derived from a single clone. This allowed rapid multiplication of plasmid copies for later extraction and purification. Foreign DNA was introduced into competent cells using chemical or electrical transformation.

Electroporation has the advantage of yielding more efficient transformation although is limited by either small volumes or pre-treatment of DNA as the procedure is sensitive to high concentrations of salt.

Chemical transformation

Suspension of cells in $CaCl_2$ solution confers chemical competence. Ca^{2+} ions are thought to create pores in cell membranes and mask the negative charges on DNA to allow easier passage across the hydrophobic cell membrane. A short heat shock applied to a cell/DNA mixture creates a thermal current that sweeps DNA into the cells.

Transformation of purified plasmid preps was carried out using subcloning efficiency (1 x 10^6 cfu/μg pUC19) DH5α competent cells (Invitrogen). Cells of increased efficiency, according to requirements were favoured for use in molecular cloning protocols (section 2.19).

Chemically competent cells (stored at -80°C) were thawed on ice and pipetted gently (50-100μl) into pre-chilled 1.5ml micro-centrifuge tubes. DNA preps (1-5μl) were added to individual aliquots of cells, stirred carefully with the pipette tip, and incubated on ice for 30 minutes. A separate transformation control containing 2.5μl of pUC19 plasmid (250pg) was also prepared. The mixture was then heat shocked for up to 1 minute in a water-filled heat block at 42°C, and transferred to ice for 2 minutes. Shocked cells were then suspended in 0.5ml SOC medium (Super Optimal Broth with Catabolite repression)

from Invitrogen (2% w/v bacto-tryptone, 0.5% w/v yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$, 20mM glucose, ddH_2O) and recovered for 1 hour at 37°C, in an orbital shaker.

Recovered cells were plated onto LB agar (section 2.7.2) containing appropriate selection additives and incubated at 37°C for up to 16 hours.

Electroporation

The principle of electroporation is based on applying a quick voltage shock across a cell membrane, which causes temporary disruption to the phospholipid bilayer allowing polar molecules (e.g. DNA) to pass into the cell. Cells (available commercially) are made electrically competent by removal of all medium salts, which ensures a charge is not conducted through the medium.

ElectroMAX DH5α-E electrocompetent cells (Invitrogen) were thawed on wet ice. In pre-chilled micro-centrifuge tubes 20µl of cells were gently mixed with 1µl (≤100ng) of purified DNA. A control sample with 1µl (100pg) pUC19 plasmid DNA was included to assess the transformation efficiency.

The cell/DNA mixture was carefully pipetted into a 0.1cm cuvette, loaded into an electroporator (BioRad GenePulser) where a 2.0kV pulse was applied. SOC medium (1ml) was immediately added to cells in the cuvette and the suspension was transferred

to a 15ml polypropylene tube and recovered for 1 hour at 37°C in an orbital shaker (225rpm).

A range of dilutions in SOC media were made with cell suspensions following recovery. Typically 50-100µl of cell suspension was spread onto selective LB plates (section 2.7.2) and incubated overnight at 37°C.

2.7.5 Extraction of plasmid DNA from bacterial cells

Plasmid DNA can be separated from bacterial genomic DNA following the basic principles of alkaline lysis, first described in 1979 (Birnboim, 1979). Cells containing the desired plasmids are cultured then separated from growth medium by centrifugation, before resuspension in a specialised solution containing glucose, RNase A, EDTA and Tris. These components serve to maintain the osmotic pressure of the cells, disrupt the cell wall, protect plasmids from damage by DNases, and degrade cellular RNA (after lysis).

Cells are then lysed in NaOH (sodium hydroxide) and SDS (sodium dodecyl sulfate), which solubilise cell membranes, break down cell walls and also convert double-stranded DNA (dsDNA) to single-stranded DNA (ssDNA) and denature most of the proteins. Isolation of plasmid DNA from genomic DNA occurs during neutralisation of

lysis buffer with potassium acetate, which decreases the alkalinity of the mixture allowing hydrogen bonding between bases of ssDNA to reform into double strands. Due to small size, circular plasmid DNA re-natures much more readily than longer strands of genomic DNA that cannot properly re-anneal. The ssDNA and denatured proteins form insoluble precipitates in this solution, which can be removed by centrifugation.

The final clean-up step enables purification and concentration of plasmid DNA from the extraneous cell debris and solution additives. This can be achieved in a number of ways including chloroform extraction, ethanol precipitation, and variations of affinity chromatography.

Plasmid DNA from bacterial culture (section 2.7.3) was isolated with Qiagen plasmid kits (Qiagen) following manufacturers guidelines, for use in subsequent sequencing, transfection and cloning. The kits comprise; bacterial cell lysis buffer, anion-exchange or silica columns that are designed to selectively bind DNA under appropriate salt and pH conditions, and appropriate washing and elution buffers to purify the immobilized DNA.

Miniprep kits were used to extract DNA from small volumes of cells (~5ml) to be subsequently used for analysis or transformation. Maxipreps, which include endotoxin removal, were used to extract larger quantities of transfection grade plasmid DNA (typically from 100ml culture).

QIAprep Spin Miniprep Kit

Cells were harvested from 5ml cultures by centrifugation at 4,000rpm (5 minutes), resuspended in 250μl buffer P1 containing RNase and transferred to a 1.5ml micro-centrifuge tube. Lysis buffer P2 (250μl) was then added to the cell mixture and incubated at room temperature for up to 5 minutes. The lysis reaction was neutralized with addition of buffer P3 (350μl) and the mixture was centrifuged for 10 minutes at 13,000rpm to pellet the cell debris. Supernatant was loaded onto a QIAprep spin column and passed through the column matrix by centrifugation (1 minute) at 13,000rpm. Two wash buffers, PB (0.5ml) and PE (0.75ml) were both sequentially loaded and centrifuged through the column at 13,000rpm for 1 minute each, followed by an additional 1 minute spin (13,000rpm) to remove any residual buffers from the column. DNA was eluted into a clean micro-centrifuge tube by adding 30μl sterile H_2O to the centre of the column, standing for 1 minute, and centrifuging for 1 minute (13,000rpm). Plasmid DNA was stored at -20°C.

QIAGEN Plasmid Maxi Kit

Cells were harvested from 100ml cultures by centrifugation at 6,000xg for 15 minutes and resuspended by vortexing in 5ml buffer P1 containing RNase. Lysis buffer P2 (5ml) was then added to the cell mixture and incubated at room temperature for up to 5 minutes. The lysis reaction was neutralized with buffer P3 (10ml) and incubated on ice for 20 minutes to allow precipitation of DNA, proteins and cell debris. The mixture was then centrifuged (14,000rpm) for 30 minutes at 4°C, and the supernatant filtered

through 3mm Whatman filter paper. A QIAtip spin column was equilibrated by applying 10ml of buffer QBT and allowing the column to empty by gravity flow. The filtrate (containing DNA) was passed through the prepared column by gravity flow allowing DNA to bind selectively to the column matrix. The column was washed twice with 30ml of buffer QC to remove impurities such as RNA, metabolites and proteins, and finally eluted to a clean 50ml centrifuge tube in 15ml sterile H_2O. Isopropanol (10.5ml) was added to the eluate and centrifuged (4°C) at 14,000rpm for 30 minutes to precipitate DNA. The resulting DNA pellet was washed in 5ml of 70% ethanol and spun (4°C) for a further 10 minutes at 14,000rpm. The ethanol was carefully decanted from the tube leaving the DNA pellet to air-dry at room temperature for ~10 minutes. The dried DNA was resuspended in 200µl sterile H_2O and stored at -20°C.

2.8.0 Yeast cell culture and manipulation

2.8.1 General culture and selection of yeast cells

Yeast cells were cultured at 30°C in a blend of yeast-extract, peptone and dextrose (YPD), and single colonies (clones) were obtained from YPD agar plates streaked with small volumes of cells.

Plasmid selection was achieved with the use of 'auxotrophic' mutated strains of yeast that cannot synthesize particular amino acids necessary for growth/metabolism on specific media. Transforming yeast with plasmids containing genes coding for the

required factor enables only the transformants to grow on the selective media. A minimal base medium can be prepared for this purpose with one, or a combination of missing essential nutrients: leucine, tryptophan, adenine and histidine.

A further method of identifying potential protein interactions was blue/white screening with X-α-Gal and IPTG (see chapter 5).

2.8.2 Preparation of YPD broth and agar plates

YPD broth was prepared using 50g YPD (Clontech) per 100ml water required. The suspension was autoclaved and allowed to cool before use or storage at 4°C.

YPD agar plates were prepared using 67g of YPD agar powder (Clontech) per 100ml water required. The suspension was autoclaved and allowed to cool to ~50°C. Any necessary selection reagents were added in a laminar flow cabinet, and approximately 20-25ml of molten agar poured into each plate. Plates were left to air dry before use or storage at 4°C.

Selective drop-out media (SDM)

SDM was prepared using 27g of minimal SDM base (Clontech) per 100ml water required. The appropriate drop-out supplement (Clontech) was added and the mixture autoclaved and allowed to cool before use or storage at 4°C.

Drop-out supplements resulted in media;

i) minus leucine

ii) minus tryptophan

iii) minus leucine/tryptophan

iv) minus leucine/tryptophan/adenine/histidine (Quadruple drop-out)

X-α-Gal/IPTG indicator plates

Molten YPD agar was allowed to cool to ~50°C before the addition of IPTG and X-Gal to final concentrations of 0.1mM and 40μg/ml respectively (section 2.7.2). The mixture was gently swirled, poured (~25ml) into 10cm petri dishes and left in a laminar flow cabinet to set before use or storage at 4°C.

2.8.3 Long-term storage and recovery of yeast

Yeast strains/clones were stored long-term at -80°C in 1.5ml micro-centrifuge tubes. One large colony was suspended and mixed evenly in YPD medium with 25% glycerol. Dependent on clone, minimal media as appropriate was used to maintain selective pressure.

On recovery, a small amount of thawed yeast was streaked onto appropriate agar plates and incubated at 30°C for ~4 days, or until individual colonies reached a size of ~2mm in diameter. Yeast strains were grown on YPD agar, and yeast clones grown on SDM agar.

2.8.4 Transformation of yeast

Yeast cells were transformed with combinations of plasmids to confirm the initial yeast two-hybrid findings.

SD medium (5ml) was inoculated with yeast-colonies 2-3mm in size and mixed vigorously by vortexing. This suspension was transferred into a 50ml volume of SD medium, incubated at 30°C for 16-18 hours with shaking (250rpm) to stationary phase OD_{600}>1.5. Overnight culture was added to 300ml of YPDA (yeast peptone dextrose adenine) and placed into an orbital shaker at 230-270rpm for 3 hours (30°C). Cells were then pelleted by centrifugation at 1,000xg (5 minutes) in 50ml polypropylene tubes and thoroughly resuspended in 50ml volumes of sterile H_20. Each tube was pooled, centrifuged at room temperature for 5 minutes (1,000xg) and supernatant discarded. Yeast pellets were resuspended in 1.5ml of freshly prepared sterile 1.1x TE/LiAc solution (1.1x TE buffer (Clontech), 0.11M lithium acetate).

Plasmid combinations as desired were mixed in 1.5ml polypropylene tubes: 0.1µg DNA-BD/bait (*WFS1*-N), 0.1µg AD/library and 0.1mg herring testes carrier DNA. Competent yeast cells (0.1ml) and sterile PEG/LiAc solution (0.6ml) (40% polyethylene glycol 3350, 1x TE buffer, 0.1M lithium acetate) were added to the plasmids and mixed with a vortex at high speed. Tubes were incubated for 30 minutes (30°C) with shaking (200rpm) before adding 70µl DMSO and gently swirling the mixture.

Plasmids were introduced to yeast cells by heat-shock for 15 minutes in a 42°C water bath, swirling occasionally. Tubes were chilled on ice (2 minutes) prior to centrifugation

at 14,000rpm for 5 seconds. Supernatant was removed, cells resuspended in 0.5ml of YPDA, spread onto appropriate selective plates (section 2.8.1) and grown at 30°C.

2.8.5 Isolation of DNA from yeast

For preliminary analysis of positive clones identified from a yeast two-hybrid screen, the Yeastmaker Yeast Plasmid Isolation kit (Clontech) was used to isolate plasmid DNA from yeast cells. Plasmid extracts were then suitable for direct use as a PCR template or for transformation into *E. coli*.

Yeast clones (taken from SD minimal agar medium or frozen stocks) were initially resuspended in 50µl Potassium Phosphate (67mM KH_2PO_4; pH7.5) in a 1.5ml micro-centrifuge tube. The cells were lysed by adding 10µl lyticase solution (5 units/µl in TE buffer), thorough mixing and incubation at 37°C for 60 minutes. 10µl of SDS (20%) was added to the mix then tubes were vortexed vigorously for 1 minute. To ensure complete lysis, the cells were also put through one freeze/thaw cycle at -20°C and vortexed again.

CHROMA SPIN Columns were used for purification of plasmids: Columns were inverted and vortexed to thoroughly resuspend the gel matrix, the cap was removed and tip broken-off. The column was then placed into a 2ml polypropylene tube and centrifuged in a swinging bucket rotor at 700xg for 5 minutes to allow the buffer to run through. The column was placed into a clean micro-centrifuge tube where samples were carefully applied to the centre of the gel matrix within the column. After centrifugation for 5

minutes (700xg) the purified DNA filtered through into the eluate, which was collected and stored at -20°C.

2.9.0 Mammalian cell culture and manipulation

2.9.1 General tissue culture

Mammalian cell lines were grown in $25cm^3$/$75cm^3$ sterile vented flasks incubated at 37°C, with 5% CO_2. All handling and manipulation was carried out in class II tissue culture hoods using sterile instruments and consumables.

Cell maintenance and propagation

Cells were typically cultured in pre-warmed Dulbecco's modified eagle media (DMEM), supplemented with 10% foetal bovine serum, 2mM L-glutamine, 1% non-essential amino acids, and 1% penicillin/streptomycin solution (Invitrogen). The volume of medium added to each flask was sufficient to provide adequate coverage of cells without liquid contact with filter (~8ml per $25cm^3$/~13ml per $75cm^3$). Cells were grown to approximately 80-90% confluency before passaging in a 1:1-1:10 ratio depending on cell growth rate and experimental requirements. Table 2.0 outlines specific requirements for various cell lines.

For passaging adherent cell lines, growth media was firstly removed from the flask and discarded, followed by gentle washing of cells with phosphate buffered saline (PBS). Cells were then lifted from the flask by trypsinization in 1x trypsin-EDTA (ethylenediaminetetraacetic acid) solution (1ml per $25cm^3$/ 2ml per $75cm^3$) and incubation at 37°C for up to five minutes. The cell suspension was diluted with an appropriate volume of completed growth media and split as required.

For passaging semi-adherent or suspended cells, media was decanted from a growing flask to a centrifuge tube. Any remaining adherent cells were detached from the flask by incubation in 0.2% EDTA (1-2ml) at 37°C for up to 5 minutes and added to the centrifuge tube. The cell suspension was centrifuged ≤1,000xg for 3 minutes (22°C) and supernatant was decanted carefully from the pellet. The cells were resuspended in fresh completed medium, and diluted/split to desired volumes and placed into fresh flasks.

2.9.2 Freezing cells for long-term storage

Mammalian cell lines were stored long-term in liquid nitrogen. Approximately 1 x 10^6 cells were mixed in 1ml aliquots of 90% growth medium or foetal bovine serum, with 10% Dimethyl sulfoxide (DMSO) cryoprotectant and transferred to cryovials. The cryovials were then placed in a cryo-freezing container, containing 100% isopropanol and initially stored at -80°C to allow gradual freezing (1°C per minute) and minimal ice-crystal damage, before transferring vials to liquid nitrogen.

Cell line	Growth media	Passaging Optimal density and splitting
AR42J Rat exocrine pancreatic tumour	RPMI-1640 Supplemented with: 10% FCS 1% glutamine (2mM) 1% P/S (100mg/ml streptomycin, 100mg/ml penicillin) 1% non-essential amino acids	1-9x10^4 cells/ml 2ml 0.2% EDTA (ethylenediaminetetraacetic acid)/75cm^3 flask
COS7 African green monkey kidney	DMEM (Dulbecco's Modified Eagle Serum) Supplemented with; 10% foetal calf serum 1% glutamine (2mM) 1% P/S (100mg/ml streptomycin, 100mg/ml penicillin) 1% non-essential amino acids	1-9x10^4 cells/ml 2ml trypsin (0.25%)/75cm^3 flask
HEK 293 Human embryonic kidney	DMEM (Dulbecco's Modified Eagle Serum) Supplemented with; 10% foetal calf serum 1% glutamine (2mM) 1% P/S (100mg/ml streptomycin, 100mg/ml penicillin) 1% non-essential amino acids	1-3x10^5 cells/ml 2ml trypsin (0.25%)/75cm^3 flask

Table 2.0. Culture conditions for mammalian cell lines.

2.9.3 Reviving frozen cells

Frozen cells recovered from storage were floated in a 37°C water bath until just thawed (~1 minute) to minimize exposure to the toxic effects of warmed DMSO. The cell suspension was transferred to a polypropylene tube, diluted with ~5ml completed medium and centrifuged for 3 minutes at ≤1,000xg (22°C). The supernatant was decanted-off and the remaining cell pellet was resuspended in ~8ml of completed medium, placed into a 25cm^3 vented flask and incubated at 37°C with 5% CO_2.

2.9.4 Cell counting

A haemocytometer viewed under a light microscope was used to count cells. The slide and coverslip were firstly cleaned with 70% ethanol before placing the coverslip over the counting grid. Cells were pelleted by centrifugation at ≤1,000xg for 3 minutes and resuspended in a known volume of appropriate growth medium. Thorough mixing dispersed the cells as evenly as possible then 70µl of the suspension was loaded into the haemocytometer through the 'v' trough. Using the microscope x10 objective, all cells present in a 1mm^2 square were counted and recorded. Results were averaged from three independent sample counts. A 1mm^2 area is equivalent to the number of cells x 10^4/ml.

2.9.5 Transfection

All transfections unless specified were achieved with Fugene 6 or Fugene HD (Promega) transfection reagent. Cells for transfection were seeded into 10cm culture dishes (~1.7 x

10^5 cells) with suitable completed growth medium (table 2.0), and allowed to establish overnight to reach optimal transfection density (typically 50-80% confluency) according to manufacturers recommendations.

Transfection mixes were typically prepared in 1.5ml polypropylene tubes. Fugene was pipetted and mixed carefully into serum-free medium without contacting the tube wall directly (3µl: µg DNA). Purified plasmid DNA (1µg) at a concentration of 1µg/µl was then added and left to incubate at room temperature for up to 30 minutes. More than one DNA complex was added to the mixture at this stage for co-transfection protocols. Following incubation the transfection mix was added onto cells in culture dishes and swirled gently to mix. Transfected cells were cultured at 37°C (5% CO_2) for a specified time dependent on protocol (generally ~24-48 hours).

2.10.0 Harvesting cells for protein and RNA extraction

2.10.1 Protein extraction

Cells cultured on plates were harvested for protein extraction using a range of lysis buffers (table 2.1). Adherent cells were washed and harvested directly from the culture dish. The growth media was firstly aspirated and discarded, followed by two careful washes in 0.1M phosphate buffered saline (PBS). Following removal of the second wash, lysis buffer was pipetted onto the cells and left for up to fifteen minutes at 4°C. A cell

scraper was used to collect the lysed cells, which were then transferred to a micro-centrifuge tube.

Semi-adherent cells were harvested as above, with the additional step of centrifuging growth medium to collect any suspended cells. Medium was removed from the culture dish and centrifuged at ≤1,000xg for 3 minutes. After the supernatant was poured away, the cell pellet was resuspended in lysis buffer and added to adherent cells subsequently collected from the plate.

The cell lysate was centrifuged (4°C) at 14,000rpm for 30 minutes to pellet large cell debris. The supernatant, containing any soluble protein was decanted into a fresh micro-centrifuge tube for further use or storage at -20°C. A sonication step prior to centrifugation was included for release of membrane-bound proteins using a Soniprep 150 sonicator (Wolflabs). This was typically carried out in 10 second bursts (x2) at 10 micron amplitude, with inversion of the tube in-between to mix the cell lysate. The sample was kept cool throughout to minimise the potential for unwanted protein aggregation.

Buffer	Stock solution ingredients (per 100ml)	Storage
RIPA	50mM Tris pH 8, 150mM NaCl, 0.1% SDS, 1mM EDTA, 0.5% deoxycholate, 0.5% IGEPAL, H_2O *	4°C
Laemmli	63mM Tris pH6.8, 2% SDS, 25% Glycerol, H_2O *	4°C
High salt	0.5M NaCl, 0.5M LiCl, 1M KSCN, 0.2% SDS, or 1% Tween 20	4°C
Low det (TNE buffer)	50mM Tris-HCl pH7.5, 150mM NaCl, 1mM EDTA, 0.1% NP40	4°C
Cytobuster	Unknown, purchased from MERCK	Room temp

* Prior to use, 1 protease inhibitor tablet (Complete Mini, Roche) per 10ml was mixed with buffer. This 'working' mixture was then stored at -20°C.

Table 2.1. Cell lysis and IP wash buffer recipes.

2.10.2 RNA extraction

Isolation of total RNA from cell samples was carried out prior to reverse transcription and real-time PCR. Dependent on protocol, extraction of RNA was performed with TRIzol reagent (Invitrogen) or an RNeasy kit (QIAGEN).

Harvesting with TRIzol

Cells grown in a monolayer were lysed directly in the culture dish by adding 1ml of TRIzol reagent per 10cm^2 surface area and pipetting to mix. Homogenized samples were incubated for 5 minutes at room temperature to allow complete dissociation of nucleoprotein complexes. The lysates were transferred to polypropylene tubes, where

0.2ml of chloroform per 1ml of TRIzol reagent was added, before vigorous shaking for 15 seconds and further incubation at room temperature for 3 minutes.

Samples were centrifuged at 12,000xg for 15 minutes (4°C), which serves to phase-separate the mixture into: a lower red phenol-chloroform phase, an interphase, and a colourless upper aqueous phase containing the RNA. The aqueous phase was carefully transferred into a fresh tube.

Isopropyl alcohol (0.5ml per 1ml TRIzol) was added to precipitate the RNA. After incubating for 10 minutes at room temperature and centrifuging (12,000xg) for 10 minutes at 4°C the RNA precipitate formed a visible pellet. Once the supernatant was carefully removed the pellet was washed with 75% ethanol (1ml per 1ml of TRIzol reagent), vortexed, then centrifuged at 7,500xg for 5 minutes (4°C).

Finally, the RNA pellet was air-dried for 5-10 minutes, dissolved in RNase-free water and stored at -70°C.

RNeasy

The RNeasy kit from QIAGEN incorporates silica-based membranes with selective binding properties to purify RNA. The maximum amount of starting material and volumes of reagents used during purification were determined according to

manufacturers' recommendations to ensure spin columns were not overloaded, which reduces the binding capacity and overall RNA yield.

Cells were harvested from culture by trypsinisation (if adherent) and centrifugation in a micro-centrifuge tube for 5 minutes at 300xg. All supernatant was carefully aspirated and discarded. Lysis buffer containing a denaturant to inactivate RNases (RLT) was added to the cell pellet and vortexed to thoroughly mix before the lysate was loaded onto a QIAshredder spin column and centrifuged for 2 minutes at full speed to homogenize.

To provide appropriate binding conditions, 1 volume of 70% ethanol was mixed with the homogenized lysate. Up to 700µl of the sample at a time was transferred to an RNeasy spin column and centrifuged (15 seconds at 8000xg). At this stage total RNA binds to the column matrix and the flow-through is gathered in the collecting tube and discarded.

Several washing steps to remove any unwanted contaminants from the membrane were then performed. Buffer RW1 (700µl) was loaded into the spin column and centrifuged (15 seconds at 8000xg), followed by two washes in 500µl buffer RPE (15 seconds then 2 minutes at 8000xg). A further full speed spin for 1 minute in a fresh tube ensured that any residual buffer was removed. RNA was eluted into a fresh collection tube with 30-50µl RNase-free H_20 spun through the column for 1 minute at 8000xg.

2.11.0 Protein quantification

A Bio-Rad *DC* Protein Assay (Bio-Rad) was used to determine the concentration of solubilised protein. The reaction between alkaline copper tartrate solution and Folin reagent when mixed with protein causes a colour change to develop according to the concentration of protein, which can be measured on a microplate reader.

A working solution A' was firstly prepared by mixing 20µl reagent S to every 1ml of reagent A. Protein standards, ranging from 0.2mg/ml-1.6mg/ml, were also prepared using bovine serum albumin (BSA) diluted in lysis buffer matched to that of protein samples.

Samples and protein standards (5µl) were pipetted onto a 96-well microtitre plate (Bio-Rad) with the well positions carefully recorded. Using a clean pipette tip for each, 25µl of reagent A' was then added to each sample followed by 200µl of reagent B. The reagents were mixed gently to avoid air bubbles and left for 15 minutes for colour to develop.

The absorbances were read at 750nm on a Bio-Rad Microplate Absorbance Reader with Microplate Manager software. A standard curve calculated from the BSA standards was used to extrapolate the relative concentration of experimental samples.

2.12.0 RNA/DNA quantification

Quantitation of DNA/RNA was performed on a Nanodrop 2000 spectrophotometer with accompanying software (Thermo Scientific), which performs efficient automated measurement of small volume samples. This instrument exposes samples of nucleic acids to ultraviolet light at a wavelength of 260nm and a photo-detector measures the light that passes through. The concentration of nucleic acid in the sample is proportional to the amount of light that is absorbed. As single and double-stranded DNA and RNA have different absorbance characteristics, a conversion factor for each may be equated according to the Beer-Lambert Law. Sample purity may also be determined by calculating the ratio of the absorbance at 260 and 280nm ($A_{260/280}$): The absorbance ratio for pure DNA is $A_{260/280}$ = ~1.8 and for pure RNA $A_{260/280}$ = ~2.

Prior to sample measurement, 2µl of ddH_2O (or control fluid) was loaded onto the apparatus pedestal for an instrument self-calibration check. With careful blotting of the pedestal in-between, samples (2µl) were then individually measured and recorded.

2.13.0 Polyacrylamide gel electrophoresis and Western-blotting

Proteins were size separated for identification using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Proteins prepared for electrophoresis

were solubilized in SDS detergent, which denatures secondary and tertiary structures, and applies a negative charge to protein proportional to its mass.

2.13.1 Preparation of polyacrylamide gel

Polyacrylamide gel was made with polyacrylamide (10-15%), 375mM Tris pH8.8, 1% SDS, 10% ammonium persulfate (APS), 0.05% TEMED (tetramethylethylenediamine) and ddH_2O. Clean gel plates were filled just over three-quarters, topped up with ddH_2O, and left to polymerize. Once set, the water was poured away and replaced with stacking gel (3% polyacrylamide, 50mM Tris pH6.8, 0.08% SDS, 10% APS and 0.06% TEMED). A comb was inserted into the stacking gel to create sample wells and left to set. After careful removal of the comb, the wells were washed with water and blotted dry with Whatman paper. The prepared gel plates were loaded into an electrophoresis tank, which was filled with running buffer (25mM Tris base, 190mM glycine, 0.1% SDS, pH 8.3).

2.13.2 Loading samples and running gel

Protein samples (whole cell lysates) were routinely loaded onto SDS polyacrylamide gels in total volumes of 20µl. These comprised protein extracts in lysis buffer (table 2.1) mixed with x2 loading buffer (4% SDS, 10% 2-mercaptoethanol, 20% glycerol, 0.004% bromophenol blue, 0.125M Tris HCl, pH 6.8). Additional preparation steps varied in accordance with the solubility/nature of the proteins involved. These included sonication, syringe-homogenisation, centrifugation, and boiling.

A protein marker ladder (Fermentas), as well as the protein samples were pipetted into designated wells, before attaching the electrodes and lid onto the gel tank. The samples were run at 150 volts for 1-2 hours, or until the size markers of interest were sufficiently separated.

2.13.3 Western blotting

Following electrophoresis gels were transferred to holder cassettes together with a piece of Hybond-P nitrocellulose membrane. These were sandwiched between several support layers of sponge and filter paper, pre-soaked in transfer buffer (25mM Tris, 192mM glycine, and 20% methanol). A foam pad and 3 layers of Whatman (3mm) filter paper were firstly placed onto the cassette, followed by the gel. The membrane was activated by immersion in methanol and positioned over the surface of the gel with care to ensure air bubbles were smoothed out. A second layer of filter paper (x3) and foam were placed on top, and the entire stack was secured into the cassette. The cassettes were loaded into an electrophoresis tank, with the membrane orientated toward the cathode. An ice cooler was also placed into the tank before filling the tank with transfer buffer, and securing the electrodes. The transfer was run at 80 volts for 1-1½hours.

After transfer, membranes were incubated with 5% dried milk in PBS-tween buffer (2mM Na_2HPO_4, 0.5mM KH_2PO_4, 1.3mM KCl, 135mM NaCl, 0.05% Tween 20, pH 7.4), which served as a blocking solution prior to washes with antibodies. Membranes were inserted into 50ml polypropylene tubes, lying flat against the surface of the tube with

attached proteins facing inwards. The tube was filled with blocking solution and placed on a tube roller for 1 hour at room temperature.

Thereafter, blocking solution was poured away and replaced with 5ml of 5% dried milk in PBS-tween containing specific primary antibodies (table 2.2) diluted according to requirements. This was incubated on a tube roller at 4°C overnight.

The following day, membranes were washed three times (5 minutes each) in PBS-tween on an orbital shaker and replaced to a 50ml polypropylene tube. Specific secondary antibodies (table 2.2) were diluted according to requirements in 5ml of 5% dried milk in PBS-tween, added to the tube, and incubated (room temperature) for 1 hour on a tube roller.

Unbound secondary antibodies were washed away with four washes (15 minutes each) of PBS-tween, on an orbital shaker.

Antibody	Raised in	Species reactivity	Epitope (residues)	Detection	Manufacturer/source
Anti-*β-actin* (monoclonal)	mouse	Human, mouse, rat	*ACTB* N-terminus	β-cytoplasmic actin N-terminal peptide	Sigma
Anti-*CLPS* (polyclonal)	Mouse	Human	Full length *CLPS*	Human colipase	Abcam
Anti-*c-myc*, clone 9E10 (monoclonal)	Mouse	Epitope-tagged proteins	410-419	Human oncogene product *c-myc* (myc epitope)	Sigma
Anti-*c-myc* (polyclonal)	Rabbit	Epitope-tagged proteins	408-425	Human oncogene product *c-myc* (myc epitope)	Sigma
Anti-*HA*, clone 12CA5 (monoclonal)	Mouse	Epitope-tagged proteins	98-106	Human Influenza virus Hemagglutinin (HA epitope)	Roche
Anti-*HA*, clone HA-7 (monoclonal)	Mouse	Epitope-tagged proteins	98-106	Human Influenza virus Hemagglutinin (HA epitope)	Sigma
Anti-*HA* (polyclonal)	Rabbit	Epitope-tagged proteins	98-106	Human Influenza virus Hemagglutinin (HA epitope)	Sigma
Anti-*INS* (monoclonal)	Mouse	Human, mouse, rat	Full length *INS*	Insulin	Invitrogen
Anti-mouse	Rabbit	Mouse	Mouse IgG	Mouse immunoglobulin	Invitrogen
Anti-mouse tagged with Alexa Fluor 594	Goat	Mouse	IgG heavy chains and all classes of mouse immunoglobulin light chains	Mouse immunoglobulin	Invitrogen
Anti-*PDI* (monoclonal)	Mouse	Human, mouse, rat	*PDI* C-terminus	Mammalian protein disulfide-isomerase	Stessgen
Anti-rabbit	Goat	Rabbit	Rabbit IgG	Rabbit immunoglobulin	Invitrogen
Anti-rabbit tagged with Alexa Fluor 488	Goat	Rabbit	IgG heavy chains and all classes of rabbit immunoglobulin light chains	Rabbit immunoglobulin	Invitrogen
Anti-*WFS1* (polyclonal)	Rabbit	Human, mouse	1-314	Wolframin	ProteinTech
Anti-*WFS1* (polyclonal)	Rabbit	Mouse, rat	100-200 mouse *WFS1*	Wolframin	Abcam
Anti-*WFS1* (polyclonal)	Rabbit	Human, mouse	*WFS1* N-terminus (1-285)	Wolframin	Dr. S., Hofmann, Institute of Diabetes Research, Academic Hospital Munich-Schwabing, Koelner Platz 1, D-80804 Munich, Germany

Table 2.2. Comprehensive list of antibodies used for: immunocytochemistry, immunohistochemistry, western blotting and immunoprecipitation.

2.13.4 Developing western blots

Western blots were developed using a chemiluminescent detection system, which is the emission of light and limited amounts of heat as a result of a chemical reaction. Secondary antibodies conjugated to horseradish peroxidase (HRP) were used as probes in immunoblots, which catalyses the oxidation of luminol in the presence of hydrogen peroxide. The light produced by luminol-based systems can be detected on radiographic film at short-exposure and ambient temperature.

Prepared membranes were placed protein-side up onto Saran wrap, which was spread smoothly over a flat bench. An ECL Plus detection kit (GE Healthcare) which uses an acridan-based substrate to produce a chemiluminescent signal was used to label western blots for development. Reagent A (25µl) mixed with reagent B (1ml) was pipetted evenly over the surface of each membrane and left for five minutes. Excess developing fluid was displaced by tapping the edge of the membrane to the bench, and the membrane was covered smoothly in Saran wrap. Wrapped membranes were placed into a developing cassette, protein-side upward, and secured at the corners with tape.

In a dark room, X-ray film (Kodak) inserted into the cassette was exposed to the membranes for time periods ranging from 1 second to several hours. Exposure times were adjusted according to signal intensity as per requirements before films were developed in an automated X-ray film processor.

2.13.5 Quantitative analysis of protein expression values

A lightbox and high resolution camera connected with Genetools software (Syngene) was used for quantitative analysis of expression values following western blot. After image capture (GeneSnap, Syngene) of developed blots, absorbance measurements (arbitrary units) were taken of individual bands/protein samples. Measurements were normalized firstly to background signal from the X-ray film, then to a control housekeeping protein (β-actin), measured from the same blot, to account for discrepancy in sample loading. Where appropriate, endogenous background protein signals were also measured and factored into the final measurements.

2.14.0 Co-Immunoprecipitation

Co-Immunoprecipitation was employed as an experimental measure to detect protein interactions *in vitro*. The principle of this method is to target and purify antigens and their potential binding partners from a protein cell lysate using specific antibodies. The 'precipitated' immune complexes can be subsequently analysed by SDS-PAGE electrophoresis and western blotting (section 2.13).

The first stage in this procedure involves the introduction of specific antibodies to whole cell lysates allowing immune complexes to form between the antibody and target antigen. Secondly, agarose beads (Sigma) with immobilized protein A are introduced into the mixture as a solid-phase support. Protein A has binding affinity for the Fc region

of immunoglobulins through interaction with the heavy chain. This enables protein-antibody complexes to be captured from the cell lysate. The assumption at this stage is that binding partner-proteins will be precipitated together with the targeted antigen. Any unbound molecules can then be washed away, before eluting the precipitated complexes and verifying the identity of the 'pulled-down' partner-proteins.

2.14.1 Pre-clearing cell lysates

Prior to immunoprecipitation, all protein lysates were pre-cleared with uncoated agarose beads. This step allows proteins with affinity to the surface of the agarose to be bound and removed, thus reducing later non-specific 'pull-downs'.

In 1.5ml polypropylene micro-centrifuge tubes, beads were prepared as standard in 40µl volumes, and equilibrated in 400µl PBS. The suspension was quickly vortexed and centrifuged (13,000rpm) for 2 minutes to pellet the beads. The supernatant was aspirated and discarded with a fine-tipped pipette taking care to cause minimum disturbance to the beads.

Cell lysis buffer (400µl) was added to the equilibrated beads and tubes were loaded onto a rotary mixer for 30 minutes at 4°C. The mixture was centrifuged (13,000rpm) afterwards for 5 minutes to pellet the beads, and the supernatant carefully aspirated and discarded.

Cell lysates (400µl) were added to prepared beads and placed on a rotary mixer for 1 hour at 4°C. The pre-cleared lysates were finally harvested by centrifugation at 13,000rpm for 5 minutes, using a fine-tipped pipette to carefully remove the purified supernatant. Aliquots of lysate samples (~30µl) were taken and stored at -20°C to provide controls for immunoprecipitation inputs.

2.14.2 Immunoprecipitation

Suitable antibodies (table 2.2) (~10µg) were added to pre-cleared protein lysates and incubated overnight at 4°C on a rotary mixer. The following day, fresh agarose protein A beads (40µl) were equilibrated in PBS and lysis buffer as described in section 2.14.1. The protein/antibody composites were incubated with the beads for 2 hours on a rotary mixer (4°C) to allow capture of the complexes onto the surface.

After this incubation period, the entire contents of the tubes were loaded into Spin-X columns (Sigma). Effective separation of supernatant and beads takes place during centrifugation for 2 minutes at 13,000rpm. The supernatant flows through a micro-porous membrane, whilst beads are retained within the column.

Removal of non-specific bound molecules from the agarose beads is assisted by a series of washes in lysis buffer. Buffer (400µl) was added to the column with the beads, incubated for a period, and washed-through during centrifugation. Standard protocols

included three washes, with incubations of 10 minutes on a rotary mixer, although stringency of washing was increased in instances of high non-specific protein binding.

2.14.3 Eluting 'pulled-down' protein

Protein-complexes were eluted from protein A beads in x5 sample loading buffer suitable for SDS-PAGE electrophoresis (10% SDS, 10mM 2-mercaptoethanol, 20% glycerol, 0.0004% bromophenol blue, 0.2M Tris-HCl, pH 6.8). Spin-X columns containing washed beads were transferred to clean 1.5ml micro-centrifuge tubes. The beads were covered with 30µl loading buffer, and heated for 5 minutes at 95°C. Loading buffer containing eluted proteins was finally separated from the beads in a 2 minute centrifugation at 13,000rpm, and stored at -20°C until required for analysis.

2.14.4 Detection of precipitated protein-complexes

Protein complexes were detected by standard SDS-PAGE and western blotting (section 2.13). In addition to 'pull-down' experimental samples, positive control lysates collected after the pre-clearing step (section 2.14.1) were included to confirm samples did contain the proteins of interest. Further controls were also carried out to optimise the protocol at various stages; collection of samples from IP flow-through, wash steps and post-elution bead-boil confirmed efficiency of antibody/antigen binding and elution.

2.15/2.16 Immunostaining

Immunostaining techniques were used to determine the presence and localisation of specific antigens in biological tissue, in whole cell mounts (immunocytochemistry) or in tissue sections (immunohistochemistry). Antibodies (table 2.2) conjugated to fluorophores were used to target specific proteins/structures, which could be visualised under a fluorescent microscope.

2.15.0 Immunocytochemistry (ICC)

2.15.1 Cell culture

Cells were chiefly cultured and transfected in 10cm tissue culture dishes according to standard procedure (section 2.9). Sterile cover slips (22 x 22mm) were placed into culture dishes before seeding with cells, which were swirled gently to ensure even distribution and adherence.

2.15.2 Fixation

Following a suitable incubation period, the dishes were removed from the 37°C incubator and allowed to cool for 5 minutes. The cover slips were then gently lifted from the dish with sharp-edged tweezers, and placed into 24-well plate chambers with cells facing upwards. To fix and permeabilise the cells, each slide was covered with 0.5ml paraformaldehyde (4%) for 15 minutes, which serves as a cross-linking agent to

preserve cell structure. Once fixed, the cells were washed three times for 5 minutes in PBS (phosphate buffered saline: 137mM NaCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4 and 2.7mM KCl in H_2O at pH 7.4).

2.15.3 Staining

Cells were incubated for 10 minutes in a solution containing 3% bovine serum albumin (BSA) and 0.1% Triton-X in PBS. This was used as a blocking agent against later non-specific antibody binding.

Antibody diluting buffer was made up with 3% BSA and 0.05% Tween-20 in PBS. Primary antibodies (table 2.2) were diluted with buffer to an optimal concentration (anti-WFS1, 1/100; anti-PDI, 1/350; anti-c-myc, 1/100) and applied to the coverslips for 1 hour at room temperature.

Unbound primary antibodies were washed away with 3 washes in PBS (5 minutes each), followed by the application of secondary antibodies (table 2.2), diluted in buffer (1/200). Alexa fluor dyes (Invitrogen) were the fluorescent probes of choice, since they have been proven more photostable, less pH sensitive, and brighter than other standard dyes of comparable excitation and emission. Coverslips were incubated in secondary antibodies at room temperature, covered with foil and kept in the dark when possible henceforth. After 1 hour, excess secondary antibodies were removed with three washes in PBS (5 minutes each).

2.15.4 Mounting

Coverslips were mounted onto microscopy slides with DAPI (4',6-diamidino-2-phenylindole). DAPI is excited under UV light producing a blue-cyan emission (~461nm), and binds strongly to DNA, so can be used as an effective cell nucleus marker.

Small drops of DAPI mountant (Vector Labs) were placed onto clean glass microscopy slides, with a pasteur pipette. Coverslips were then gently lifted from plate wells with sharp-edged tweezers and placed with cells facing downwards over the DAPI, taking care to avoid underlying air bubbles. The slides were labelled accordingly and stored at 4°C, allowing the mountant to set prior to viewing.

2.16.0 Immunohistochemistry

2.16.1 Tissue preparation

Rat pancreata, provided courtesy of Dr S Jacques, Molecular Neuroscience Group, University of Birmingham, were excised from animals euthanized by CO_2 narcosis. Tissue was immediately immersed in increasing gradients of sucrose in PBS (10%, 20% and 30%) until observed to sink, embedded in OCT mounting medium (Thermo Scientific), rapidly frozen on dry ice, and stored at -80°C.

A Bright 5030 Cryostat (Bright) was used to cut 15µm sections of tissue, which were mounted onto charged slides (Surgipath) and stored at -20°C.

2.16.2 Fixation and staining

Frozen sections were warmed to room temperature then fixed in absolute ethanol for 1 minute. The slides were placed into a Coplin jar and washed twice (5 minutes each) in PBS followed by a 10 minute wash in PBS with 0.1% Triton-X-100 to permeate the tissue.

The sections were isolated using a hydrophobic pap-pen (Vector Labs). By drawing around the tissue a heat-stable, water repellent barrier was created, which keeps staining reagents localised to a particular section, allowing more than one section per slide and necessity for much less antibody per section.

Slides were placed in a darkened humidified chamber at room temperature for 30 minutes covered (100µl per section) in blocking solution (1% bovine serum albumin, 0.1% Triton-X-100 in PBS). Following gentle aspiration of blocking solution primary antibodies (table 2.2), diluted to an appropriate concentration in buffer (3% BSA and 0.05% Tween-20 in PBS) were then applied to each section (100µl) and incubated in the chamber overnight (16 hours) at 4°C.

PBS was used to wash the slides three times (5 minutes each) in a Coplin jar. After aspiration of the PBS, secondary antibodies (table 2.2) diluted in buffer to an

appropriate concentration, were applied (100µl) to each section and incubated for 1 hour in the chamber at room temperature.

Unbound secondary antibodies were washed off with 3 x 5 minute PBS washes in a Coplin jar. Excess fluid was gently aspirated from the sections, a drop of DAPI mountant (Vector Labs) was applied over the tissue and a coverslip lain on top. The prepared slides were stored in the dark at 4°C allowing the mountant to set ready for viewing.

2.17.0 Fluorescent microscopy

An Axioplan 2 microscope with Zeiss Neofluar lenses, an AxioCam HRC Camera and Axiovision software (Carl Zeiss Ltd) was used to view slides and capture images. To increase the resolving power at high magnification (x1000) immersion oil was placed between the specimen and the objective lens. In these instances, the edges of coverslips were sealed with varnish to ensure they remained in position.

2.18.0 Measuring protein stability with cyclohexamide

Cyclohexamide (CX), a biochemical produced by bacterium *Streptomyces griseus*, was used to measure the relative differences in protein half-lives between wild-type and variants of wolframin. CX is a known inhibitor of eukaryotic protein synthesis and has

been shown to block the elongation phase of protein translation by binding to ribosomes and inhibiting eEF2-mediated translocation (Obrig, 1971).

HEK293 cells were seeded onto 10cm plates; for each experimental condition/sample, one plate was allocated for each time interval. Cells were grown to ~80% confluency (section 2.9.0) and transfected (section 2.9.5) with constructs containing wild-type *WFS1* and *WFS1*-variants. After 24 hours incubation, CX in 200µl pre-warmed DMEM growth medium (table 2.0) was added to all plates to give a final concentration of 100µM, which were swirled to mix taking care not to disturb the cells. One plate from each condition/sample set was harvested in RIPA lysis buffer (section 2.10.1, table 2.1) immediately following CX treatment (designated '0 hours') and stored at -80°C. The remaining plates were left to incubate (37°C) and harvested at subsequent time intervals according to requirements.

Once all plates were harvested, measurements of protein concentration in each sample were made (section 2.11), and the samples were loaded equally onto SDS-polyacrylamide gels (section 2.13.0) and electrophoresed for western blotting (section 2.13.3). Quantitative analyses of protein expression values were performed using Genetools/GeneSnap software (Syngene; section 2.13.5). The stability of each wolframin variant was calculated by comparing the expression values at a given time interval relative to that at '0 hours', designated 1 (100%).

2.19.0 Molecular Cloning

Molecular cloning techniques (summarised in fig. 2.3) were used to generate multiple copies of defined DNA sequences, within vectors suitable for inducing protein production. Sequences relating to genes/fragments of interest were initially sourced from Geneservice (Source Bioscience) or stocks available in the Medical and Molecular Genetics Dept., University of Birmingham. Basic cloning principles involve; cutting both vector and insert of interest with matched restriction enzymes, ligating the insert into the linearised vector, transforming the new 'clone' to competent bacteria to enable rapid, reliable multiplication, then finally selection of successful transformants followed by thorough checks to confirm clones contain the correct insert sequence in the correct orientation.

2.19.1 Obtaining and verifying starting material

Where DNA fragments and gene sequences of interest were available from plasmid stocks in the laboratory, full sequencing verification was firstly carried out. Alternatively new 'IMAGE' clones were purchased from Geneservice. These derive from a collection of vectors containing cDNAs representing expressed genes from various organisms. Sequence verified clones containing full open reading frames were selected from the Mammalian Gene Collection. These were delivered in *E.coli* cells growing on agar slopes containing selection antibiotics and were stored at 4°C on arrival.

IMAGE clones were grown up on agar plates containing appropriate selection antibiotics for specified vectors (section 2.7.2). An inoculation loop was lightly scraped across the agar slope to pick up cells, and streaked over the surface of an agar plate. The inoculated plate was left to incubate for >16 hours at 37°C. Following this period, individual colonies were picked and seeded into 5ml LB broth (with antibiotic), and cultured overnight at 37°C in an orbital shaker. Cells were collected from the culture suspension the following day by centrifugation at ~3,000rpm for 5 minutes, which served to pellet the cells. After decanting supernatant, plasmids were extracted and purified using a Qiagen Spin Miniprep Kit (section 2.7.5).

The purified DNA was sequenced (section 2.4) with primers spanning the start site and 3'-end of the insert, if full IMAGE cloning data was not provided. Restriction sites surrounding the insert were verified for subsequent subcloning steps.

2.19.2 PCR amplification of DNA fragments/genes

PCR amplification of the desired sections was completed with appropriate primers (section 2.1.1), electrophoresed on agarose gel and purified (section 2.3.0-1).

Additional restriction sites could be added to existing sequences during the PCR reaction if required. The primers for this purpose were designed according to standard protocol (section 2.1.1) with additional features: restriction sites were added to the 5'-end of a

primer in the appropriate orientation, and a polylinker sequence was also introduced upstream of the restriction site to provide structural support for the restriction enzyme when binding and cutting DNA (fig. 2.1A).

Only sequence matched bases (not including restriction site and polylinker) were included for calculation of primer melting temperature (Tm). Further consideration was given to fragments destined for expression vectors containing an N-terminal tag. Dependent on the reading-frame of the tag sequence, one or two additional nucleotides were added to the primer sequence to ensure ligation of the new gene in the correct reading frame (fig. 2.1B).

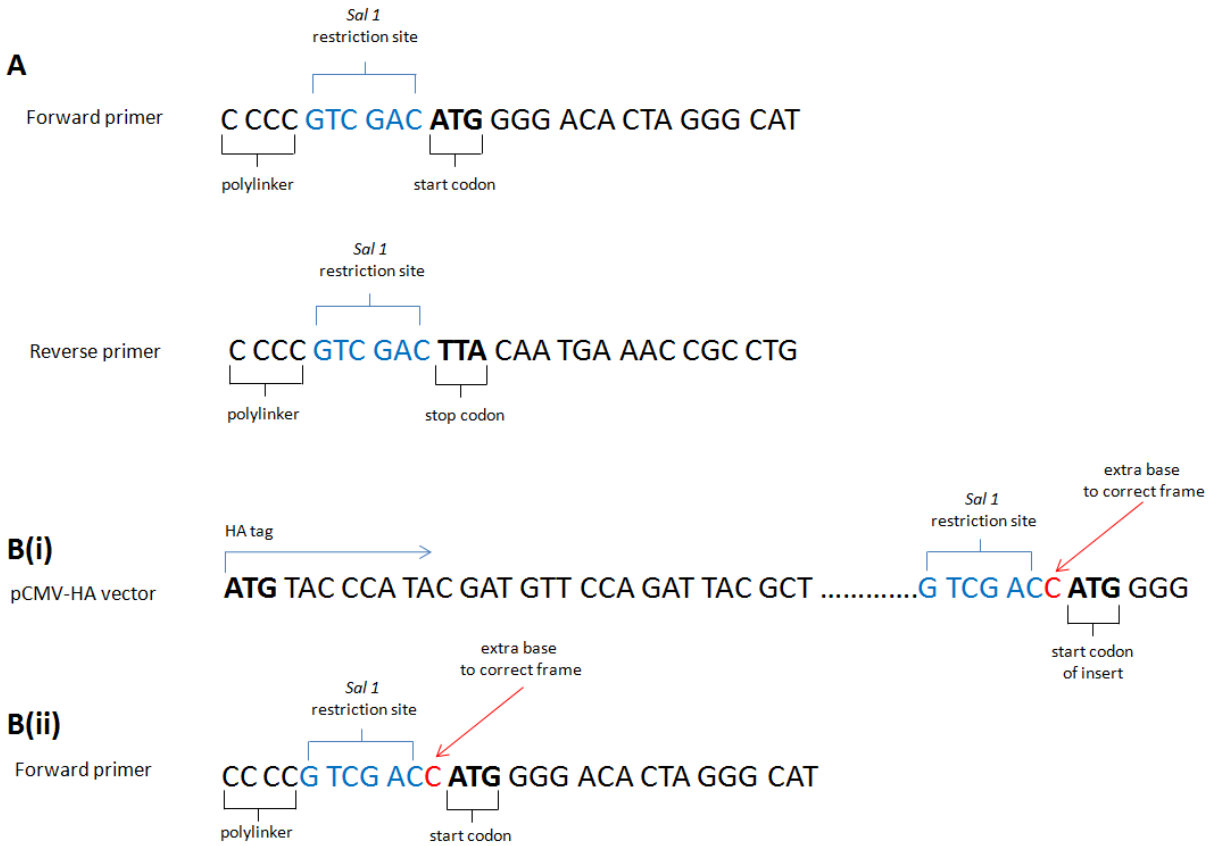


Figure 2.1. Primer design for insertion of restriction sites during PCR. (**A**) Example primer design for addition of *Sal 1* restriction site to include polylinker nucleotides to stabilise enzyme/DNA binding for cutting. (**B**) Primer design for insertion of DNA into tagged (N-terminal) expression vectors.(**i**) Example sequence of pCMV-HA vector detailing reading frame and necessity for extra nucleotide to maintain correct reading frame of cloned gene. (**ii**) Primer design including extra nucleotide and polylinker sequence.

2.19.3 Restriction digestion/ligation

Transference of cDNA fragments from one vector to another was achieved using an enzyme digestion/ligation technique. The full insert sequence was initially mapped for the presence of internal restriction sites using Webcutter (Heiman, 1997). With reference to publicly available vector maps or sequencing data generated in the laboratory, compatible restriction sites were then identified in both 'donor' (containing insert) and recipient plasmids. When selecting appropriate restriction enzymes, consideration was also given to the efficiency of enzymes in a matched buffer during double-digestion, and the possibility of creating unwanted complementary DNA overhangs that could anneal incorrectly or in the wrong orientation.

Inserts were cut from a donor vector or purified from a PCR reaction; recipient vectors were linearised during digestion. This was typically carried out in 50µl reaction volumes comprising: purified plasmid DNA, BSA, restriction enzymes, and suitable buffer according to specific manufacturers' recommendation. Reactions were incubated at 37°C for 2 hours. The digested products were electrophoresed on agarose gel (section 2.3.0) to separate fragments according to their respective sizes. Using an appropriate reference DNA size ladder (Invitrogen) DNA fragments/linearised vectors were distinguished, excised from the gel under UV light, and gel purified (section 2.3.1).

Ligation of purified fragments into linearised plasmids was achieved by incubating the DNA products together with T4 DNA ligase, and buffer containing adenosine tri-

phosphate (ATP), an essential cofactor. DNA ligase, derived from the T4 bacteriophage catalyses the formation of phosphodiester bonds between juxtaposed 5' phosphate and 3' hydroxyl termini in duplex DNA. This enables digested DNA with complementary 'free' ends to be joined together. Reactions were carried out overnight at 4°C in 0.5ml polypropylene tubes, and routinely comprised; 0.1-1μM DNA, 4μl ligation buffer (x5), T4 DNA ligase (3 units), and ddH_2O up to a final volume of 20μl. Optimisation of ligation was carried out if necessary with various insert: vector ratios, using the formula below as a guideline:

$$\frac{\text{kb size of insert x ng of vector}}{\text{kb size of vector}} = \text{1: 1 ratio}$$

2.19.4 Intermediate vectors

The pGEM-T vector system (Promega) and CloneJet PCR cloning kit (Fermentas) were used as high efficiency intermediate vectors in most PCR cloning applications. These linearised vectors allow straightforward TA ligation of PCR products, which can then be subsequently amplified and subcloned into expression vectors as desired. Available multiple cloning sites (MCS) useful for subsequent steps largely influenced selection of intermediate vectors (fig. 2.2).

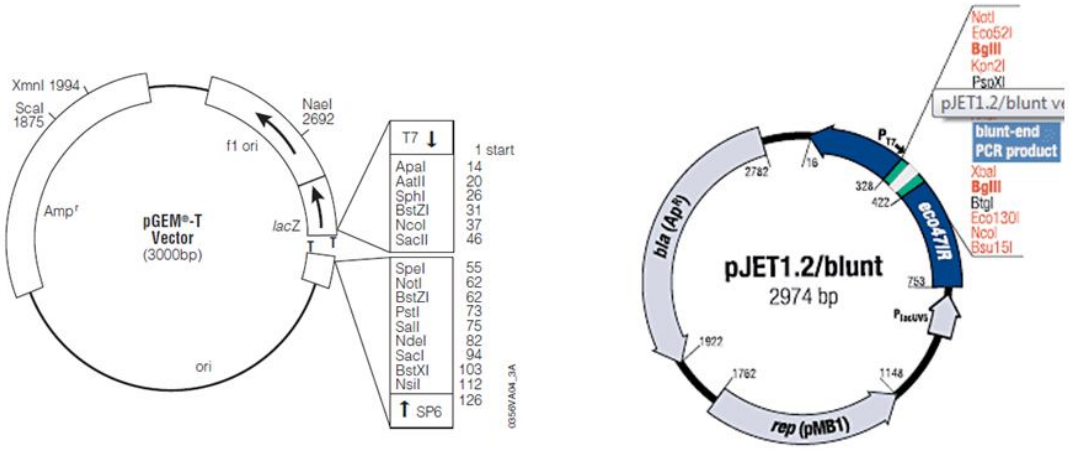


Figure 2.2. Intermediate cloning vector maps for pGEM-T (Promega) and pJET1.2 (Fermentas).

pGEM-T

The pGEM-T vector system was suitable for direct ligation of PCR products generated with *Taq* DNA polymerase. Ligation reactions were incubated for 1 hour at room temperature and comprised: 5µl rapid ligation buffer (x2), 1µl pGEM-T vector (50ng), insert DNA (variable), T4 DNA ligase (3 units), and DNA/nuclease-free H_20 up to a final volume of 10µl. Molar ratios of insert to vector were begun at 1:1, but optimised using the following formula as a guideline according to protocol:

$$\frac{\text{ng of vector x kb size of insert}}{\text{kb size of vector}} \quad x \quad \text{insert: vector molar ratio} = \text{ng of insert}$$

A-tailing reaction

PCR products generated with proofreading enzymes such as *Pfu* DNA polymerase produce 'blunt-ended' products. These require an additional *A-tailing* modification prior to ligation with pGEM-T to add overhanging adenosine residues. A 10µl reaction including: 1-7µl DNA, 1µl *Taq* DNA polymerase buffer with $MgCl_2$ (x10), dATP to a final concentration of 0.2mM, 5 units of *Taq* DNA polymerase, and purified H_20, was mixed and incubated for 15-30 minutes at 70°C. Up to 2µl of this reaction mixture was used for subsequent ligation.

CloneJET

The CloneJet PCR cloning kit (pJET1.2 vector) is suitable for direct ligation of PCR products generated with proofreading *Taq* polymerase, such as *Pfu*, which produce blunt-ended products. Ligation reactions comprised: 10µl reaction buffer (x2), 0.15 pmol DNA, 50ng pJET1.2 cloning vector, T4 DNA ligase (5 units) and nuclease-free H_2O to a total volume of 20µl. The mixture was vortexed briefly before incubation for 5 minutes at room temperature.

Blunting reaction

Sticky-ended PCR products generated with non-proofreading DNA polymerases such as *Taq*, required removal of unwanted adenosine overhangs prior to ligation with pJET1.2 vector. A 'blunting' reaction to remove the overhanging nucleotides comprised: 10µl reaction buffer (x2), 0.15pmol DNA, 1µl DNA blunting enzyme (Fermentas) and purified

H_20 up to a final volume of 20μl. Following incubation for 5 minutes at 70°C the reaction was chilled on ice. Directly to the blunting reaction, 5 units of T4 DNA ligase and 50ng pJET1.2 vector were added, vortexed, then incubated at room temperature for 5 minutes to ligate.

2.19.5 Transformation, selection and small-scale culture plasmid extraction

Chemically competent cells were transformed (section 2.7.4) with pUC19 control plasmid and products of ligation reactions. Up to 10μl of ligation mix was incubated with 100μl α-select silver efficiency competent cells (Bioline), which produce >1x10^8 cfu/μg pUC19. Cells with increased efficiency (Bioline α-select gold, ~10^9) were used to increase success of problematic transformations.

Recovered cells were grown overnight on suitable selective plates at 37°C, and several of the resulting colonies were independently picked and seeded for >16 hours in 5ml of selective LB broth. Plasmids were recovered and purified from clonal cultures using the Qiagen Spin Miniprep kit (section 2.7.5).

2.19.6 Verification of clones

The purified plasmid DNA preps were checked for the presence of the correct insert sequence. Using suitable enzymes each prep was digested (section 2.2) and electrophoresed (section 2.3.0). Clones exhibiting the correct digestion pattern were then sequence verified if required (section 2.4).

2.19.7 Re-transformation, selection and large-scale culture plasmid extraction

Clones containing the correct sequence inserts were re-transformed (section 2.7.4) into *E.coli*. A typical reaction included 50µl of sub-cloning efficiency chemically competent DH5α cells (Invitrogen) mixed with ≤1µg plasmid DNA (miniprep). A 1-5ml starter culture was inoculated to 100ml of selective LB broth and grown for ~16hours before plasmid extraction using the QIAGEN Maxiprep kit (section 2.7.5). Purified plasmids were again sequence verified (section 2.4) prior to transfection or further manipulation.

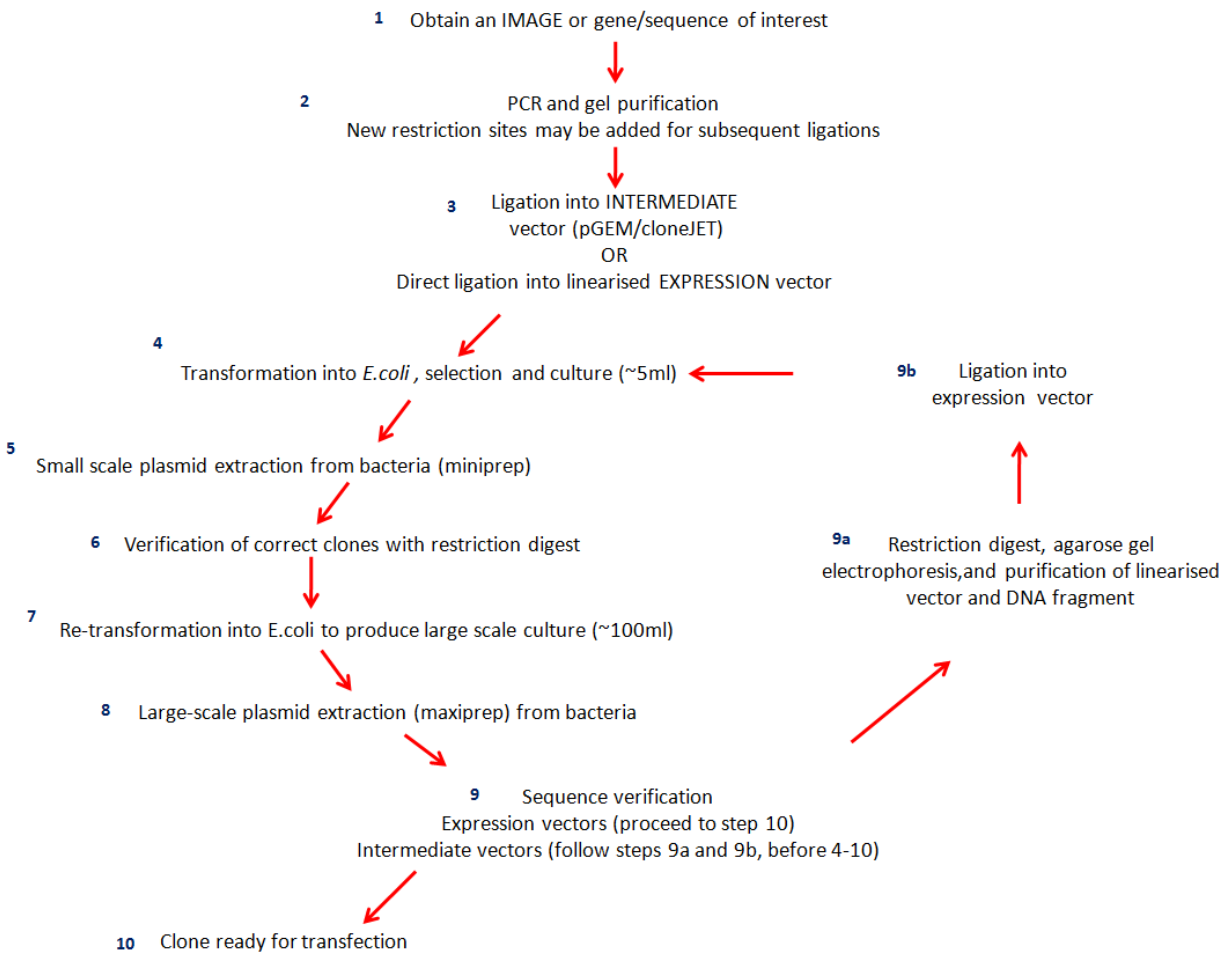


Figure 2.3. Principle steps in molecular cloning

2.20.0 Bioinformatics

Publicly available resources were accessed throughout including: gene and protein annotation databases, sequence alignment and primer design software, and predictive analysis tools (table 2.3).

Programme	Application	Reference
BioEdit	Sequence alignment	www.mbio.ncsu.edu/bioedit/bioedit.html
Ensembl Genome Browser	Comprehensive annotation on eukaryotic genomes	www.ensembl.org
NCBI	Biotechnology information resource	http://www.ncbi.nlm.nih.gov/guide/
Primer3	Primer design software	http://frodo.wi.mit.edu/primer3/
Quikchange Primer Design	Mutagenesis primer design software	www.stratagene.com/qcprimerdesign
UCSC	Genome browser for reference sequences	http://genome.ucsc.edu/
UniProt	Protein sequence databank	www.uniprot.org
Webcutter 2.0	Mapping restriction sites	http://bio.lundberg.gu.se/cutter2

Table 2.3. Bioinformatics resources.

CHAPTER 3

MUTATION ANALYSIS OF WOLFRAM SYNDROME PATIENTS

3.0 Introduction

Collection of molecular genetic data is the foundation for understanding any inherited disorder. From a clinical perspective, knowledge of the genetic mutations carried by patients is essential for both postnatal and prenatal genetic diagnosis, and carrier detection. Defining the exact nature of disease causing mutations is also often useful for planning treatment and preventing development of complications and may provide access to gene therapy should this be available.

In the laboratory, identifying and associating genetic mutations to phenotype forms the basis for elucidating complex pathogenic mechanisms at the protein and cellular level. Comparing specific mutation to protein function can help us to identify structurally and functionally important protein regions, sometimes relevant to a particular pathway or system. In the Wolfram gene for example, heterozygous A716T and K863T variants are among a growing number of *WFS1* alleles associated with non-syndromic hearing loss (Bespalova, 2001: Young, 2001; Fujikawa, 2010), suggesting these regions of the protein are particularly important for auditory function. Other *WFS1* variants have been linked specifically with susceptibility to diabetes (Sandu, 2007; Franks, 2008; Fawcett, 2010) and psychiatric illness (Sequeira, 2003; Koido, 2005; Swift, 2005).

In Wolfram syndrome (WS), a small number of families from regions in the Middle-east (Amr, 2007) have been found to harbour a homozygous missense mutation in *WFS2* (*CISD2*/*ZCD2*), although the majority of cases are attributed to mutations in *WFS1*. A wide range of mutations have been described throughout the *WFS1* gene with no apparent mutational hotspots. Previous attempts to correlate genotype and phenotype have been limited by small sample numbers, although there does appear to be some relationships between mutation type and disease severity. Several reports suggest that clinical expression of WS is more complete in patients harbouring inactivating mutations only (Smith, 2004; Cano, 2007; D'Annunzio, 2008) and presentation of clinical features is likely to occur at a relatively earlier age (Cano, 2007). Patients carrying one or more missense variants may present with a relatively milder phenotype (Van Den Ouweland, 2003; Gasparin, 2009; Zymslowska, 2011).

The West Midlands Regional Genetics Laboratory provides a genetic testing service primarily for National Health Service patients within the region, but also carries out testing for private consultants and for healthcare institutions nationally and internationally. Part of the service offered is genetic diagnosis for Wolfram syndrome and the patient registry currently lists approximately 100 individuals, referred to the clinic over the last 20 years. The testing methods include MLPA (multiplex ligation-dependent probe amplification) to identify large chromosomal abnormalities and mutation detection by direct sequencing analysis. At the time of this study a number of listed families (26 probands) had only one heterozygous, or no *WFS1* mutations identified.

3.1 Specific aims and approach

We aimed to complete the Wolfram patient database by screening all suitable patients for *WFS1* mutations, before looking to identify potential mutations in other candidate genes. Following the mutation screen we planned to add our data to the larger West Midlands cohort and search for any genotype-phenotype correlations. We hoped to extend and update the current mutational spectrum for WS by completing the mutational and genotype-phenotype analysis of the largest single cohort to date.

DNA samples from WS patients were obtained from the Regional Genetics Laboratory and a *WFS1* mutation screen was performed by PCR (2.1.0-2.1.1) and direct sequencing (section 2.4.0). Primers (appendix 2.1) spanned all coding exons (2-8) and splice sites; exon 8 (2609 base pairs) was sequenced in 5 overlapping sections (A-E).

Any patients found to be negative for *WFS1* mutations were then screened for possible mutation of *WFS2*, then for mutations in *ATP1B1*, described in section 3.3.0. Identified mutations were confirmed by re-sequencing both forward and reverse complements, and novel variants were also sequenced in a panel of normal control chromosomes.

Results

3.2.0 Mutation analysis of *WFS1*

Following full sequencing of the *WFS1* coding sequence, a total of 10 mutations were identified in 8 individuals (table 3.0). Nine mutations were found in exon 8, which accounts for ~68% of the gene coding sequence, and a single mutation was identified in exon 7.

Patient number	Exon	Nucleotide change	Amino acid change	Mutation type	Protein domain	Zygosity
Patient 1 D95.7672	8	c.2411 T>C	L804P*	Missense	ERD	Compound heterozygote
	8	c.2068 T>C	C690R	Missense	ERD	
Patient 2 D96.0073	8	c.2099 G>A	W700X	Nonsense	ERD	Heterozygote
Patient 3 D01.09361	8	c.1597 C>T	P533S*	Missense	TM6	Heterozygote
Patient 4 D03.09228	8	c.1230_1233delCTCT	V412fsX440	Frameshift	TM2	Heterozygote
Patient 5 D094.2054	8	c.862-1G>A IVS7-1G>A	*	Splice	CD	Compound heterozygote
	8	c.2654 C>T	P885L	Missense	ERD	
Patient 6 D04.15546	8	c. 1400_1402delTGC	Del468L	Deletion	TM4	Homozygote
Patient 7 D03.09215	7	c.740 T>C	F247S*	Missense	CD	Heterozygote
Patient 8 (unknown)	8	c.1532 T>C	L511P*	Missense	TMEL	Homozygote

Protein domain abbreviations

ERD – Endoplasmic reticulum domain, **CD** – Cytoplasmic domain, **TM** – Transmembrane domain, **TMCL** – Transmembrane domain cytoplasmic loop, **TMEL** – Transmembrane domain Endoplasmic reticulum loop.

Table 3.0. *WFS1* mutations identified. * Denotes novel mutations

The majority of *WFS1* mutations discovered were missense: C690R, P885L, L804P*, P533S*, F247S* and L511P*, including 4 previously undescribed variants (denoted by *) (fig 3.0).

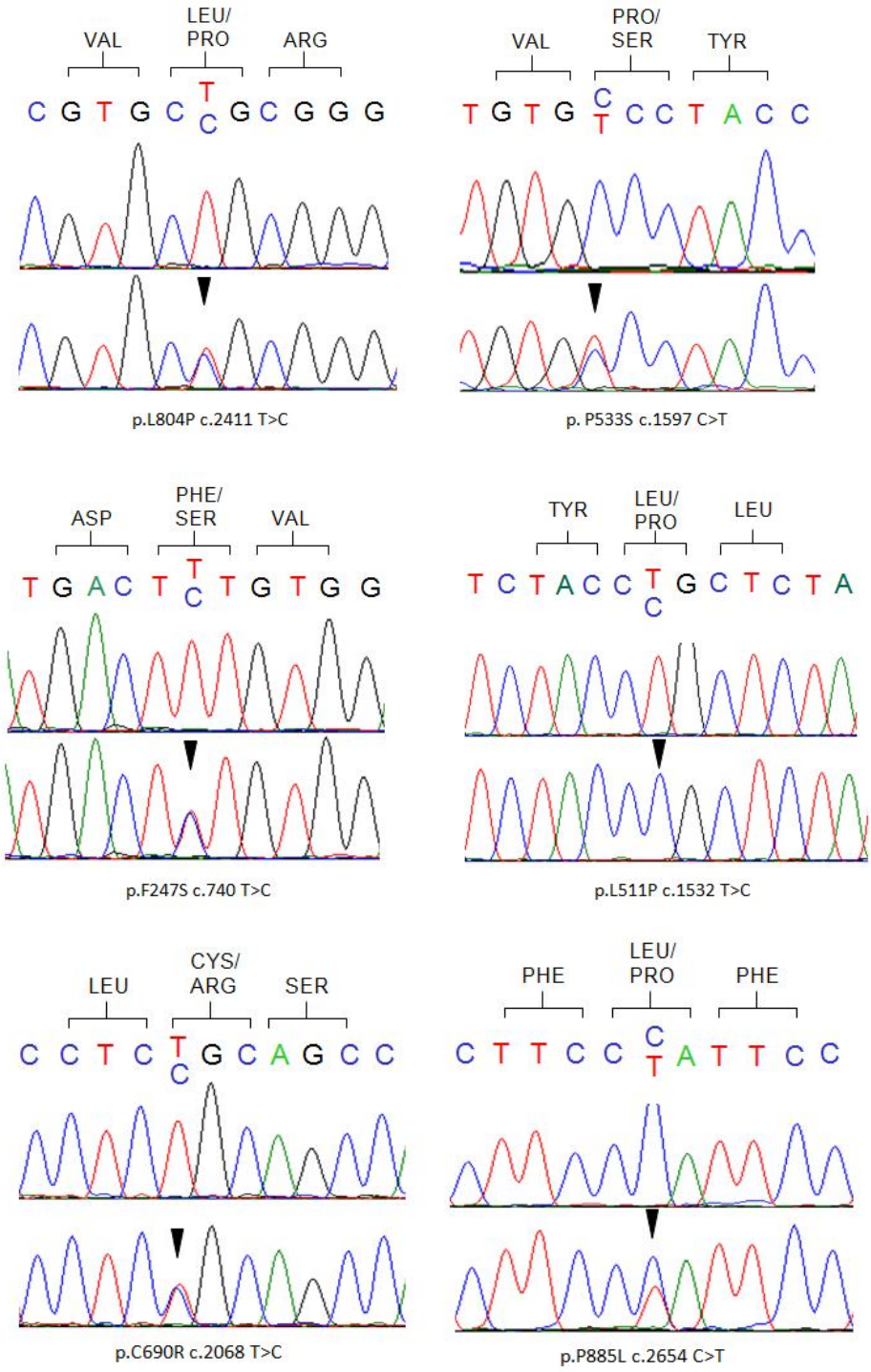


Figure 3.0. *WFS1* mutations. Electropherograms showing missense mutations (L804P, P533S, F247S, L511P, C690R and P885L) found in the *WFS1* gene matched against normal control sequences.

A further novel finding was discovered in patient 5, who was compound heterozygous for a substitution (IVS7-1G>A) at the splice acceptor site (consensus sequence AG>AA) flanking the end of intron 7 (fig 3.1). The likely outcomes of 3' splice acceptor site mutation are either exon skipping or utilisation of an available downstream legitimate or illegitimate cryptic splice site (Krawczak, 2007). In the case of our WS patient, this mutation is therefore likely to cause partial or complete absence of exon 8.

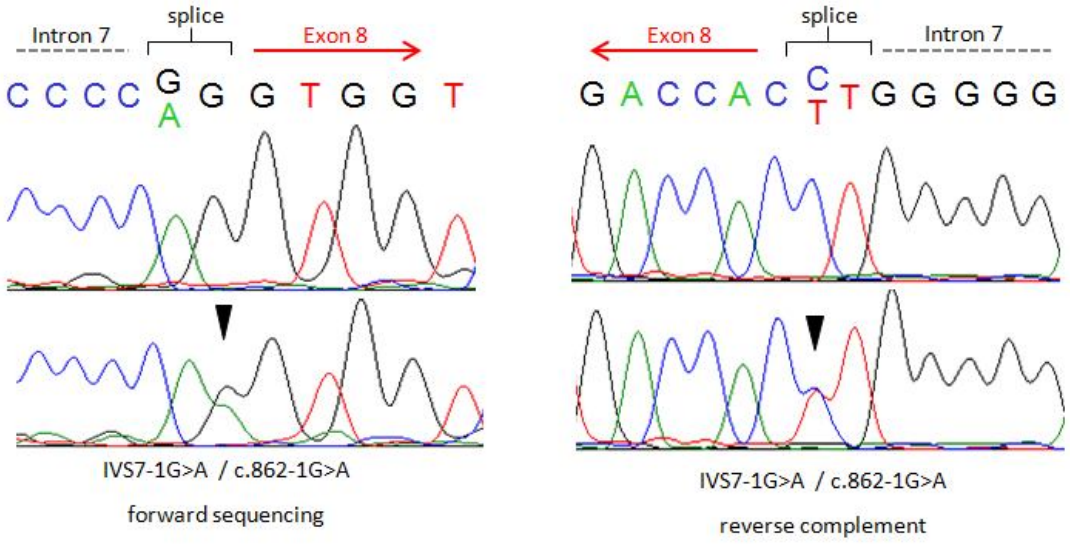


Figure 3.1. *WFS1* mutation. Electropherograms showing a novel splice site mutation (IVS7-1G>A) found in the *WFS1* gene matched against normal control sequences.

Deletions were found in patients 4 and 6. The first was a heterozygous 4bp deletion causing frameshift and formation of a premature stop codon at position 440, the second was a homozygous 3bp (out-of-frame) deletion of nucleotides 1400-1402. Finally, a single inactivating mutation (W700X) was found in patient 2 (fig 3.2).

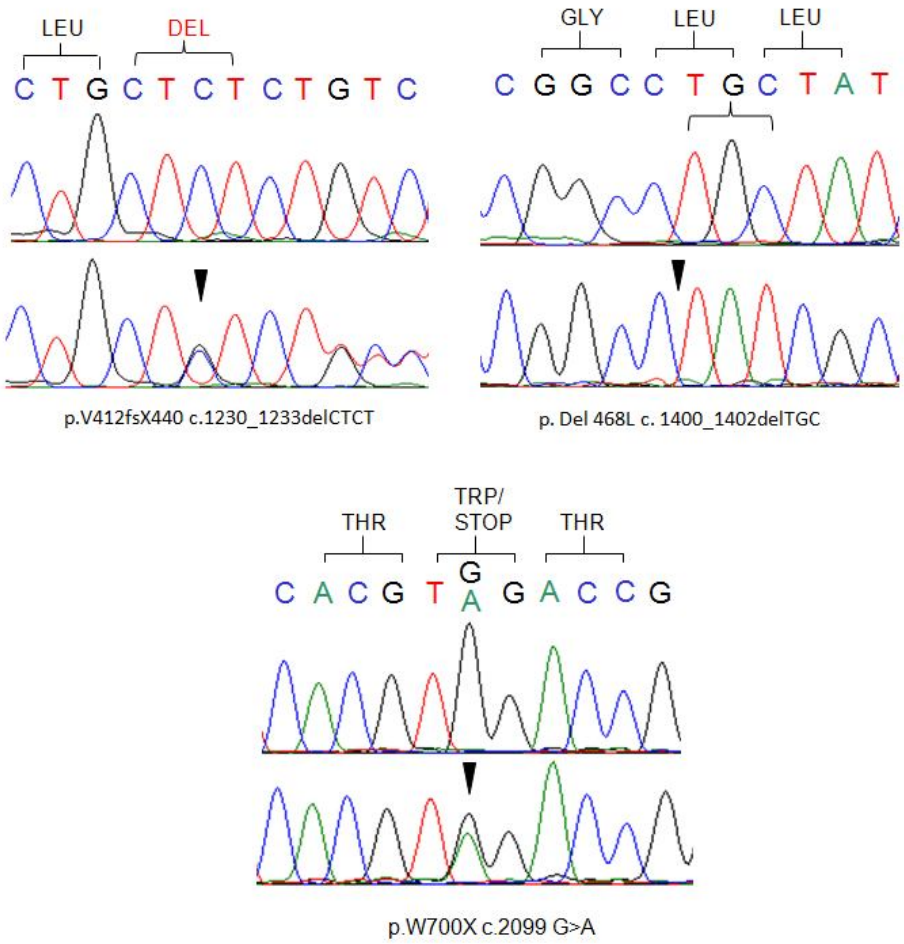


Figure 3.2. *WFS1* mutations. Electropherograms showing 2 deletion mutations and a nonsense mutation found in the *WFS1* gene matched against normal control sequences.

The position of all the discovered *WFS1* variants are illustrated in figure 3.3, a schematic showing the hypothetical amino acid configuration of the wolframin protein (adapted from Hardy, 1999). The mutations are spread throughout the length of the protein, with 1 in the cytosolic amino-terminal domain, 3 within predicted transmembrane domains, and 4 in the endoplasmic reticulum luminal domain (C-terminal). The red star denotes the start of exon 8 where the splice mutation in patient 5 was found.

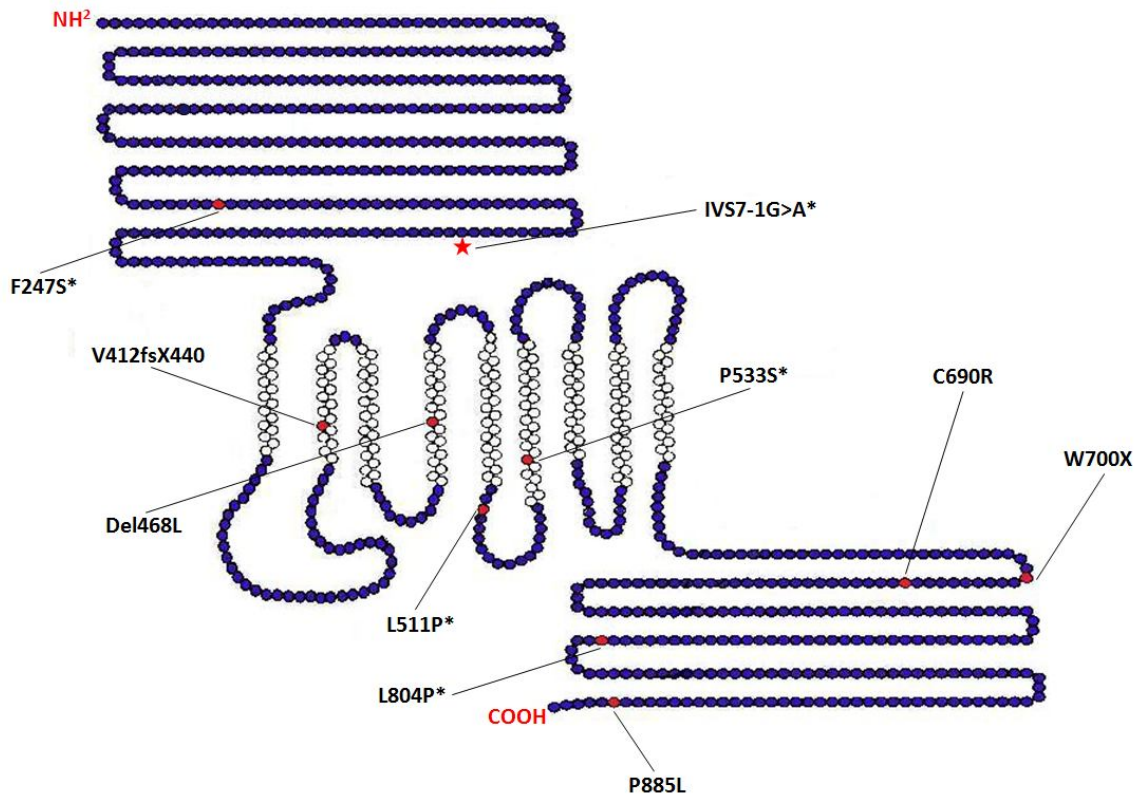


Figure 3.3. Location of mutations across the wolframin protein. Dot diagram of the predicted amino acid configuration of wolframin highlighting the location of *WFS1* mutations found in our cohort (adapted from Hardy, 1999). * Denotes novel findings.

3.2.1 Novel *WFS1* mutations

All novel *WFS1* variants were screened in a panel of ≥100 normal control chromosomes (100-200) and were found to be absent in all cases. A series of bioinformatic analyses were also applied to some of these variants looking for any physiochemical data that may or may not imply these are functionally important residues. Conservation tracks comparing homologous *WFS1* sequences in humans, other mammals, birds, zebrafish (*Danio rerio*) and the fruitfly (*Drosophila melanogaster*) show that all codon variants we found in our patients occur in highly conserved amino acids, suggesting these residues have structural or

functional importance. All codons are conserved in the mammalian and bird species, P553 and L804 are also conserved in species as phylogenetically distant as zebrafish and fruitflies (fig 3.4).

Programmes predicting the phenotypic effects of amino acid substitution, Polyphen and SIFT (Sorting Intolerant from Tolerant), were also applied to the missense mutations. Based on assemblage of functional and physiochemical annotations, both sets of results independently predicted the phenotypic consequences of the mutations as damaging (table 3.1).

WFS1 variant	SIFT score	Predicted phenotypic effect	Polyphen score	Predicted phenotypic effect
F247S	0.01	deleterious	2.603	probably damaging
P533S	0	deleterious	2.396	probably damaging
L511P	0.01	deleterious	2.243	probably damaging
L804P	0	deleterious	2.243	probably damaging

Table 3.1. Predicted phenotypic effects of *WFS1* mutations. SIFT and Polyphen predict the phenotypic effects of amino acid substitutions according to different sets of physical, chemical and functional data. A SIFT score <0.05 is predicted as deleterious. A Polyphen score of <0.5 is predicted as benign. According to these parameters, all novel *WFS1*-mutations discovered in our patients are considered pathogenic.

3.2.2 Patient Clinical Data

Patients were referred on the basis of childhood onset diabetes mellitus and optic atrophy. All available patient clinical data are tabulated in appendix 3.

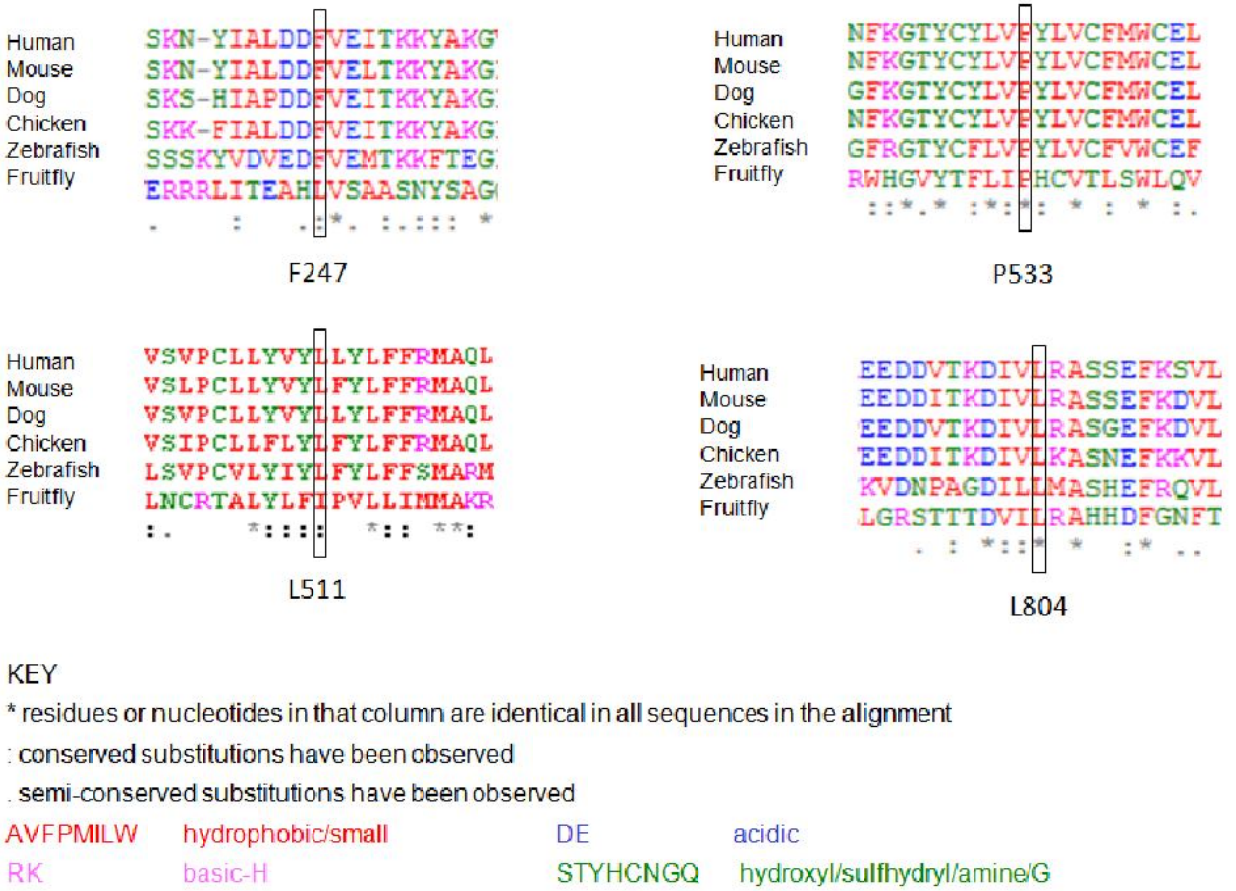


Figure 3.4. Conservation of *WFS1* residues across species (Clustal W). Each of the codons found to carry novel mutations are highly conserved residues suggestive of structural or functional significance.

3.3.0 Mutation analysis of *WFS2 (Zcd2)*

Identification of a homozygous *WFS2*-E37Q mutation

Available DNA from listed WS patients who did not screen positive for *WFS1* mutations was examined for mutations in *WFS2* (15 families). All three coding exons were sequenced in full (see appendix 2.2 for primer list) and a single individual was found to carry the homozygous missense change, E37Q, in exon 2 (fig. 3.5, table 3.2). The same mutation was previously reported by Amr *et al* (2007) in affected members of 3 consanguineous Jordanian families, and found to be absent in a large cohort of control chromosomes sourced from both matched and unmatched populations. The G>C transversion at nucleotide 109 codes for substitution of glutamic acid with glutamine, and disrupts an *Xmn1* restriction site, which is illustrated in electrophoresed digest products of our patient and normal control DNA (fig. 3.6).

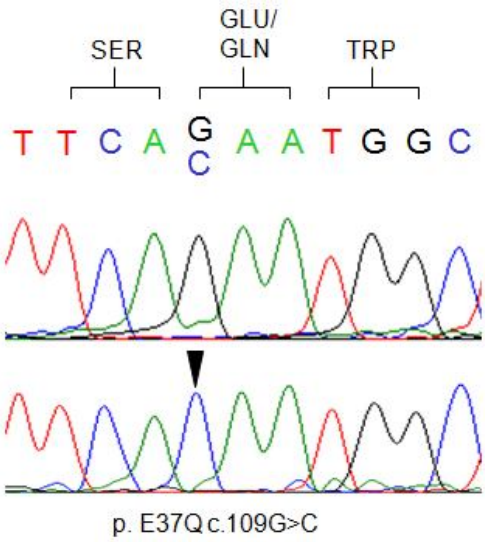


Figure 3.5. Mutation of *WFS2*. Electropherogram illustrating the *WFS2* point mutation (E37Q) found in patient D01.11844.

Gene	Exon	Nucleotide change (c.)	Amino acid change (p.)	Mutation type	Zygosity
WFS2	2	109G>C	E37Q	Missense (splice enhancer)	Homozygote

Table 3.2. Genetic data for *WFS2*-mutation. Patient D01.11844 was found to be homozygous for a *WFS2* mutation (E37Q), which encodes a splice enhancing sequence in exon 2.

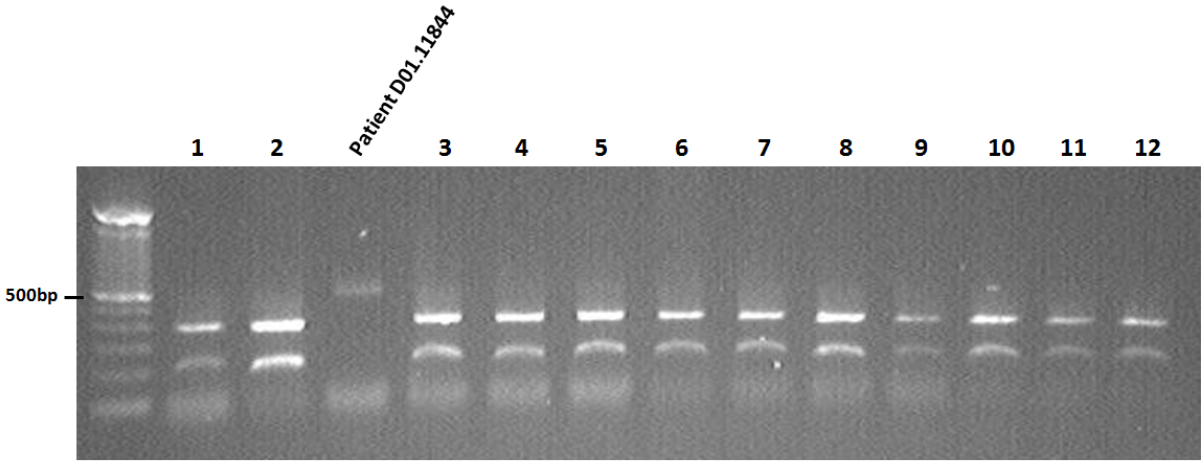


Figure 3.6. Abolition of *Xmn1* restriction site by *WFS2*-mutation. Agarose gel (1.5%) electrophoresis of *WFS2* restriction fragments digested with *Xmn1*. PCR amplified genomic DNA from patient D01.11844 and 12 control individuals demonstrates abolition of the unique *Xmn1* recognition site (GAANNNNTTC) by the E37Q mutation.

Reverse-transcription-PCR was performed on RNA from affected individuals from the Jordanian cohort by Amr and colleagues (2007) and compared to that of normal individuals. Primers encompassing full length *WFS2* generated amplicons of expected size in control individuals (551bp), whereas products generated from patient *WFS2* cDNA were only 336bp in length. Sequencing products from patient cDNA showed that exon 2 was completely missing from the final transcript. The schematic in figure 3.7 illustrates the consequences of this mutation at the transcription level. Panel A represents the normal splicing pattern of *WFS2*. Nucleotide 109 (G) sits 6bp downstream from the intron-exon boundary at the 5' end of exon 2. Following the substitution of G>C exon 2 is skipped during the splicing stage of RNA processing (panel B). The remaining exons, 1 and 3, are then joined together (panel C) causing a frameshift in exon 3 and creation of a premature termination codon after the first 8 nucleotides. The result is truncation of ~75% of the protein.

Intron-exon boundary signals and intronic *cis* and *trans*-acting elements are known to direct the spliceosome during processing of precursor RNA to mature spliced messenger RNA, containing exons only. Other genetic elements embedded within exons, known as exonic splice enhancers (ESEs) or silencers (ESS) are also partly involved in either promoting or inhibiting the splicing process. These motifs are typically 6-8bp in length (Mersch, 2008) and consist of both common and variable sequences which means they are often difficult to identify. In *WFS2*, although the region of exon 2 surrounding the 109G>C substitution does not conform to any known ESE sequences, this region has been identified as a functional ESE (Amr, 2007).

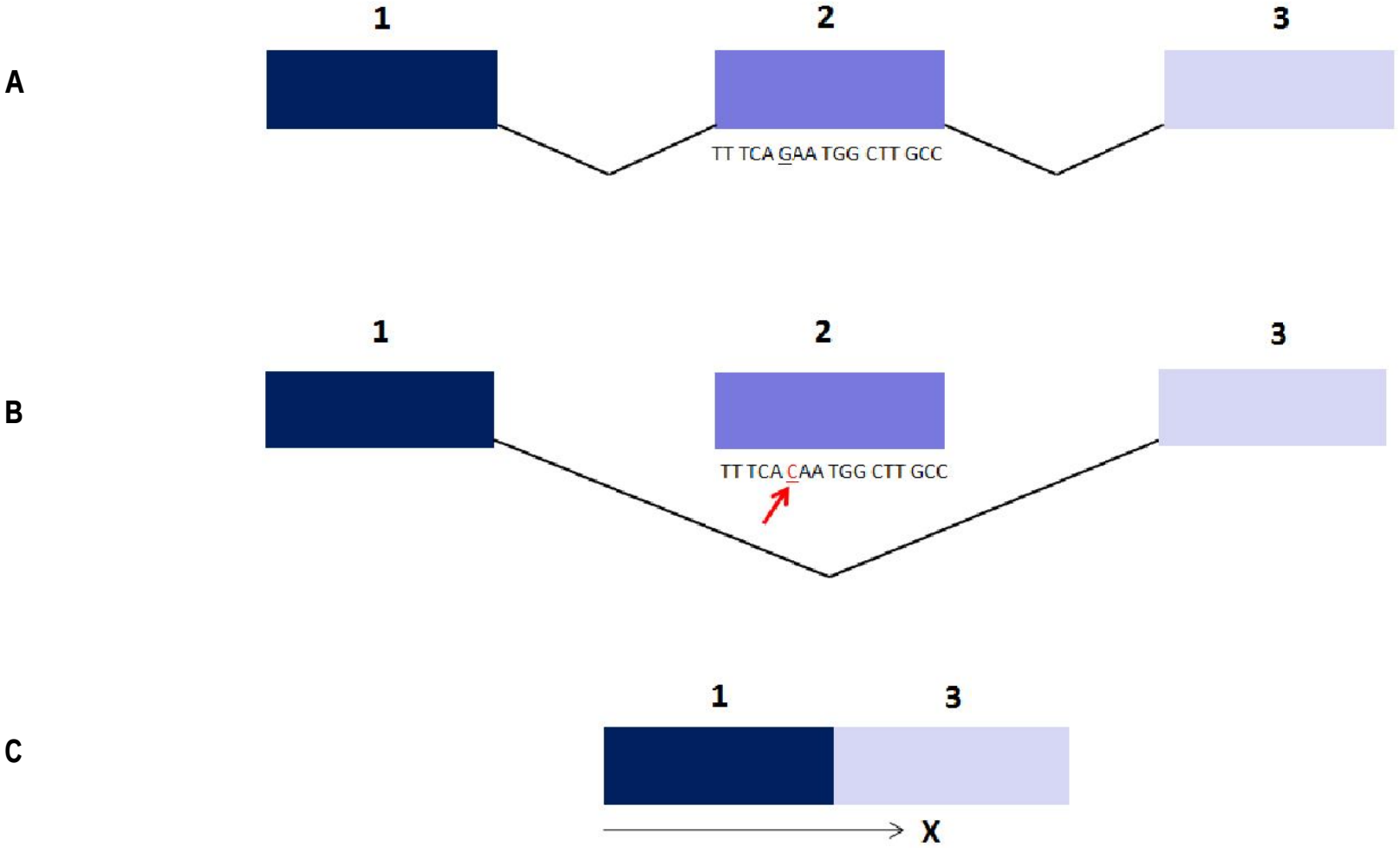


Figure 3.7. Splicing aberration caused by WFS2-mutation. Schematic diagram of splicing aberration in the *WFS2* gene introduced by the 109G>C point mutation. RT-PCR of patient lymphoblastoid RNA by Amr *et al* (2007) demonstrated that the mutation causes exon 2 to be completely skipped, suggesting nucleotide 109 is part of a functional exonic splice enhancer motif. **(A)** Nucleotide 109 sits within 6bp of exon 2 splice acceptor site (5') **(B)** Substitution of G>C at nucleotide 109 causes exon 2 to be skipped **(C)** Exons 1 and 3 are spliced together resulting in frameshift and creation of a premature stop codon, which truncates the protein product by ~75%.

3.3.1 Clinical information

The patient harbouring the *WFS2* mutation was a male, 18 years old at the time of report, who presented with a typically severe WS phenotype (table 3.3). Diabetes mellitus, optic atrophy, hearing loss, diabetes insipidus and renal complications were all diagnosed by the age of 11 years. No neurological abnormalities were reported, although these symptoms do not typically manifest until the third decade. Unlike the other affected individuals with *WFS2* mutations (El-shanti, 2000; Amr, 2007), this patient has not yet developed peptic ulcers and gastrointestinal bleeding tendency, which appears to be a unique feature of *WFS2*-associated WS.

		Age of onset (years)					
Patient number	Age (time of report)	Diabetes mellitus	Optic Atrophy	Deafness	Diabetes insipidus	Renal/Urinary	Other
D01.11844	18	11	4	4	11	11 (kidney transplant)	End-stage renal disease, psychiatric abnormalities

Table 3.3. Age of onset (years) of clinical features in a patient with WFS2-associated WS. Patient D01.11844, homozygous for *WFS2*-E37Q presented with a typical spectrum of WS clinical features including the first reported case of diabetes insipidus in *WFS2*-associated WS. Peptic ulcer disease, a characteristic only seen in WS patients with *WFS2*-mutations, was not diagnosed at the time of this report although the patient was only 11 years of age.

Combining clinical information from our patient and the four original affected families carrying *WFS2* mutations (El-shanti, 2000), we noted the age of diabetes mellitus onset ranged from 5-15 years, and optic atrophy from 4-22 years (table 3.4). All affected individuals, with respect to age, developed the full spectrum of WS clinical characteristics

plus gastrointestinal disease. While suspected in some, none of the original cohort was diagnosed with diabetes insipidus, although this feature was present in our patient.

Age of onset (years)						
Diabetes mellitus	Optic atrophy	Deafness	Diabetes insipidus	Renal/urinary	Peptic ulcer/bleeding tendency	Neurological/other
10 (5-15)	13 (4-22)	10.5 (4-25)	?	23 (11-35)	20.5 (15-26)	Depression (suicide)

Table 3.4. Median age (and range) of onset in years of clinical features in *WFS2*-associated WS families.

Little neurological data are available from the original families although central respiratory failure was stated as the cause of death in one individual suggesting that neurological manifestations are a part of *WFS2*-associated WS as well. Further support for this comes from an independent study by a group at the Saudi Centre for Organ Transplantation (Hasan, 2000) who reported an isolated case of WS in a 15 year old boy born to consanguineous Jordanian parents. Although no genetic information was available at the time, the ethnic origins of this individual combined with the presentation of upper gastrointestinal bleeding tendency and duodenal ulcers in addition to the other WS clinical features, suggests he may in fact also carry a *WFS2* mutation. An ataxic gait pointed toward neurological manifestations and following cranial magnetic resonance imaging (MRI) generalised brain atrophy and a small sized pituitary gland relative to the patients' age were apparent.

3.3.2 Haplotype analysis

All of the *WFS2*-associated cases of WS have been reported in families of Jordanian descent. Previous microsatellite analysis of *WFS2* families revealed a common haplotype for markers D4S3256, D4S1531, and D4S1564 in all affected individuals across a critical region around the *WFS2* locus (El-Shanti, 2000), suggesting they may have shared ancestry.

The *WFS2* patient in our Birmingham cohort is also of Jordanian background, so we tested whether this individual shared the defined haplotype with the other families. We were kindly sent some DNA from a member of the original cohort courtesy of Rita Shiang, Virginia Commonwealth University, USA, and genotyped our patient along with this individual and an unaffected normal control (section 2.5.0). The results did in fact confirm that our patient carries the same haplotype as the other affected families, pointing towards a founder effect (see appendix 5 for raw data).

3.4.0 Mutation analysis of *ATP1B1*

Since we have ruled out a number of WS patients as carriers of *WFS1* or *WFS2* coding sequence mutations, we investigated the possibility of mutation in an additional candidate gene *ATP1B1*. This 7 exon gene codes for the glycoprotein subunit (β1) of sodium potassium ATPase a ubiquitously expressed integral membrane protein, which is responsible for maintaining electrochemical gradients of Na^+ and K^+ across the plasma membrane.

ATPase is a high energy consumer, closely linked to energy utilisation and glucose uptake and inhibition of this enzyme has been linked with apoptosis and neuronal cell death (Lees,

2001). The Na^+/K^+ ATPase β1 subunit has also been shown to directly interact with the C-terminus of wolframin, and reduced expression of Na^+/K^+ ATPase β1 subunit seems to be linked to mutation or absence of wolframin (Zatyka, 2008a).

Due to the relatively small size of this gene (2608 bps), the connections it has with *WFS1*, and the functional relevance this protein has to many of the primary systems affected by WS, we proposed this gene as a reasonable candidate for mutation analysis. Following direct sequencing of *ATP1B1* coding exons (see appendix 2.3 for primers list) however, we ruled out mutations in this gene as the causative factor for WS in our patients, as no mutations in any of the screened individuals (10 patients) were identified.

3.5.0 Mutation analysis of *SLC19A2*

Thiamine (vitamin B1) is an essential nutrient for normal cell function, growth and development, and an important coenzyme in intermediary metabolism. Low intracellular thiamine levels impair energy production pathways and have been linked to ER and oxidative stress, and apoptosis (Wang, 2007). Systemic deficiency is often attributed to inadequate dietary intake in developing countries, chronic alcoholism or diabetes mellitus and can lead to cardiovascular and neurological disorders. Localised tissue-specific thiamine deficiency is associated with the pathology of thiamine responsive megaloblastic anaemia (TRMA), a rare, recessively inherited, monogenic disorder.

TRMA is primarily characterised by juvenile-onset megaloblastic anaemia, non insulin-dependent diabetes mellitus, and sensorineural deafness (Rogers, 1969), although other clinical findings such as heart, neurological, and retinal defects have also been described in a number of cases (Neuwirth, 2009). TRMA patient cells are highly sensitive to thiamine depletion, although treatment with pharmacologic dosages of vitamin B1 can ameliorate anaemia, arrest hearing loss and delay the onset of diabetes (Mathews, 2009).

The gene associated with TRMA, designated *SLC19A2*, was mapped to chromosome 1q 23.3 (Neufeld, 1997) and codes for a 497 amino acid transporter protein THTR1, which is essential for the cellular uptake of thiamine. Loss of THTR1 has been shown to result in morphological and oxidative damage, and all known cases of TRMA to date (25 families) have been associated with mutational dysfunction of this protein (Neuwirth, 2009).

During the course of this study, one listed patient on our WS database with no known mutations (patient D05.25758), presented with sideroblastic anaemia. This individual was 9 years of age, and had other clinical features typical of WS including juvenile onset diabetes mellitus, deafness and eye disease, although the latter was found to be cone rod dystrophy unlike the characteristic optic atrophy seen in WS patients (table 3.5). In light of this new development, the clinical diagnosis shifted from what initially appeared to be WS to something conforming more closely to TRMA. Our aim was to screen this individual for mutations in *SLC19A2* to confirm TRMA with a genetic diagnosis. All 6 coding exons of *SLC19A2* were sequenced in our patient, an affected sibling, and control DNA (see appendix 2.4 for list of primers).

Patient number	Age (time of report)	Sex	Diabetes mellitus	Optic atrophy	Deafness	Other
D05.25758	11	M	1.5	No	3 months	Eye disease: Cone rod dystrophy (2), Sideroblastic anaemia (9)

Table 3.5. Age of onset (years) of clinical features in a patient with *SLC19A2*-associated TRMA. Patient D05.25758, diagnosed with thiamine responsive megaloblastic anaemia (TRMA) presented with sideroblastic anaemia at age 9 following initial diagnosis of WS.

3.5.1 Confirmation of inactivating *SLC19A2* mutation in patient with TRMA phenotype

A c.484C>T point mutation was discovered in both patient D05.25758 from our cohort and his affected brother. This mutation substitutes an arginine residue with a termination codon at amino acid 162 (R162X), truncating the protein by approximately two-thirds (fig 3.8; table 3.6). This mutation has been described before in 3 other families of Turkish (Raz, 2000), Pakistani, and Japanese origins (Labay, 1999), and possibly reflects a mutational hotspot since none of these families appear to share a common ancestor (Raz, 2000).

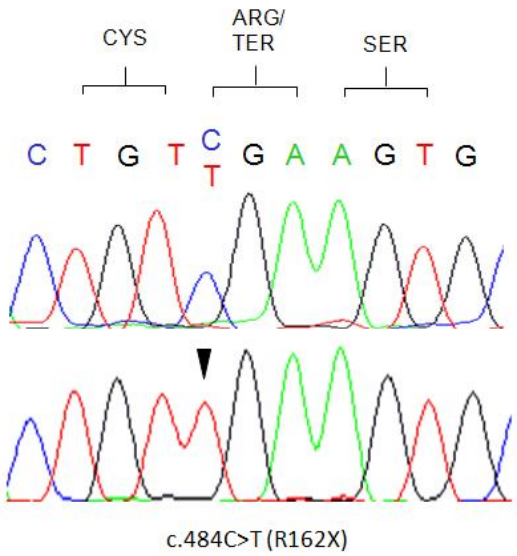


Figure 3.8. Mutation of *SLC19A2*. Electropherogram illustrating the *SLC19A2* homozygous point mutation (R162X) found in patient D05.25758, diagnosed with thiamine responsive megaloblastic anaemia.

Gene	Exon	Nucleotide change (c.)	Amino acid change (p.)	Mutation type	Zygosity
SLC19A2	2	c.484C>T	R162X	Nonsense	Homozygote

Table 3.6. Genetic data for *SLC19A2* mutation. Patient D05.25758, was found to carry a homozygous nonsense mutation in *SLC19A2* (R162X), confirming a genetic diagnosis of thiamine responsive megaloblastic anaemia (TRMA).

As already documented, not all TRMA patients present with the same severity of phenotype or the typical triad of associated features (Neuwirth, 2009), even when comparing those individuals carrying identical *SLC19A2* mutations (table 3.7) making diagnosis and detection tricky. Depending on the order and presentation of characteristics the spectrum of TRMA clinical features can be similar to those in WS. Our patient also had visual abnormalities, a feature not always present in TRMA syndrome, further complicating initial diagnosis.

Reported	Anaemia	Diabetes mellitus	Deafness
Raz, (2000)	3	3	3
Labay, (2009)	2	2	birth
Labay, (2009)	1 month	6.5	6

Table 3.7. Age of onset (years) of clinical features associated with *SLC19A2*-R162X. Data from three previously reported TRMA families harbouring the same *SLC19A2* nonsense mutation found in our affected individual (patient D05.25758).

3.6.0 Sequencing analysis summary

From a total of 26 probands we have successfully identified 10 mutation types (16 chromosomes): 8 in the *WFS1* gene, 1 in *WFS2* and 1 *SLC19A2* mutation. Only a single heterozygous *WFS1* coding sequence mutation was found in several of the listed patients and 16 individuals screened negative for any coding sequence mutations of *WFS1*, *WFS2,* or *ATP1B1*. However, large deletions and intronic/promoter sequence variants have not been ruled out.

Perhaps revisiting clinical records and observations might suggest an alternate diagnosis, since WS does share features common to several other diseases such as maternally inherited diabetes and deafness (King, 1992), TRMA and Alstrom syndrome (Alstrom, 1959). It is also possible that other, as yet unidentified genes are responsible for WS in these patients, as locus heterogeneity is already a known facet of this disorder.

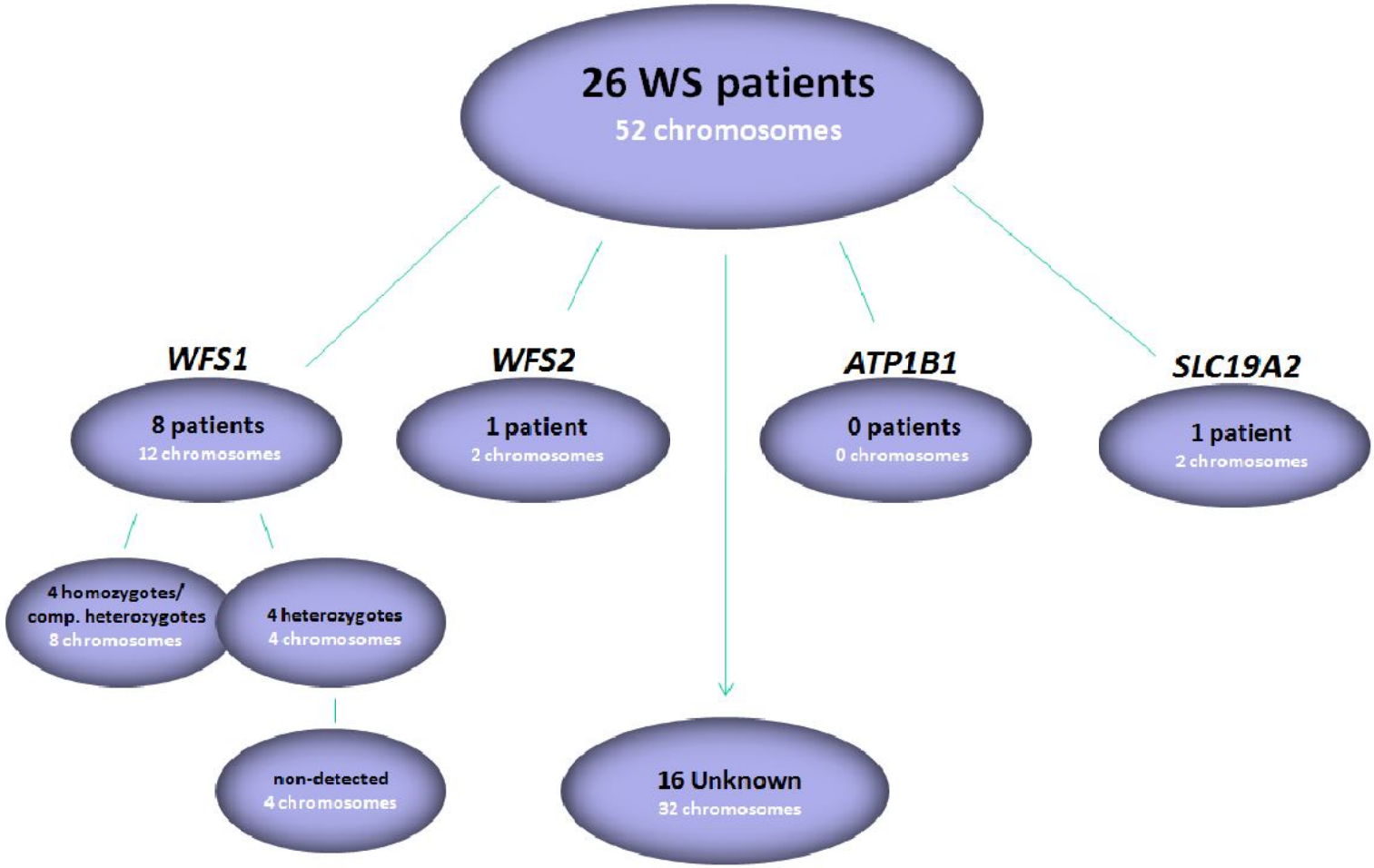


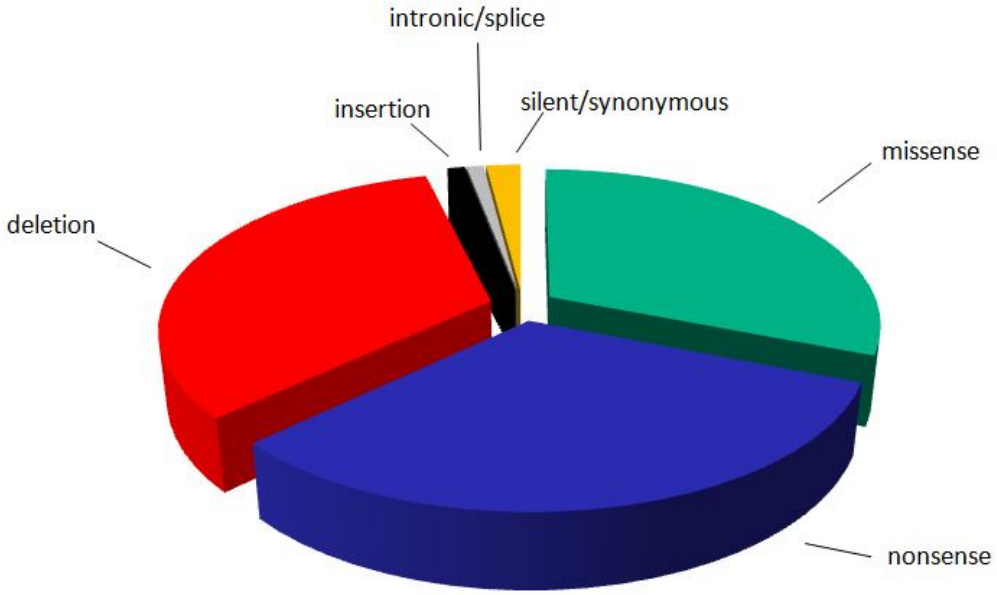
Figure 3.9. Summary of WS mutation analysis. From a total of 26 probands (52 chromosomes): mutations were found in 10 individuals, accounting for 16 chromosomes. *WFS1* mutations were found in 8 patients: 2 homozygotes, 2 compound heterozygotes and 4 patients with only one mutation detected. One patient was found to be homozygous for a *WFS2* mutation and a further patient was homozygous for *SLC19A2* mutations. No mutations were detected in 16 individuals, who were also screened for possible mutation of *ATP1B*1.

3.7.0 Genotype-phenotype correlation

The West Midlands Wolfram syndrome patient registry to date, lists 93 individuals from 83 families (appendix 3.0-3.1). Excluding those individuals we found to harbour *WFS2* and *SLC19A2* mutations, we carried out some genotype-phenotype analysis and correlation on available data from probands only.

From a total of 162 chromosomes, 121 mutations have so far been identified in the *WFS1* gene. The most commonly found are missense, nonsense and deletion mutations which occur in almost equal frequencies, 31%, 32% and 33% respectively. Insertions, and splice mutations account for only 2% of the total and other unique synonymous changes, perhaps pathogenic, were discovered in 2% of cases (fig 3.10). One or more *WFS1* mutations have been found in 67 individuals, where almost half (45%) are compound heterozygotes, 30% are homozygous and the remainder have only one mutation detected.

Where patient clinical data were available the median age of onset was calculated for distinct clinical features (table 3.9). From a total of 42 probands, diabetes mellitus presented at a median age of 6 years (range 1.5-30) and optic atrophy at 10 years (range 1-18). Of those individuals who presented with additional features, the mean ages of onset were: deafness at 9.5 years (range 0-39), diabetes insipidus at 13 years (range 2.5-40), renal and urinary tract complications at 22 years (range 10-46) and neurological abnormalities at 30.5 years (range 9-44). Deafness was reported in 61% of cases over 30 years of age, and diabetes insipidus and renal/urinary disturbances were found independently in 78% of these individuals.



Total number of mutations	121	% of total
missense	37	31
nonsense	39	32
deletion	40	33
insertion	1	1
intronic/splice	1	1
silent	3	2

Figure 3.10/Table 3.8. Proportions of mutation-types found in the West Midlands cohort (n=93). Nonsense, missense and deletion mutations occur in almost equal frequencies.

In order to relate genotype to phenotype, patients were categorised into three main groups according to the mutation-types they carried. The first group were those found to be either homozygous or compound heterozygous for inactivating mutations, the second homozygous or compound heterozygous for missense changes and the third were compound heterozygotes for a missense and an inactivating mutation. Both nonsense and deletion-frameshift changes were considered 'inactivating'.

Onset of optic atrophy was calculated at a median age of 11, 10 and 8.5 years for each group respectively, all of which correlate closely with the overall median age of presentation for the whole cohort. The median age of diabetes mellitus onset was 5 years both in group 1 (n=13) and 3 (n=7), where at least one inactivating mutation was present. Group 2 (n=4), carrying missense mutations only, presented with diabetes mellitus at a median age of 10.5. When combining groups 2 and 3 (group 4; n=11) to represent all patients carrying at least one missense *WFS1* mutation, the median age of optic atrophy and diabetes mellitus onset was 10 years (table 3.9).

Group (n=number of patients)			Diabetes mellitus	Optic atrophy
		Entire cohort	6	10
1	n=13	inactivating-inactivating	5	11
2	n=4	missense-missense	10.5	10
3	n=7	inactivating-missense	5	8.5
4	n=11	missense-inactivating/missense	10	10

Table 3.9. Genotype-phenotype correlation in WS. Median age of onset (years) for primary WS clinical features between individuals harbouring selected combinations of inactivating and missense *WFS1* mutations. Patients carrying at least one missense mutation present with diabetes mellitus at a later median age than those with inactivating mutations only.

Our observations follow a similar trend to the findings of Cano *et al* (2007) who performed a meta analysis of combined clinical and genetic data from six independent studies of WS patients (n=96). They too found that patients with one or more *WFS1* missense mutations

seemed to have a later age of diabetes mellitus onset. Those carrying two missense mutations had a median age of onset of 12 years, compared with one missense mutation (median age of onset: 9 years) and no missense mutations (median age of onset: 5 years). Unlike our study they also noticed a slightly earlier age of optic atrophy presentation in patients with one or more inactivating mutations (median: 10 and 11 years respectively) compared to those with no inactivating mutations (median: 14 years).

Despite the relatively small numbers of suitable patients, we compared those individuals over 30 years with one or more missense mutations (n=6) to those with inactivating mutations only (n=7) to see whether any differences in overall clinical expression of WS were apparent (fig 3.11). Interestingly a greater proportion of individuals carrying inactivating mutations appeared to have a more complete spectrum of complications with 71% of patients presenting with deafness and 100% of patients with diabetes insipidus compared to those individuals with one or more missense mutations who presented with deafness over the age of 30 in only 66% of cases and diabetes insipidus in 44%. Renal and urinary tract complications were reported in roughly equal proportions in both groups (83% and 86% respectively). Neurological findings were also detailed in all instances except one where the patient was only in fact 30 years of age.

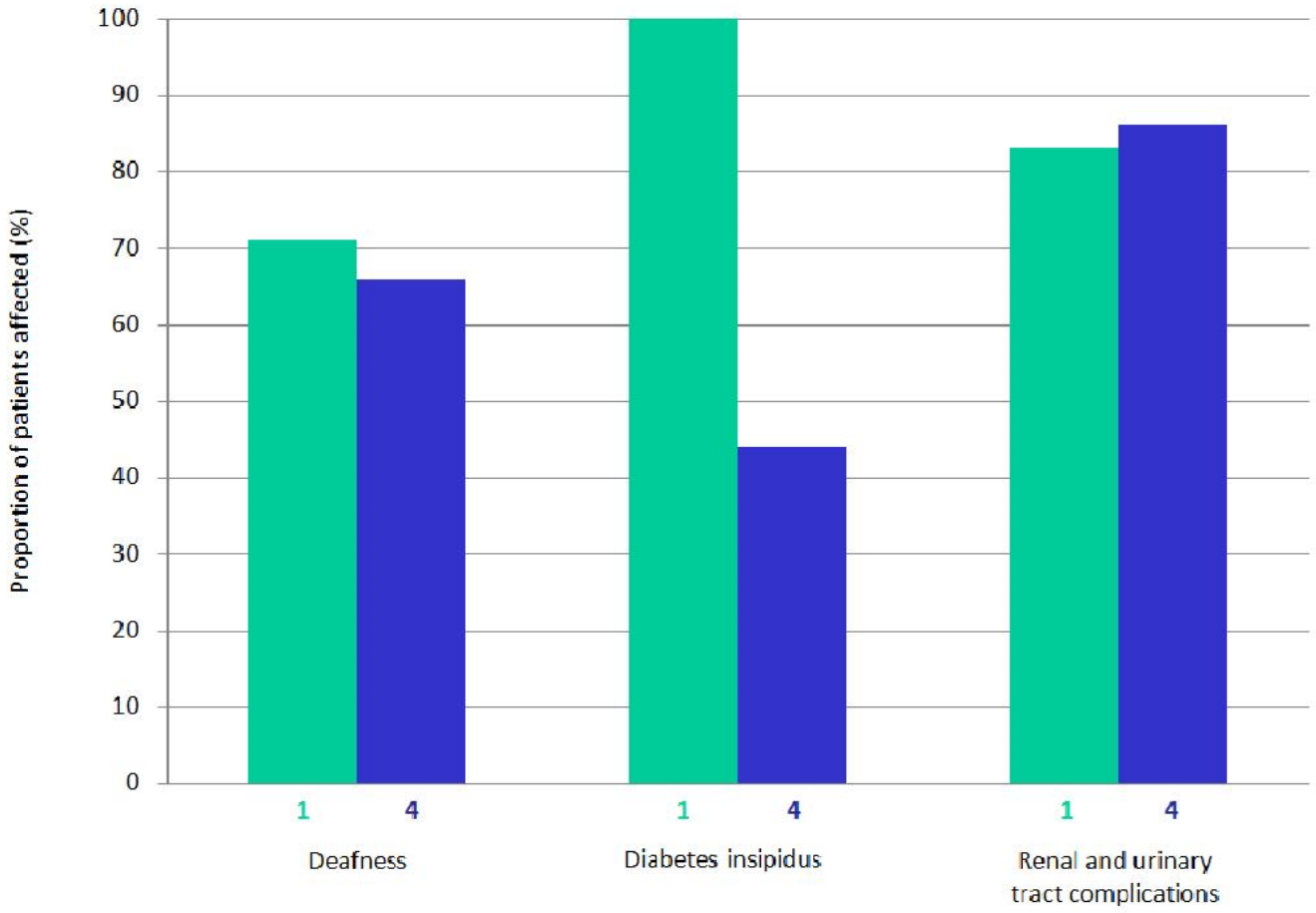


Figure 3.11. Genotype-phenotype correlation in WS. Comparative presentation of additional WS clinical features between individuals carrying inactivating mutations only (group 1, n=13) and those with at least one missense mutation (group 4, n=11). Proportion of affected individuals over 30 years of age.

3.8.0 Summary and discussion

Our findings have further completed existing genetic data on the West Midlands Wolfram syndrome patient database by uncovering disease-causing mutations in 10 individuals. A total of 10 distinct *WFS1* variants were found in 8 patients, 5 of these mutations have previously been described, and 5 are novel findings. The suspected pathogenicity of these novel changes are supported by predictive physio-chemical data, conservation status of associated amino acids, and the absence of these mutations in a control panel of normal chromosomes.

We also report the first case of *WFS2*-associated WS in the UK population. Conforming with previously reported families carrying this mutation this individual has a Middle-Eastern ethnic background and shares a defined disease-haplotype with other affected individuals, suggestive of common ancestry. As the only reported *WFS2* mutation to date, this WS-associated allele perhaps only poses a disease risk to culturally and geographically isolated populations of Jordanian-Bedouin background. Contrary to earlier reports, where diabetes insipidus was absent in affected individuals (El-Shanti, 2000; Amr, 2007) we also document the presence of diabetes insipidus as a feature of *WFS2*-associated WS.

One listed individual previously thought to have WS presented with anaemia during the course of this study. The clinical diagnosis was re-evaluated and designated as suspected Thiamine Responsive Megaloblastic Anaemia (Rogers syndrome). We were able to confirm a genetic diagnosis by screening the thiamine transporter gene, *SLC19A2,* which carries mutations in every other reported case of this syndrome. Our patient was homozygous for a nonsense change, previously described in several other affected families.

Coding exons of the *ATP1B1* gene were also sequenced for potential mutation in all individuals that screened negative for known WS-gene mutations. This relatively small gene was considered a reasonable candidate as a number of interesting links between this gene and WS pathology can be made in addition to the direct interaction between the protein products of this gene and wolframin. We did not however discover any significant sequence changes in *ATP1B1*, therefore confirming that coding variants in this gene are not responsible for WS in our cohort.

From a total of 93 patients listed on the regional WS database, genetic mutations have yet to be identified in 16 patients (4 of which have no more available DNA). We have ruled out coding mutations in *WFS1*, *WFS2* and *ATP1B1* as causative factors for WS in these cases although the presence of intronic or promoter sequence variants or large deletions in these genes are a possibility.

The mutational spectrum of *WFS1* is expanded by our findings, bringing the total number of unique *WFS1* mutations reported to date to over 530. Together with the clinical data presented, we further highlight not only the genetic heterogeneity of WS but the phenotypic variability of this disease and likeness to other genetically distinct, but phenotypically similar disorders such as TRMA. Continued molecular analysis allows us to refine the diagnostic criteria for WS and overlapping syndromes sharing similar phenotypes.

Our results support the notion that correlation between mutation type and disease severity may exist. In agreement with earlier comparative analyses we found in our cohort that people harbouring inactivating *WFS1* mutations present with diabetes mellitus at an earlier age (median: 5 years) than those with at least one missense mutation (median: 10 years). We did not observe any notable differences in age of optic atrophy onset although patients

with two inactivating mutations seem to present with a wider spectrum of additional clinical features than those with at least one missense mutation. Taken together with previous evidence these trends suggest that some missense mutations may retain partial function.

3.8.1 Further work

Making extensive genotype-phenotype correlations in rare disorders is limited primarily by sample numbers, but in the case of complex diseases such as WS analyses are further complicated by genetic heterogeneity and variability of clinical manifestations. It is important to take into consideration patient ages when making correlations between groups where possible within the constraints of limited sample sizes. Building on current data archives will gradually improve our understanding of WS natural history and continual detailed documentation of patient information in the clinic throughout disease progression to adulthood will also help us to perform more informative and detailed genotype-phenotype correlation on larger age-matched groups.

Those individuals with WS phenotypes but no identified mutations might provide an interesting panel for discovering potentially novel WS-causative genes if they exist, although DNA from sufficiently large numbers of family members would be necessary for gene mapping studies. Due to the functional relatedness of mitoNEET (*Cisd1*) to *WFS2* (*Cisd2*) we propose this gene as an interesting candidate for association with WS. Perhaps *ATF6α,* described in chapter 1 (section 1.5.1) could also be considered due to the regulatory relationship wolframin has been shown to have on this important ER-stress sensing protein. The application of next generation sequencing technology, which is becoming increasingly

more accessible and cost-effective would also allow whole-exome or even whole genome capture, which may lead to the discovery of new disease genes.

Our data suggest that some wolframin variants with *WFS1* missense mutations might retain some capacity to function, which may be the basis for delayed onset and presentation of fewer clinical features in some individuals carrying these types of mutation. From a sequencing analysis perspective we may be able to glean some further clues about the phenotypic consequences of individual *WFS1* missense mutations by gathering and comparing the clinical data from patients homozygous for missense mutations with those from patients who are compound heterozygous for missense mutations.

Mechanisms behind phenotypic variability can only be explained by functional analyses of the gene product, wolframin. The ultimate correlation would be to determine precise functional roles of specific regions of the wolframin protein so that targeted therapies might be developed. A selection of wolframin 'missense' variants will be characterised in chapter 4 forming the principle stage of our functional investigation.

CHAPTER 4

CHARACTERISATION OF SELECTED *WFS1*-VARIANTS

4.0 Introduction and aims

It has been postulated that mutations in the *WFS1* gene result in loss of wolframin function. Evidence suggests that mutations may cause reduced steady-state protein levels and/or a range of alterations to the characteristics and behaviour of wolframin. However, a comprehensive understanding of the relationship between *WFS1* genotype and resulting protein functionality has yet to be reached. Previous *in vitro* studies have demonstrated the potential for missense mutations to result in wolframin aggregation, instability and mis-localisation (Hofmann, 2003 and 2006; Yamaguchi, 2004; Fonseca, 2005). Analysis of fibroblasts from patients carrying various mutations has also shown in some cases significantly lower expression levels of wolframin, or indeed an apparently complete absence of wolframin (Hofmann, 2003 and 2006; Philbrook, 2005; Fonseca, 2005), suggesting mutations may induce production of unstable, readily degraded proteins.

If the fate of mutated wolframin may vary depending on the nature of the mutation it is also possible that impairment of function may be variable too. Even if all Wolfram syndrome-associated mutations culminate in similar end-points, ultimately cell degeneration caused by uncontrolled ER-stress, the pathogenic mechanisms leading up to the system break-down may be different. The select few mutations that have been characterised and analysed for function certainly do not reflect the full spectrum of mutations reported in Wolfram

patients. Each study also has its own technical and analytical limitations such as comprehensive clinical data, detection methods, and adequate controls.

Our genotype-phenotype data (chapter 3), in concordance with findings from other groups (Smith, 2004; Cano, 2007; D'Annunzio, 2008) suggest that WS patients with at least one *WFS1* missense mutation have a relatively later median age of onset than those patients harbouring inactivating mutations and are less likely to present with the full spectrum of WS-associated clinical features. These observations, together with reports of relatively 'mild' WS phenotypes in some patients carrying *WFS1* missense mutations (Gomez-Zaera, 2001; Van Den Ouweland, 2003; Gasparin, 2009; Zymslowska, 2011) form the basis of our hypothesis that:

Some wolframin variants may retain partial function.

From a therapeutic perspective, identifying even limited functional wolframin in some patients may be important for future developments in therapeutic intervention, such as introduction of targeted chemical chaperones.

The aims of this chapter were to explore the possibility that selected *WFS1* missense mutations may still allow some residual wolframin protein function. The foundation of this work was to select a representative range of *WFS1* mutations for study and to characterise these in terms of expression levels, localisation and stability, before looking to investigate further any suitable candidates for specific changes in function.

An *in vitro* 'over-expression' model was the approach used in this study for a number of reasons. Firstly, this allowed the freedom to select specific *WFS1* mutations to investigate,

which were not available from any cell or tissue banks at this time. These mutations were selected to represent a wider spectrum of *WFS1* mutations than previously investigated, where clinical data were available from homozygous patients. Some previous studies have been limited by the use of heterozygous models, which are difficult to interpret when considering relationship to phenotype. The selection could also be made to reflect a spread of mutations found across the *WFS1* gene or protein as desired. Using well established cell lines and over-expressing our proteins of interest had the advantages of lower technical complexity and ease of replication in the laboratory, enabling the generation of more efficient and cost-effective preliminary data. Finally, once *WFS1* mutations in mammalian expression vectors are constructed and characterised, they could be used in a range of subsequent functional assays, for which we have already validated cell lines within the Birmingham laboratory and well established protocols suitable for our purposes.

How well a model system such as this one reflects conditions in Wolfram patients is the obvious limiting factor when considering a practical experimental approach. Translating any information generated in isolated populations of immortalised cells to conditions *in vivo* is tentative. For instance, the demands placed on cells by vastly over-expressing proteins could have countless and variable effects on their overall behaviour. We can account in part for these effects by making comparisons where appropriate with endogenously expressed wolframin, and can also (at least) state with reasonable credence that protein variants over-expressed in the same conditions as wild-type have the 'potential' to behave similarly or differently. Although evidence indicates that some wolframin protein variants may be undetectable and others expressed at reduced levels in affected individuals, an assumption

is made that all of the selected wolframin variants in our panel are expressed to some degree in humans, and without tissue from patients this conjecture cannot be confirmed.

The initial characterisation of our chosen wolframin variants was therefore aimed at identifying the 'potential' for differences to be observed *in vitro*, and thus laying a foundation on which to build further detail for confident translation to the human condition.

4.1.0 Selection of *WFS1* gene mutations

The selection of *WFS1* mutations for this study was based on reported phenotype. Assuming some wolframin protein variants have residual function we planned to compare mutations seen in relatively mild cases of Wolfram syndrome to those reported in severely affected patients. If mutations seen in patients with a milder phenotype retain some partial function, we might expect to observe fewer abnormalities in their characteristics and function compared to mutations found in patients with a severe clinical phenotype. Classification of patient phenotype as mild or severe was defined with reference to the median age of onset of the primary WS clinical features; juvenile-onset diabetes mellitus (6 years) and optic atrophy (10 years) calculated from both our cohort (chapter 3) and from the work of Domenech (2004) who reported average ages of diabetes mellitus onset at 5.8 years and optic atrophy at 11 years from the combined data of 5 studies (Hardy, 1999; Gomez-Zaera, 2001; Colosimo, 2003; van den Oweland, 2003; Domenech, 2004).

We have chosen to concentrate our attention on missense mutations since truncated transcripts and proteins resulting from other types of nonsense and frameshift mutations are vastly different in character than wild-type and generally found to be highly unstable.

Attributing a given phenotype to specific mutations is less complex when the individual is homozygous for these changes, so the variants selected for investigation here have all been reported in homozygous individuals with a clear record of clinical data. Consideration was also given where appropriate, to the position of the substitution along the length of the protein, in an attempt to include variation found in a range of protein domains.

4.1.1 Clinical information

Table 4.0 summarises the selection of mutations used in this study and the clinical features reported. The chosen amino acid substitutions relating to typical severe WS phenotypes are leucine to proline at codon 511 (L511P), and glycine to alanine at codon 736 (G736A). In chapter 3, the novel homozygous mutation L511P was discovered in a consanguineous Turkish family with 7 affected members under the age of 17, all presenting with a severe phenotype. Diabetes mellitus presented at a mean age of 4 years and optic atrophy at 8.5 years. Other clinical features including sensorineural hearing loss, renal and urinary tract abnormalities, and diabetes insipidus were also noted to different degrees in the older individuals. The proband listed in table 4.0 was a 10 year old female who presented with diabetes at age 3, optic atrophy at 9 years, and had also developed neurogenic bladder, chronic renal insufficiency and diabetes insipidus at the time of the report. The second ‘severe’ homozygous mutation (G736A) was reported in an 8 year old Spanish girl from a

consanguineous background (Domenech, 2004). She had developed diabetes mellitus and optic atrophy before age 6 and was also diagnosed with possible onset of diabetes insipidus.

A relatively mild case of WS was reported in a 39 year old Brazilian patient (Gasparin, 2009) homozygous for a glutamic acid to lysine substitution at codon 158 (E158K). The onset of diabetes and optic atrophy in this individual was during adulthood, 27 years and 19 years respectively, and the additional reported clinical features including deafness, urinary tract abnormalities, and erectile dysfunction also presented relatively late. Secondly, an arginine to histidine change at codon 558 (R558H) was selected for our study since a 38 year old French individual homozygous for this mutation also appeared to have milder WS (Cano, 2007). He developed diabetes mellitus at 29 years and optic atrophy at 35 years of age. Although he also presented with neurogenic bladder at 33 years and suffered from depression, no other clinical features were identified. A more recently reported case of an individual homozygous for *WFS1*-R558H (Chaussenot, 2011) was another French patient, 40 years of age. Again, this individual presented with late-onset diabetes mellitus and optic atrophy (32 years), but also presented with deafness and neurological complications (cerebellar ataxia, nystagmus, dementia and moderate cerebellar atrophy) at 40 years.

Interestingly R558H has been reported previously in compound heterozygous individuals also carrying inactivating mutations. An affected individual from an Italian family was compound heterozygous for R558H and a 6bp deletion (p.R178_Q179del) and had a typically 'severe' spectrum of WS early-onset features (Colosimo, 2003). Another two affected sibs (17 and 18 years of age) from a Canadian family compound heterozygous for R558H/E864X, had presentation of diabetes mellitus and optic atrophy at 6-10 years and 13-14 years respectively, but had no other reported complications (Smith, 2004).

Amino acid variant	Diabetes mellitus	Optic atrophy	Other complications	Relative severity	Reported
E158K	27	19	deafness (39) urinary (33) erectile dysfunction	mild	Gasparin, 2008
L511P	3	9	urinary (10) diabetes insipidus (10)	severe	Prince, 2011
R558H	29	35	urinary (33) depression	mild	Cano, 2007
R611H	Type 2 diabetes risk allele				Minton, 2002; Sandu, 2007; Florez, 2008; Franks, 2008
G736A	5.5	5	diabetes insipidus (8)	severe	Domenech, 2004
E737K/R818C	20	33	urinary cerebrall atrophy tremor	mild	Gomez-Zaera, 2001

Table 4.0 Selected *WFS1* missense variants and associated clinical features.

The variants E737K and R818C have been included in this investigation, although to what extent each of these contributes to WS pathogenesis is unknown. The patient listed in table 4.0 is a 38 year old female of Spanish origin (Gomez-Zaera, 2001). Her parents are first-cousins and she is homozygous for both of these variants. She had relatively late onset primary diagnostic features, but also developed urinary tract and neurological symptoms. A sibling with the same genotype also presented with almost identical features and died following food aspiration at the age of 42. The authors of this study suggest the causative mutation is more likely R818C since this substitution results in more significant chemical alteration than E737K, although neither of these variants were seen in control chromosomes. An earlier Japanese study (Ohtsuki, 2000) looking at potential links between *WFS1* variant carrier status and psychiatric illness did discover E737K in 10% of the patients, although also noted the presence of this variant in 7% of control individuals. R818C has also been documented in other studies (Smith, 2004) but not yet excluded as a benign polymorphism.

Perhaps the most interesting *WFS1* variant is codon 611. Although accepted as a common single-nucleotide polymorphism, H611 (c.1832G>A) has generated extensive interest as a potential susceptibility allele for type 2 diabetes (Minton, 2002; Sandu, 2007; Florez, 2008; Franks, 2008). The R611 allele remarkably, has also been highlighted as a possible marker for mood disorders (Zalsman, 2009), medication overuse headache (Lorenzo, 2007), and type 1 diabetes (Awata, 2000). To date, no functional analysis of this variant has been carried out. We therefore aimed to understand more about how this particular *WFS1* allele could contribute to complex disease, by primarily generating biochemical data about the characteristics of this variant, then seeking to identify any obvious limitations in function.

4.1.2 Physio-chemical properties of wolframin 'missense' variants

The hypothetical structure of the wolframin protein and location of *WFS1* substitutions of our interest are illustrated in figure 4.0 (adapted from Hardy, 1999). The tabulated data (table 4.1) include the physio-chemical properties of the wild-type and substituted amino acids. The N-terminal end of the wolframin protein sits on the cytosolic side of the ER and residue 158 is situated in this domain. The substitution of glutamate (E) with lysine (K) replaces an amino acid that is negatively charged at a physiological pH with one that is positive. The same substitution E>K at residue 737 has been selected, although this amino acid is located in the carboxy-terminal portion of the protein, within the ER lumen.

Leucine (L) is hydrophobic in nature and therefore generally seen in the interior of proteins. Amino acid L511 is likely to reside in, or near to the hydrophobic cross-membrane section of wolframin. Substitution with proline (P) at position 511 could have a significant impact on the overall protein structure, since this is the most rigid of the naturally occurring amino acids. This inflexibility is due to linkage between the side-chain and nitrogen atom, and thus thought to play an important role in determining protein conformation.

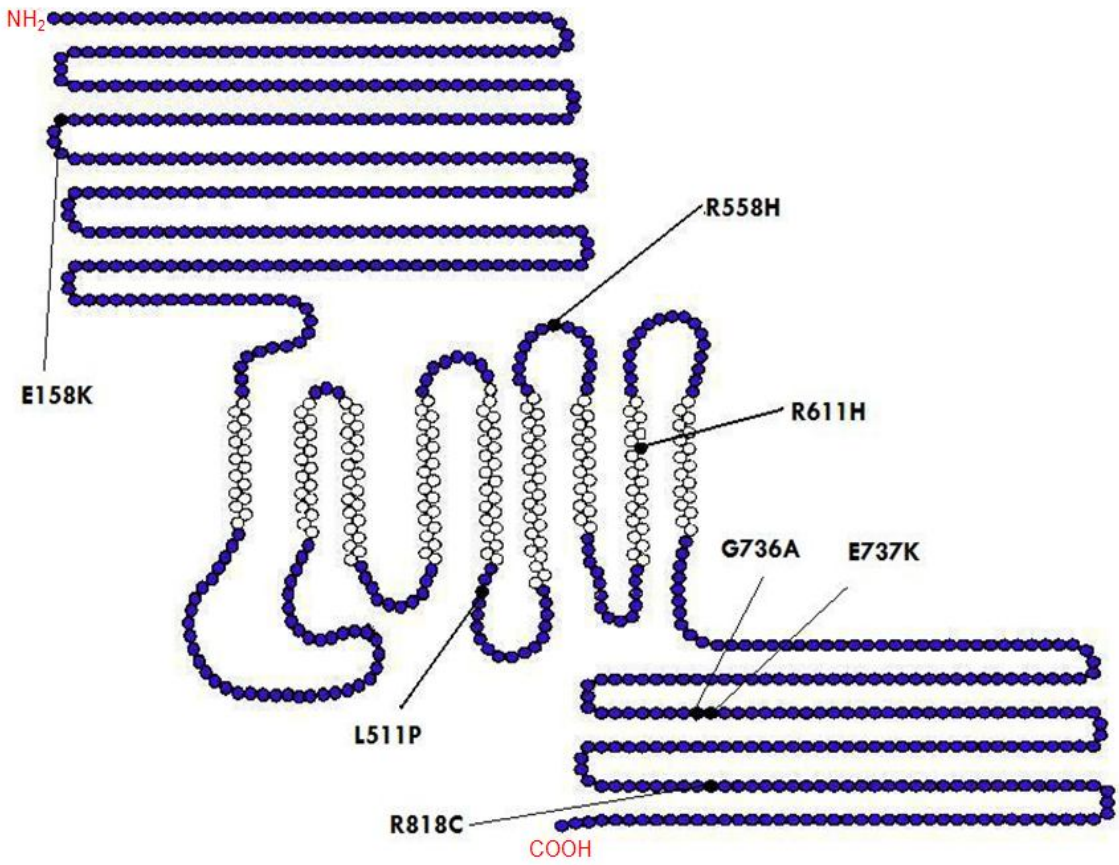


Figure 4.0 Dot diagram showing the predicted amino acid configuration of the wolframin protein and location of selected sequence variants. Adapted from Hardy (1999).

WFS1 variant (p.)	*WFS1* variant (c.)	Exon	Position	Wt variant	Substitute variant
E158K	472G>A	5	CD	Polar Hydrophilic charged (-)	Polar hydrophilic charged (+)
L511P	1532T>C	8	TMCL	Aliphatic hydrophobic neutral	Hydrophobic neutral
R558H	1673G>A	8	TMLL	Polar hydrophilic charged (+)	Aromatic polar hydrophilic charged (+)
R611H	1832G>A	8	TM	Polar hydrophilic charged (+)	Aromatic polar hydrophilic charged (+)
G736A	2206G>C	8	LD	Aliphatic neutral	Aliphatic hydrophobic neutral
E737K	2209G>A	8	LD	Polar Hydrophilic charged (-)	Polar hydrophilic charged (+)
R818C	2452 C>T	8	LD	Polar hydrophilic charged (+)	Polar hydrophobic neutral

Table 4.1 *WFS1* missense variants – Genetic and physiochemical data

KEY: CD = cytoplasmic domain, TM = transmembrane domain, CL = cytoplasmic loop, LL = luminal loop, LD = luminal domain.

The other amino acid thought to be conformationally important is glycine (G). In contrast to proline, this small residue lacks a side chain which confers a great deal of flexibility, and is often located in turn regions of proteins allowing polypeptide chains to form a tightly packed configuration. The residue at position 736 is present in the ER luminal domain and has been substituted with alanine, a hydrophobic molecule very different in nature. It is interesting to note here that the two amino acids implicated in overall protein conformation (leucine and glycine) are those thought to be involved in the pathogenesis of WS in both our severely affected patients.

The arginine (R) to cysteine (C) change found at codon 818 replaces a hydrophilic, charged amino acid with a neutral residue. The sulphur groups of cysteines allow the formation of disulphide bonds and dimerisation, associated with stabilising protein structure. Again, this points toward a particularly significant disruption to the structure and function of wolframin, and as mentioned earlier may be more damaging than E737K.

The final changes, arginine (R) to histidine (H) at 558 and 611 are both thought to be positioned in, or close to the transmembrane domains. Being hydrophilic, perhaps residue 611 is more likely to reside in the cytosolic or luminal loops as opposed to the proposed position on the diagram. The substitution R>H does not alter the physio-chemical characteristics of the residue drastically, since one basic cation is replaced with another. Charged residues can take part in solvent binding and electrostatic interactions with complementary ions. It may be that some of these interactions could be hampered but not abolished following the above substitutions, which would support our hypothesis that some *WFS1* mutations may encode partially functioning proteins. This seems a likely situation

when considering the H611 allele, attributed to a diabetic predisposition and not WS, but a possibility also for the R558H mutation, which has been found in a 'mild' case of WS.

Results

4.1.3 Bioinformatic analysis of wolframin 'missense' variants

Bioinformatic analyses of sequence variants can provide clues and supporting evidence to aid our understanding of protein function. Comparative alignments of nucleic acid and protein sequences across a range of species can give us an indication of how functionally important a particular sequence is. Those occurring farthest back in the phylogenetic tree and conserved despite evolutionary speciation are thought to represent particularly important structural and functional elements.

Alignments of wolframin protein sequences were made using ClustalW (ebi.ac.uk) (fig 4.1). Sequences from humans and other mammalian species were compared to zebrafish (*Danio rerio*) and fruitflies (*Drosophila*). Codons 158, 558, 736, and 737 are identical through all species. Codons 511 and 611 differ only in the fruitfly sequence, being substituted conservatively at 511. Codon 818 is the least conserved, aligning identically in the mammalian species only.

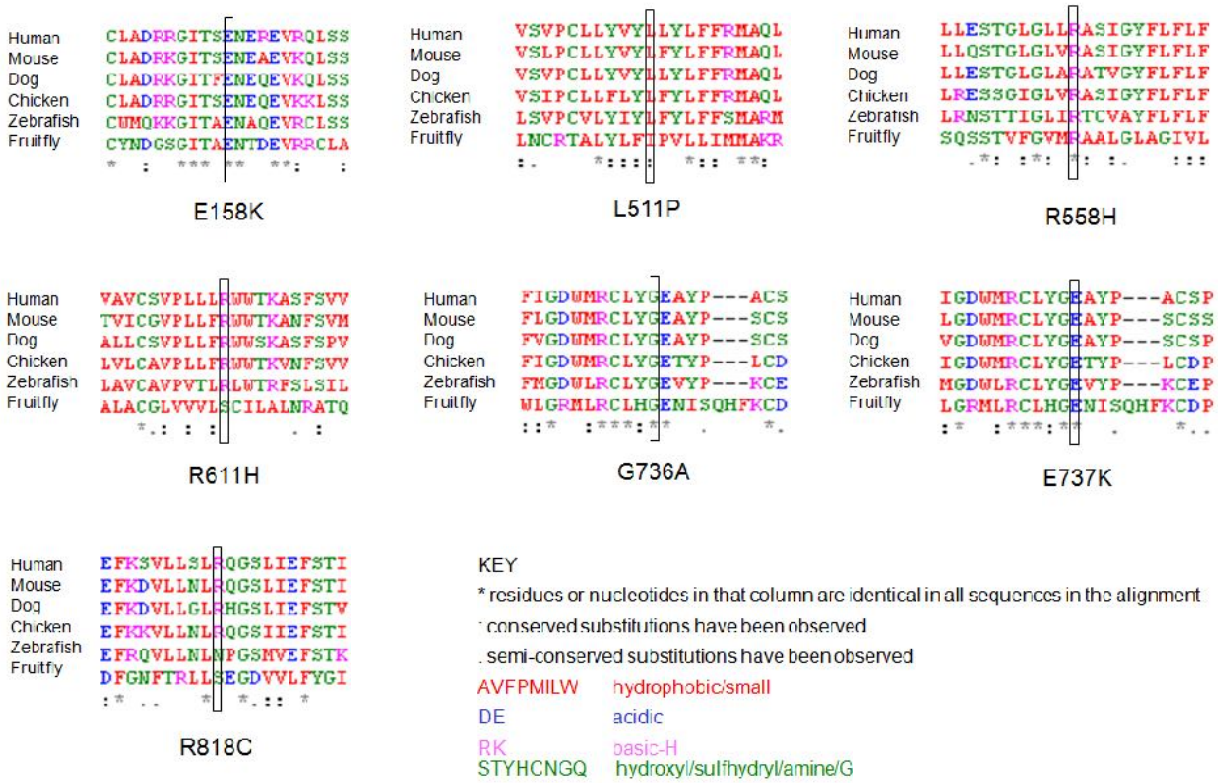


Figure 4.1 Conservation of wolframin across species (ClustalW). Amino acid residues selected for study are all conserved across mammalian species, some as phylogenetically distant as the fruitfly signifying functional importance.

SIFT (sorting intolerant from tolerant) and Polyphen (polymorphism phenotyping) are programmes designed to predict the probable phenotypic effects caused by gene sequence changes. Multiple sequence alignment, protein function annotation, and information about the physical properties of amino acids are combined to analyse a query sequence.

The resulting output from SIFT is a comprehensive table of possible substitutions at each amino acid residue and a normalised probability score from 0-1 representing the predicted phenotypic effects of that change. Scores less than 0.05 are more likely to be deleterious, and those greater than or equal to 0.05 are predicted to be tolerated. The scores generated from our *WFS1* substitutions are listed in table 4.2. All of these changes, with the exception of E158K have a score indicative of deleterious phenotypic consequences. E158K, found in a mildly affected WS patient, resulted in a SIFT score of 0.08, and is classified as tolerated.

Polyphen interprets quantitative probability data to produce a qualitative prediction outcome of sequence variants on phenotype: benign, possibly damaging, or probably damaging. The results from our panel of *WFS1* mutations suggest that all 8 variants are damaging. L511P has the highest score (2.243) and is predicted with the greatest degree of confidence to result in a deleterious outcome of 'probably' damaging. The remaining variants all produce a score indicative of 'possibly' damaging.

WFS1 variant	SIFT Score	Predicted phenotypic effect	Polyphen Score	Predicted phenotypic effect
E158K	0.08	tolerated	1.609	possibly damaging
L511P	0.01	deleterious	2.243	probably damaging
R558H	0	deleterious	1.979	possibly damaging
R611H	0.04	deleterious	1.852	possibly damaging
G736A	0	deleterious	1.824	possibly damaging
E737K	0	deleterious	1.609	possibly damaging
R818C	0.02	deleterious	1.965	possibly damaging

Table 4.2 Bioinformatic analysis of selected wolframin variants using SIFT and Polyphen. Phenotypic effects of amino acid substitutions are predicted according to different sets of physical, chemical and functional data. A SIFT score <0.05 is predicted as deleterious. A Polyphen score of <0.5 is predicted as benign.

The output from these programmes cannot be directly compared since different parameters and data sets are used to calculate the predictions. Both sets of results do however agree that all selected *WFS1* mutations are likely to have a deleterious effect on phenotype, other than SIFT analysis of E158K, which was considered to be tolerated. It is reassuring to note that E158K has a relatively lower Polyphen score compared to the other variants, as does E737K, supporting our clinical information about the 'milder' WS phenotype relating to E158K and the previously suggested benign effects of E737K.

All of the *WFS1* mutations in our study may induce adverse phenotypic effects, although to variable degrees of severity. It is therefore anticipated that each of the codons in question has a crucial role to play in normal structure and/or functioning of wolframin. As expected, these amino acids are highly conserved throughout a range of mammalian species, some even in more distantly related animals. Additionally, the statistical analyses lend support to associations made between genotype and phenotype, predicting (generally) the outcome of these substitutions to be deleterious.

4.2.0 Cloning and site-directed mutagenesis of *WFS1*

In order to characterise the selected *WFS1* mutations *in vitro*, we prepared some mammalian expression vectors containing wild-type *WFS1* and the chosen range of variants. The vector of choice was pcDNA3.1, which promotes over-expression of untagged proteins.

Human Wt-*WFS1* sequences in pCMV-myc expression vectors were available for our use in the Birmingham laboratory (courtesy of Dr Malgosia Zatyka). From these plasmid preparations, human *WFS1* cDNA was sub-cloned (section 2.19) into pcDNA3.1(+) using

EcoR1 and *Not1* restriction sites (fig 4.2). Following sequence verification (section 2.4.0), *WFS1*-pcDNA maxipreps (section 2.7.5) were used as templates for site-directed mutagenesis (SDM) PCR (section 2.6.0) to generate the sequence variants. Introduced mutations were also confirmed by sequencing. Unwanted changes may be introduced to the vector 'backbone' sequence during the mutagenesis reaction, potentially interfering with promoter activity or antibiotic resistance genes. To avoid the arduous task of full sequencing, each of the mutated *WFS1* sequences were subsequently inserted to fresh pcDNA3.1 vectors, which had not been subject to a mutagenesis PCR reaction. For a more detailed description of the protocol please refer to section 2.6.0 (primers list in appendix 2). Electropherograms depicting wild-type and mutated *WFS1* sequences are shown in figure 4.3 and a comprehensive description of nucleotide and codon substitutions are described in table 4.3.

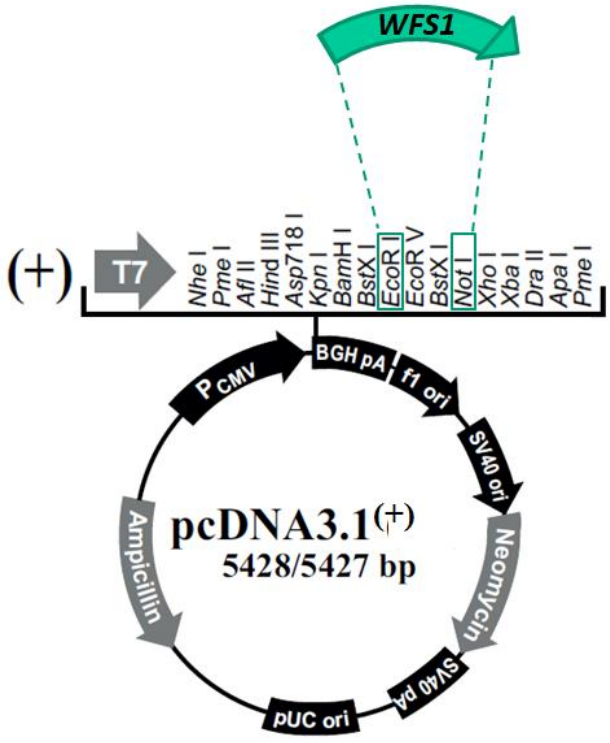


Figure 4.2 Cloning of human *WFS1* into mammalian expression vector pcDNA3.1(+).

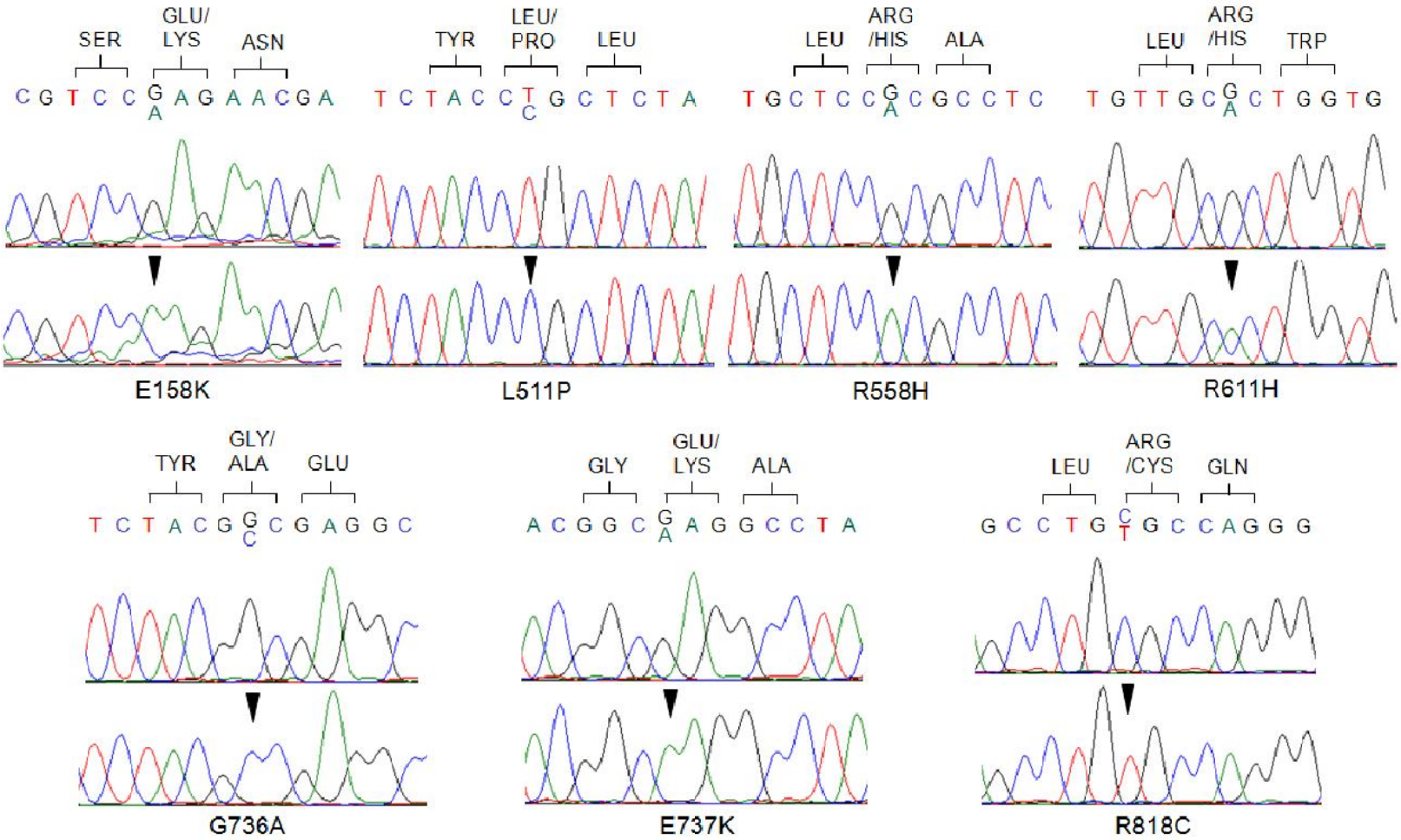


Figure 4.3 Site-directed mutagenesis of human *WFS1*. Electropherograms illustrating introduced point mutations.

Variant	Amino acid	Codon
E158K	glutamic acid>lysine	GAG>AAG
L511P	leucine>proline	CTG>CCG
R558H	arginine>histidine	CGC>CAC
R611H	arginine>histidine	CGC>CAC
G736A	glycine>alanine	GGC>GCC
E737K	glutamic acid>lysine	GAG>AAG
R818C	arginine>cysteine	CGC>TGC

Table 4.3. Site-directed mutagenesis of human *WFS1.* Amino acid and codon substitutions

4.3.0 Steady-state expression of wolframin variants

4.3.1 Specific aims and methods

The first step in characterising our *WFS1* constructs was to examine their relative steady-state expression. We aimed firstly to verify over-expression of these proteins in our model system validating their use in further transfection experiments. Some evidence suggests that wolframin expression in WS patients is markedly reduced or entirely absent (Hofmann, 2003 and 2006; Philbrook, 2005; Fonseca, 2005), therefore our second objective was to determine whether our missense mutations result in reduced steady-state expression of wolframin.

HEK293 cells were used in these experiments as they provide a quick-growing, easily manipulated, and readily transfected human cell line. Purified *WFS1*-pcDNA plasmids (1μg) were transfected with FugeneHD reagent (Roche) into HEK293 cells grown to 80-90%

confluency (section 2.9.5). Following a 24 hour incubation period, the cells were harvested in lysis buffer (2.10.1), sonicated on ice (2 x 10 seconds), and centrifuged for 30 minutes (4°C) at 14,000rpm. The supernatant was harvested and measured for protein concentration on a BioRad 750nm Microplate Absorbance Reader (section 2.11.0). Samples (10µg/lane) were electrophoresed on SDS-PAGE gels and western blotted (section 2.13.0). Semi-quantitative absorbance measurements of the resulting blots were then calculated using GeneSnap/GeneTools software (Syngene; section 2.13.5).

The influences of technical variability, particularly transfection efficiency, were minimised by repeating the experiment a number of times.

4.3.2 Steady-state expression of wolframin variants harvested in RIPA buffer

The first set of transfections (n=4) were harvested in RIPA buffer for western blotting. Wolframin protein was detected with polyclonal α-*WFS1* antibody (Proteintech) and variance in gel protein loading was corrected for with measurements of β-actin, detected with α-β-actin antibody (Sigma). Figure 4.4A is a representative blot illustrating steady-state expression of wild-type wolframin and selected variants following transfection into HEK293 cells. Variation between the wolframin mutants is apparent, and variability between experiments was also observed to different degrees. Figure 4.5 shows the mean expression values corrected for β-actin from all of the experimental repeats and standard error. The corresponding quantitative measurements are detailed in table 4.4 (see appendices 7 and 8 for raw data). The figures show that aside from *WFS1*-L511P, the steady-state expression of all the wolframin variants fall into a range between 73.4% (± 24.6%) and 118.9% (± 26.7%) relative to wild-type wolframin. The mean expression level of *WFS1*-L511P however, is only 46.2% (± 19.0%) relative to wild-type. Despite considerable deviation from the mean value,

consistently lower levels of protein were detected from *WFS1*-L511P throughout all conditions, indicating that this variant may give rise to depleted steady-state expression of wolframin.

Two-tailed t-tests comparing wild-type-*WFS1* with each of the wolframin variants lend support to our initial observations. Statistically significant differences (p.=0.047) could be found between the steady-state expression levels of wild-type-*WFS1* and *WFS1*-L511P.

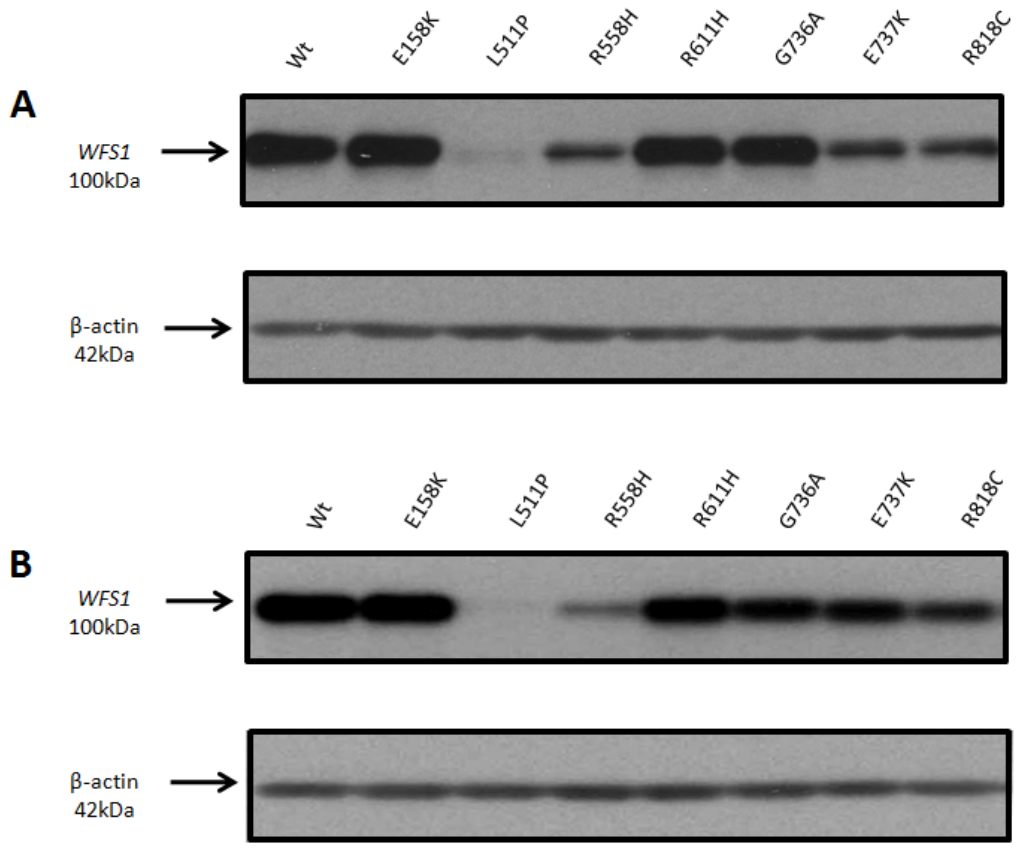


Figure 4.4 Steady-state expression of wolframin variants harvested in RIPA buffer. Clone set 1. **(A)** Typical SDS-PAGE immunoblot analysis of samples probed with anti-*WFS1* Ig (Proteintech) and anti β-actin (Sigma). **(B)** Selected samples re-run and probed with an alternative anti-*WFS1* Ig (courtesy, Dr S. Hofmann) and β-actin (Sigma). Comparative signal detection suggests that mutations do not influence antibody binding efficiency.

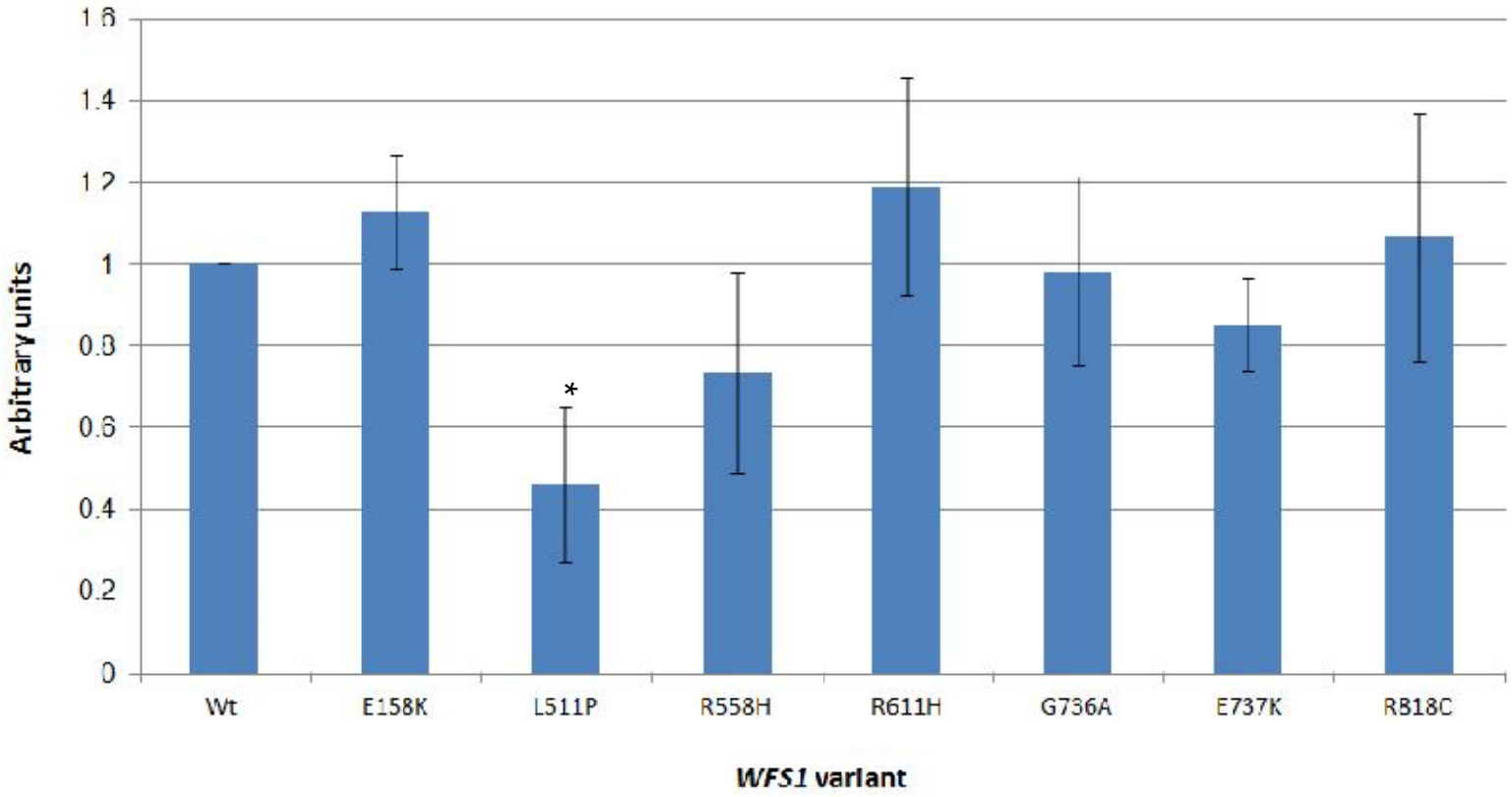


Figure 4.5 Mean steady-state expression wolframin variants relative to wild-type (n=4). Clone set 1 harvested in RIPA buffer. Mean steady-state expression of *WFS1*-L511P is significantly lower than wild-type (Wt) *WFS1* (p.=0.047). The remaining variants show no significant differences in expression.

WFS1 variant	Mean relative expression Arbitrary units (n=4)	STD error	T test (P.value)
Wt	1	0	-
E158K	1.127652	0.139403	0.367957
L511P	0.461953	0.189779	0.046646
R558H	0.734445	0.245622	0.300435
R611H	1.188946	0.266582	0.473075
G736A	0.983187	0.230494	0.93818
E737K	0.850727	0.114011	0.227761
R818C	1.064962	0.304151	0.82111

Table 4.4 Mean steady-state expression wolframin variants relative to wild-type (n=4) in arbitrary units. Clone set 1 harvested in RIPA buffer. T-Tests comparing differences in expression of wolframin variants to wild-type.

Testing possible confounders of low expression detection

4.3.3 Antibody specificity

To confirm that our observations were a realistic reflection of expression levels, we aimed to explore a number of other possible influencing factors. The first consideration was whether or not introduced *WFS1* mutations affect the binding capacity of the antibody. For example, the L511P mutation may alter the epitope site, resulting in reduced binding efficiency and a weaker signal following immunoblotting. To address this question we re-ran a selection of the samples on SDS-PAGE gels and blotted using a second polyclonal anti-*WFS1* Ig (Hofmann, 2003) raised to a slightly different range of target peptides (amino acids

1-285). The resulting blots corresponded well with the initial findings (fig 4.4B), where the only significant difference in expression was seen in *WFS1*-L511P. Thereafter our initial antibody of choice was assumed appropriate for subsequent experiments.

4.3.4 Quality of plasmid DNA preparations

To ensure the quality of the DNA preps had no bearing on transfection efficiency or expression levels we produced fresh maxipreps (see section 2.7.5) from master stocks. The subsequent sets of experiments were direct repeats of the initial transfections and protein harvest in RIPA buffer. Relative steady-state expression of wolframin variants (n=3) were measured once again and the resulting data substantiated our earlier findings (fig 4.6, table 4.5). Again, *WFS1*-L511P differed significantly in steady-state expression relative to wild-type *WFS1* (11% ±7.2%, p.=0.004).

There was a considerable degree of variation between experiments, particularly notable in *WFS1*-R611H, which deviated from the mean significantly. We additionally found that *WFS1*-R818C expression was significantly lower than wild-type *WFS1* in this set of experiments (59% ±7.9% p.=0.024). The discrepancies found here, also notable for *WFS1*-L511P, where the steady-state expression of clone set 2 was found on average to be ~35% lower relative to wild-type wolframin than clone set 1, perhaps highlight the importance of further controls to account for differences in transfection efficiency.

Despite these fluctuations, none of the results refuted our initial findings, and moreover we were able to confirm independently in our system that *WFS1*-L511P expression levels differ significantly relative to wild-type *WFS1* (see appendices 9 and 10 for raw data).

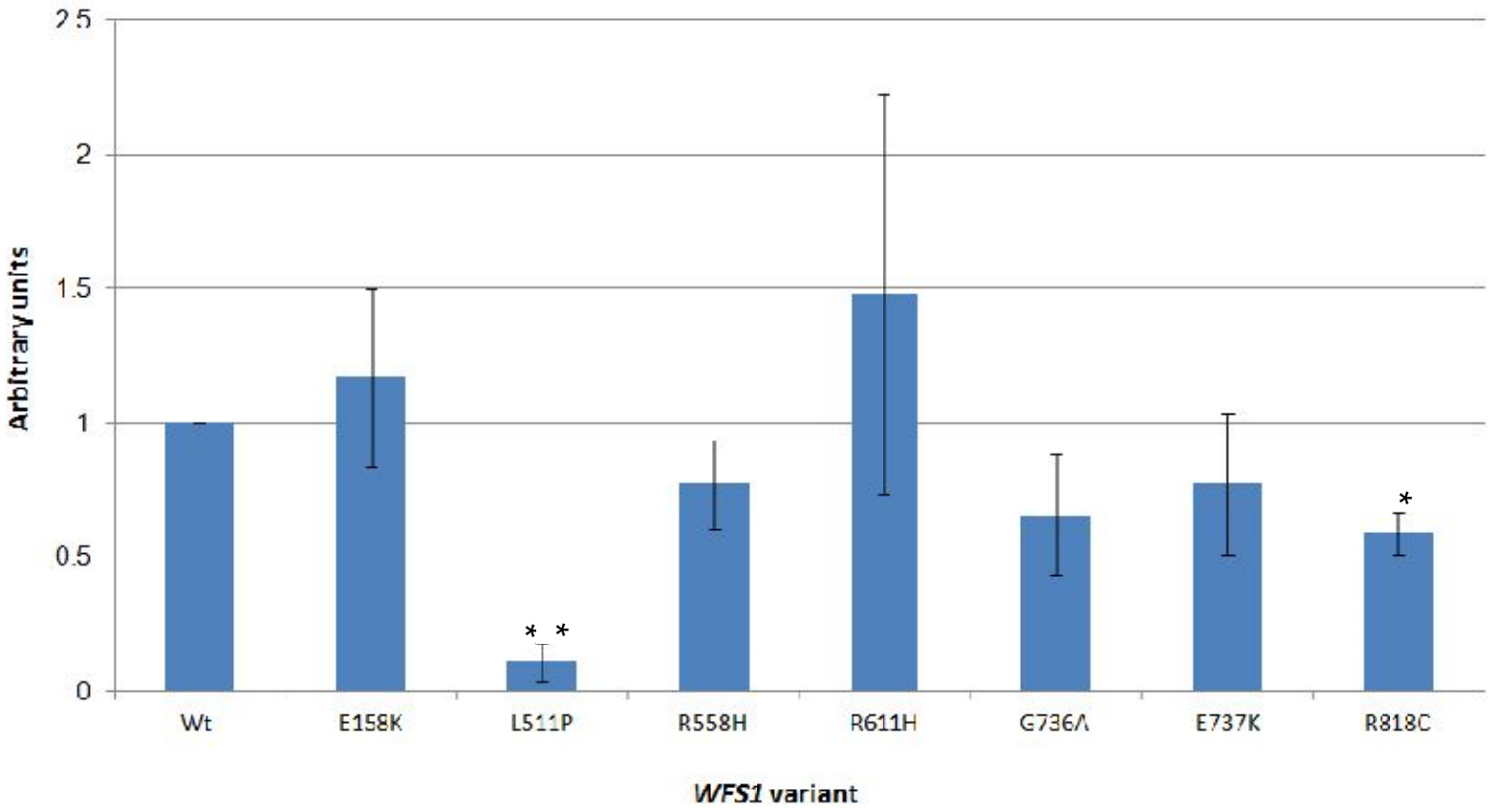


Figure 4.6 Mean steady-state expression wolframin variants relative to wild-type (n=3) in arbitrary units. Clone set 2 harvested in RIPA buffer. Mean steady-state expression of *WFS1*-L511P and *WFS1*-R818C is significantly lower than wild-type (Wt) *WFS1* (p.=0.004 and p.=0.024 respectively). The remaining variants show no significant difference.

WFS1 variant	Mean relative expression Arbitrary units (n=3)	STD error	T test (P.value)
Wt	1	0	-
E158K	1.166538	0.330703	0.600243
L511P	0.110128	0.071518	0.004278
R558H	0.772607	0.164493	0.232519
R611H	1.479586	0.747182	0.514149
G736A	0.655815	0.226108	0.203288
E737K	0.7717	0.263097	0.399241
R818C	0.589061	0.079265	0.023918

Table 4.5 Mean steady-state expression wolframin variants relative to wild-type (n=3). Clone set 2 harvested in RIPA buffer. T-Tests comparing differences in expression of *WFS1* variants to wild-type.

4.3.5 Differences in wolframin solubility

Finally, we investigated potential differences in protein solubility, which could result if mutations sufficiently alter the physio-chemical characteristics, and overall conformation of a protein. It is conceivable that our introduced *WFS1* mutations could change the chemistry of the protein such that the proportion of total wolframin content released to the supernatant during sample preparation may differ significantly between variants, obscuring the accuracy of our results. We endeavoured to address this issue by firstly comparing the amount of wolframin in the insoluble fraction of cell debris following protein harvest in RIPA

buffer. Secondly we directly repeated the earlier transfection experiments, this time harvesting the protein in an alternative lysis buffer (Laemmli).

4.3.6 Crude measures of wolframin solubility in RIPA buffer

Steady-state wolframin expression was measured in both the soluble and insoluble fractions of samples harvested in RIPA buffer (n=3). The soluble proteins were prepared for western blotting as before following sonication and centrifugation. The insoluble fraction of cell debris, which forms a small pellet after centrifugation, was resuspended in RIPA buffer and mixed with vigorous pipetting. Despite the technical limitations of homogenising and loading insoluble material accurately, this method was intended to give a crude indication of any major inconsistencies in solubility between the variants.

Mean expression of each wolframin variant relative to wild-type was calculated over three repeat experiments for both soluble and insoluble fractions (table 4.6) and the resulting figures were compared (fig 4.7). No significant differences in solubility were found between wild-type *WFS1* and any of the variants: L511P, R611H, G736A, E737K, or R818C. However, a significantly smaller proportion of insoluble *WFS1*-E158K (p.=0.026) and *WFS1*-R558H (p.=0.002) proteins were measured in comparison to wild-type.

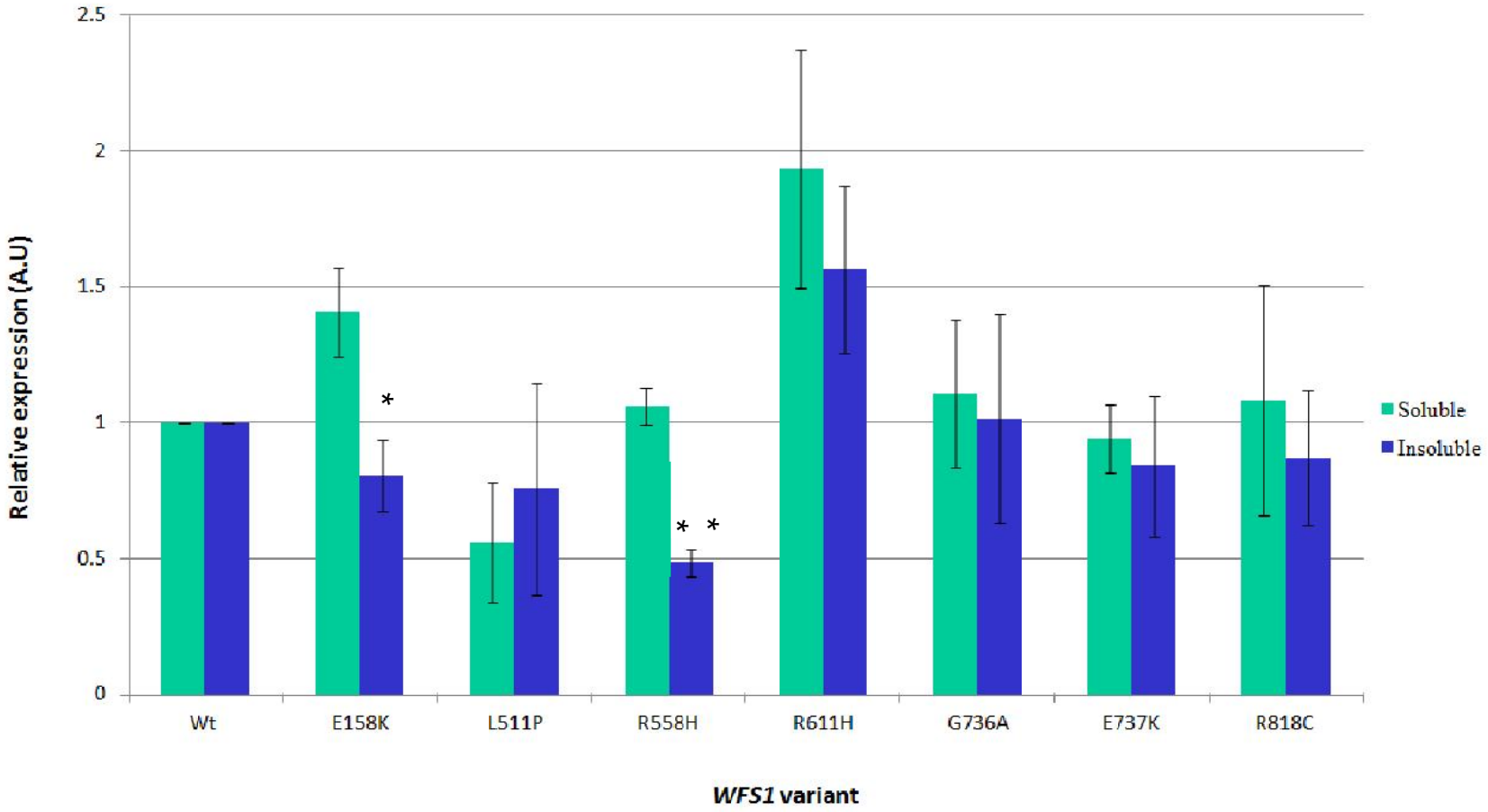


Figure 4.7 Relative steady-state expression of wolframin variants in both soluble and insoluble fractions of RIPA buffer lysates (n=3), measured in arbitrary units. *WFS1*-E158K and *WFS1*-R558H expression levels in insoluble fractions were found to be significantly lower than wild-type (Wt) *WFS1* (p.=0.026 and p.=0.002 respectively). The remaining variants show no significant difference.

WFS1 variant	Soluble	STD error soluble	Insoluble	STD error insoluble	t-tests (P.value)
Wt	1	0	1	0	-
E158K	1.406818	0.16148	0.80578	0.129087	0.025506
L511P	0.559116	0.219817	0.756536	0.38924	0.624419
R558H	1.060473	0.067853	0.485287	0.049123	0.001629
R611H	1.935425	0.437693	1.564174	0.309044	0.448862
G736A	1.10979	0.272102	1.016532	0.384474	0.821573
E737K	0.940718	0.124176	0.839255	0.257697	0.694454
R818C	1.082418	0.425406	0.868427	0.248395	0.629228

Table 4.6 Mean steady state expression of wolframin variants in soluble and insoluble fractions relative to wild-type wolframin measured in arbitrary units. Transfected HEK293 cells were harvested after 48 hours in RIPA buffer.

Although rudimentary, these findings give credence to our earlier observation of lower *WFS1*-L511P steady-state expression by providing evidence that this variant is not obscured from detection in an insoluble fraction of cell debris. Representative immunoblots (figs. 4.8A and B) highlight the relatively low expression of *WFS1*-L511P relative to wild-type. These data also raise a question about the 'actual' steady-state expression levels of *WFS1*-E158K and *WFS1*-R558H, given that relatively smaller proportions of these mutant proteins were detected in the insoluble fraction compared to wild-type. When looking at the mean measurements for these variants in the soluble state, we observed equal or slightly higher levels of protein expression compared to wild-type of no significant difference. If a larger proportion of total wolframin is soluble in the case of these two variants, our calculation of steady-state expression levels could be over-estimated (see appendix 11-13 for raw data).

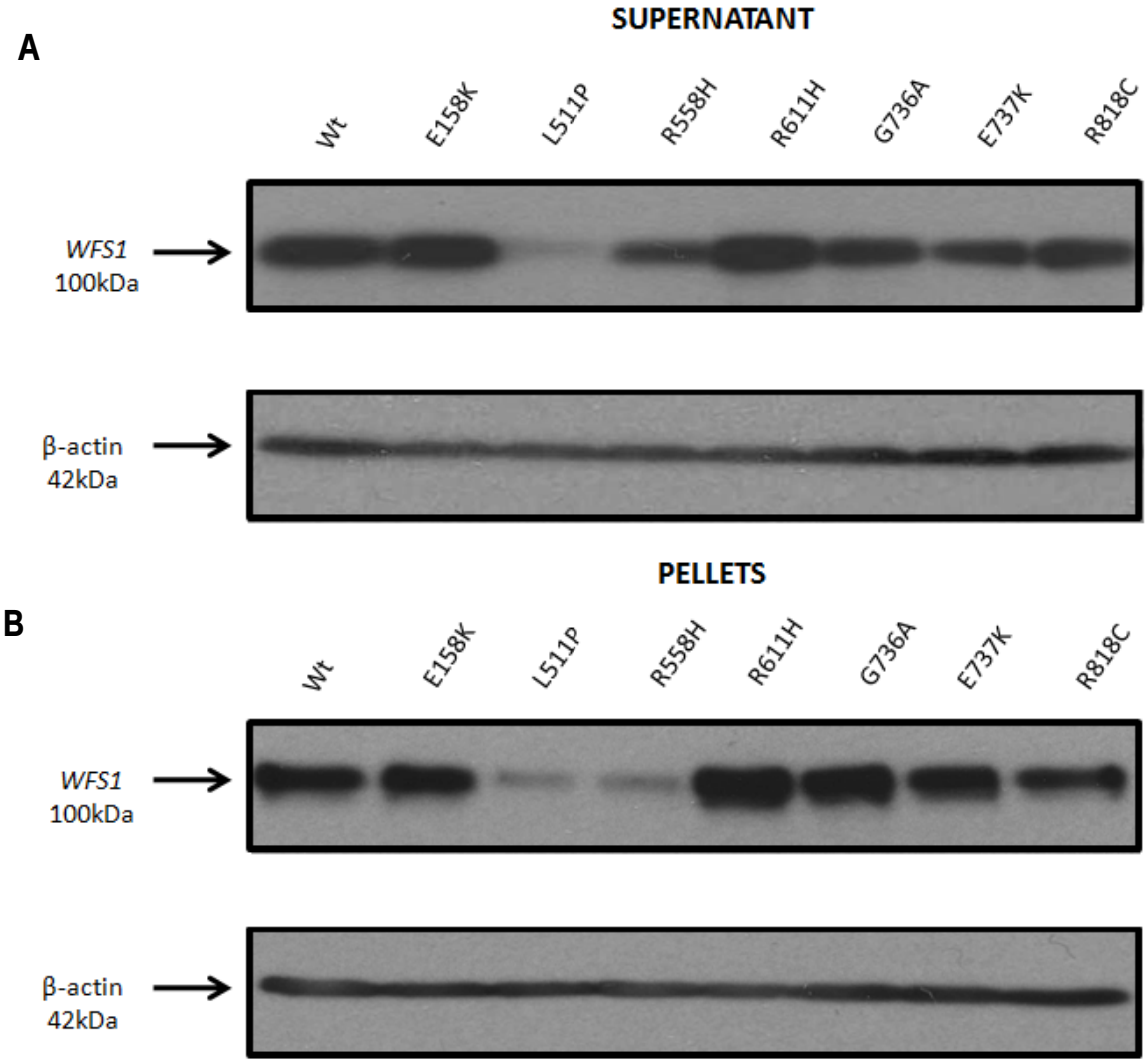


Figure 4.8 Steady-state expression wolframin variants in the soluble and insoluble fractions of RIPA buffer lysate. Typical SDS-PAGE immunoblot analysis of samples probed with anti-*WFS1* Ig (Proteintech) and anti β-actin (Sigma).

4.3.7 Steady-state expression of wolframin variants harvested in Laemmli buffer

Harvesting our wolframin variants in Laemmli buffer allowed us to further assess the potential differences in solubility. Unfortunately we were unable to produce such 'clean' western blots with cells lysed in this buffer compared to those in RIPA buffer. Nevertheless, we repeated transfections three separate times and compared the average expression of the wolframin variants relative to wild-type looking for any obvious differences (figs. 4.9 and

4.10, table 4.7). An appreciable degree of variation between experiments was observed, likely attributed in a large part to the poor quality of the blots. Within the limits of our data, we could however see general expression trends correlating with our earlier experiments using RIPA buffer.

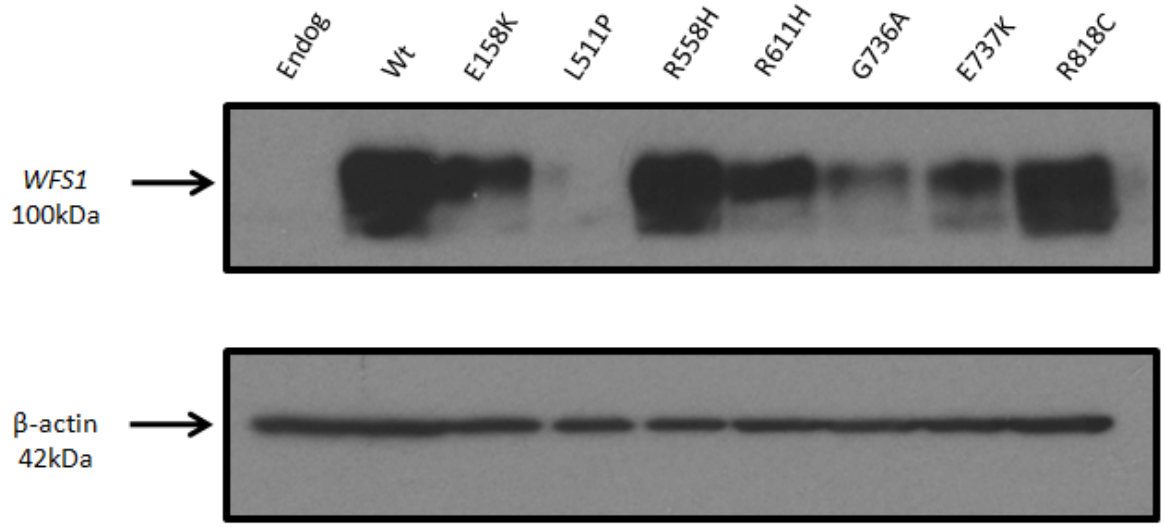


Figure 4.9 Steady-state expression wolframin variants harvested in Laemmli buffer. Typical SDS-PAGE immunoblot analysis of samples probed with anti-*WFS1* Ig (Proteintech) and anti β-actin (Sigma) demonstrating reduced steady-state expression of *WFS1*-L511P.

Distinctly lower levels of *WFS1*-L511P expression relative to wild-type were measured in all three experiments, just above the threshold of statistical significance (31% ±23%, p.=0.069). Steady-state expression measurements for the remaining variants in Laemmli did not differ significantly from those of wild-type wolframin, falling into a relative range of 61%-97%. In contrast to the crude solubility data (section 4.3.6), we recorded no marked differences in the expression of *WFS1*-E158K or *WFS1*-R558H.

Together these findings reinforce our earlier observations in RIPA buffer, further reducing the likelihood of solubility having a confounding effect on our initial results. Most importantly, we have provided additional evidence indicating that the *WFS1*-L511P variant is expressed at a lower steady-state level than wild-type (see appendices 14-15).

WFS1 variant	Relative expression (AU)	STD error	t-test (P.value)	
Endog	0.009223	0.005306	1.91E-05	Wt/Endog
Wt	1	0	-	-
E158K	0.866405	0.102884	0.252729	Wt/E158K
L511P	0.308386	0.233886	0.0685	Wt/L511P
R558H	0.968677	0.33373	0.918986	Wt/R558H
R611H	0.613311	0.177376	0.116305	Wt/R611H
G736A	0.739458	0.295863	0.393589	Wt/G736A
E737K	0.776227	0.145943	0.201186	Wt/E737K
R818C	0.83228	0.303278	0.568051	Wt/R818C

Table 4.7 Mean steady-state expression wolframin variants relative to wild-type (n=3) in HEK293 cells harvested in Laemmli buffer. T-Tests comparing differences in expression of wolframin variants to wild-type.

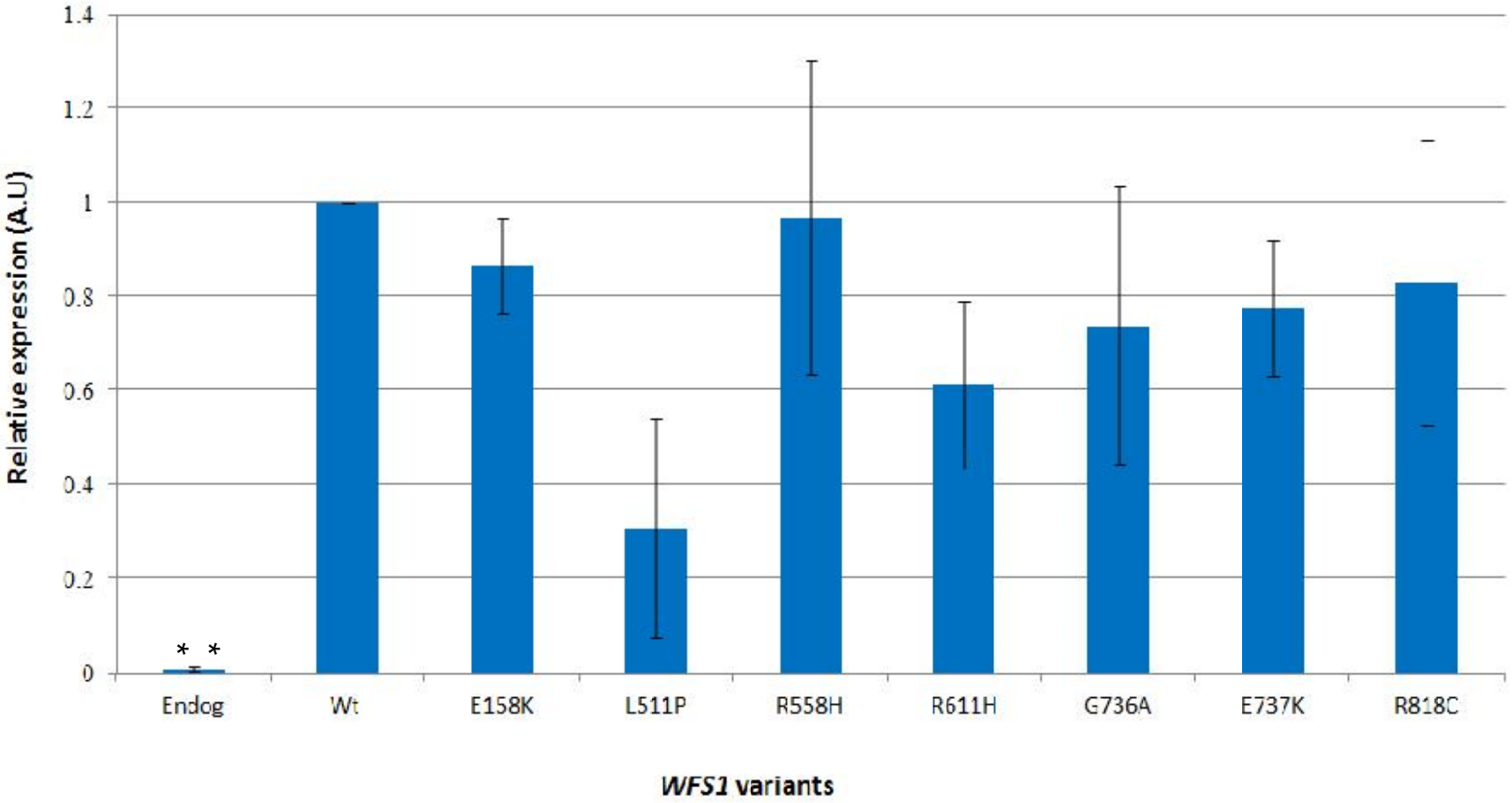


Figure 4.10 Mean steady-state expression wolframin variants relative to wild-type (n=3) harvested in Laemmli buffer.

4.3.8 Summary and discussion

In contrast to the findings of Hofmann *et al* (2006), these experiments provide the first evidence that steady-state levels of some over-expressed wolframin 'missense' variants, even those representing relatively severe WS phenotypes, are comparable to those of wild-type. The majority of our selected mutants demonstrated the capacity for normal expression *in vitro*. One variant, associated with a severe WS phenotype (*WFS1*-L511P) was expressed at a significantly lower level than wild-type. These findings suggest that some, but not all mutations result in reduced expression of *WFS1*. This supports the hypothesis that loss of wolframin function in WS may result from a range of pathomechanisms dependent on genotype.

We substantiated our initial results by ruling out a number of possible influential factors including DNA quality, efficiency of antibody binding, and solubility. We also performed experiments from the stage of transfection a sufficient number of times (n=10) to minimise the effects of variable transfection efficiency.

Each of these additional lines of investigation could be taken further in a number of ways. The wolframin variants could be cloned into tagged expression vectors for the purposes of confirming antibody specificity, although we did perform experiments with two different antibodies reaching the same results. Attempts could be made to measure differences in solubility more accurately by applying a range of buffers and treatments to insoluble proteins. Since we do not attempt to make comments on wolframin solubility in physiological conditions, but merely validate our steady-state observations, we are satisfied that our crude data about relative solubility support our findings and follow a comparable trend. The question of transfection efficiency could also be addressed using quantitative

measures such as the bicistronic reporter assay. In this system the gene of interest (eg *WFS1*) is cloned into an expression vector containing an internally controlled reporter gene, such as GFP (green fluorescent protein). Both genes are expressed equally from the same promoter, so that measurements taken from the reporter gene give us quantitative transfection efficiency information about the gene of interest.

The next step in this investigation is aimed at understanding the mechanisms behind low *WFS1*-L511P expression by firstly looking at relative protein stability. Previous study of selected wolframin missense variants revealed that all of the mutants investigated did have lower steady-state expression than wild-type, which was attributed to a reduction in protein stability (Hofmann, 2006). For this reason we shall firstly compare relative stability of *WFS1*-L511P to wild-type wolframin before considering other potential mechanisms for reduced expression such as transcription disruption, or translation inefficiency due to codon bias. We will additionally seek to further characterise the remaining wolframin variants in readiness for future functional assays.

4.4.0 Comparative stability of wild-type wolframin and selected variants

Background and aims

Earlier studies revealed that wild-type wolframin is a stable protein with a half-life of ~24hours (Yamaguchi, 2004; Hofmann, 2006). Pulse-chase experiments showed that mutations can reduce the stability of wolframin to various degrees, and that cellular depletion of wolframin is mediated by proteasomal degradation. Hofmann *et al* (Hofmann, 2006) demonstrated that wild-type wolframin was stable for a chase period of 6 hours, whereas *WFS1*-P504L and *WFS1*-P724L were reduced to 50%. Other missense and nonsense mutants (R629W, P885L, W700X and F883X), degraded at an even greater rate, closer to background levels after a 6 hour chase period.

Interestingly, the results of these stability assays did not correlate well with the steady-state levels of the same mutants, where *WFS1*-P724L for example, was shown to be one of the most stable variants, yet was detected at comparably low steady-state levels in relation to the others. All of the mutants studied by Hofmann *et al* (Hofmann, 2006) however, were reportedly expressed at lower steady-state levels than wild-type, and all of these products were shown to be relatively unstable. The authors therefore suggested that protein instability is the basis for loss of wolframin function in WS, which is supported by the lower levels of wolframin detected in patient fibroblasts.

In contrast to these findings, only one variant from our panel of *WFS1* mutants (L511P) appears to result in lower steady-state protein expression. Consequently we expect that the majority of our mutants may differ in stability also and reflect a more ‘normal’ pattern of degradation. With consideration to our hypothesis, we predict mutations relating to mild

WS phenotypes and/or those retaining a degree of function to resemble wild-type wolframin more closely. We anticipate that the most likely variant to engender protein instability from our selection is *WFS1*-L511P, and perhaps a rapid rate of degradation explains the lower expression levels observed in our earlier experiments. It is possible however, that other candidates from our panel are equally, or even more unstable, but perhaps expression levels are compensated by enhanced protein production, undetectable by observing steady-state alone.

This section of work investigates the relative stability of our selected wolframin variants. We aim to further characterise our mutants for possible future functional assays, under the hypothesis that some variants have residual function. We also wanted to investigate further the underlying mechanisms behind the reduced steady-state expression levels seen in *WFS1*-L511P to ascertain whether this observation was in fact due to a reduction in protein stability.

4.4.1 Specific methods and approach

To measure relative differences in protein half-life between our *WFS1* variants, we used cyclohexamide in a time-course experiment followed by western blotting. Cyclohexamide (CX), produced by the bacterium *Streptomyces griseus*, inhibits protein biosynthesis in eukaryotic organisms. By targeting specifically the 60S subunit of ribosomes, CX blocks protein translation allowing measurements of protein to be made without the confounding effects of continuing protein synthesis.

This chemical is inexpensive, works rapidly, and has been used extensively in biomedical research for measuring protein degradation, including some work on wolframin (Fonseca,

2010). Although toxic, the application of CX in our experiments eliminates the need for harmful radioactivity, which is necessary for the alternative 'pulse-chase' method, and the protocol itself is less technically complex.

Plates of HEK293 cells were transiently transfected (section 2.9.5) with variants of *WFS1*-pcDNA3.1 (1μg). After 24 hours incubation the cells were treated with CX (section 2.18.0), and harvested in RIPA buffer at time intervals for western blotting. The first samples were harvested immediately following treatment and regarded as time 'zero'. All subsequent harvest intervals could then be measured against time 'zero' to calculate relative degradation of protein over time. Samples were frozen immediately following harvest until the time-series was completed, then sonicated and centrifuged in preparation for western blotting (section 2.13.0). Western blots were probed for *WFS1* with anti-*WFS1* Ig (Proteintech, table 2.2), and also anti-β-actin Ig (Sigma, table 2.2) to enable calculation of equal protein loading. The resulting blots were measured using GeneSnap/GeneTools software (Syngene) to obtain semi-quantitative figures of absorbance (see appendix 16-23 for raw data). For each *WFS1* variant, levels of wolframin over the time course, corrected for endogenous wolframin expression, were compared to levels at time 'zero' to establish the relative degradation. Results from the mutants were compared to those of wild-type *WFS1* to determine any differences in protein stability.

Optimising the protocol to suit our system required a number of pilot experiments to ascertain appropriate time intervals, treatment, and transfection conditions.

4.4.2 Pilot experiments and method development

We harvested protein in a series of initial experiments at time intervals up to 48 hours following transfection with wild-type *WFS1* and treatment with CX at varying concentrations. The earlier reports of Yamaguchi (2004) and Hofmann (2006) showed wild-type wolframin to be stable for at least 6 hours, and reduced to 50% by around 24 hours. Both of these data sets were generated with the pulse-chase method and we therefore aimed to establish a pattern of wild-type wolframin degradation following CX treatment.

Treated cells were firstly harvested at intervals of 0, 2, 4, 6, 12, 24 and 48 hours to give a considerably wide set of intervals from which to gauge an appropriate time range. We expected to observe stable wild-type wolframin between 6 hours and 24 hours, but also wanted to explore the potential for these experiments to run over extended periods in the event that some *WFS1* variants may prolong the half-life of wolframin, or that we may observe more subtle differences in protein stability beyond the 24 hours chase-period.

To ensure our over-expression system was a valid model, we compared the effects of CX on cells both untransfected and transfected with wild-type *WFS1*-pcDNA. We also carried out a range of experiments where cells were transfected with serial dilutions of wild-type *WFS1* plasmid DNA to establish an appropriate level of induced over-expression. We wanted to ensure the differences between endogenous wolframin and over-expressed wolframin could be distinguished, whilst confirming that enhanced protein expression was not sufficient to 'overcome' the action of CX. Additional preliminary experiments compared cells treated with a range of CX concentrations from 50μM to 500μM to look for balance between effectiveness and toxicity. Previous published works of CX assays (Blackinton,

2005; Fonseca, 2010) and manufacturers recommendations (Sigma-Aldrich) were used as guidelines.

A negative control experiment was conducted using a myc-tagged construct containing *WFS1*-N-terminal domain (amino acids 1-321) provided courtesy of Dr Malgosia Zatyka. This represented a nonsense mutation producing severely truncated products, which were expected to be highly unstable. Full length *WFS1*-myc was also included here to allow valid comparisons with the same probe (anti-myc Ig, Sigma).

4.4.3 Pilot Results

Together the pilot studies enabled us to devise a suitable protocol for observing stability of over-expressed *WFS1* in HEK293 cells following 100µM CX treatment. We were able to demonstrate clear differences in *WFS1* expression levels between transfected and untransfected HEK293 cells following western blotting (fig 4.11) but also corresponding patterns of degradation in endogenous wolframin and over-expressed wild-type wolframin (fig 4.12).

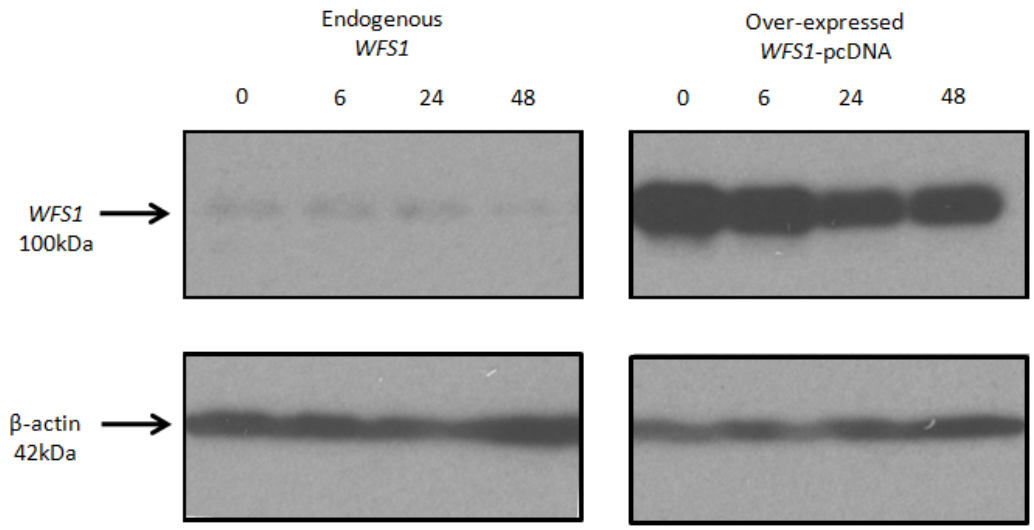


Figure 4.11 Cyclohexamide pilot experiments over a 48 hour chase period. Example immunoblots of wolframin expression in untransfected and transfected HEK293 cells showing clear differences in signal intensity between endogenous and over-expressed *WFS1*. Blots developed after identical exposure time.

In subsequent experiments we could therefore confidently conclude that endogenous protein signals do not interfere with those of our introduced *WFS1* variants, and *WFS1*-pcDNA provides a suitable 'over-expressed' positive control. In agreement with earlier studies these initial experiments suggested that wild-type wolframin remained relatively stable beyond 6 hours.

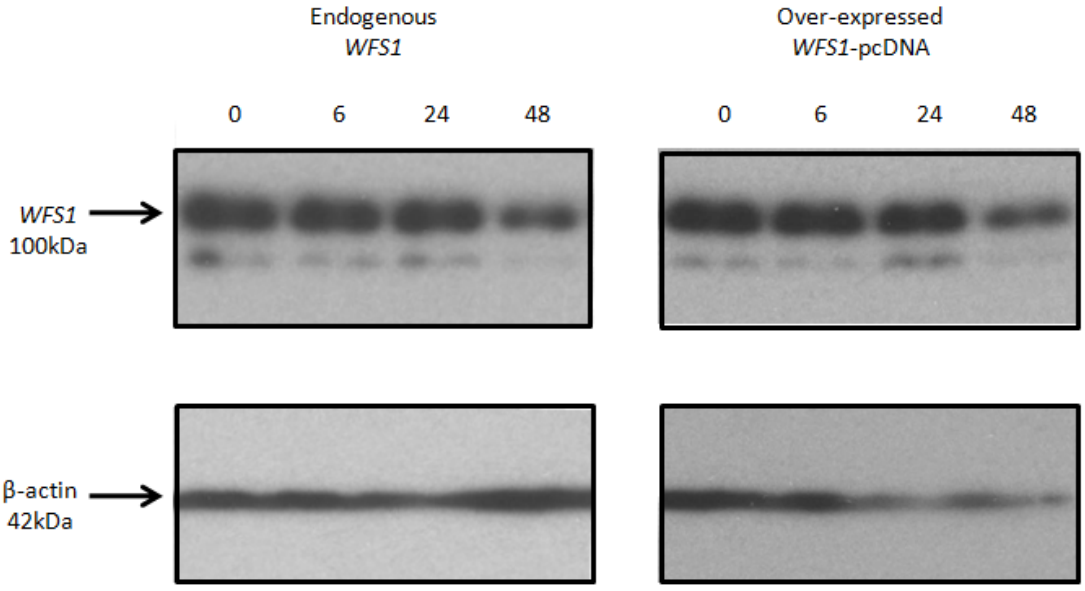


Figure 4.12 Cyclohexamide pilot experiments over a 48 hour chase period. Example immunoblots of wolframin expression in untransfected and transfected HEK293 cells showing comparable patterns of stability. Blots developed after different exposure times.

As predicted, our preliminary results indicated that the *WFS1*-N-term products were highly unstable, reduced to background levels after 6 hours and wild-type *WFS1*-myc behaved similarly to untagged wild-type-pcDNA, remaining stable for periods exceeding 6 hours (fig 4.13).

For the most part, extending the time-course up to 48 hours proved inappropriate as in most instances cells incubated with CX for this period did not survive (appendix 16A-B). We also found a marked degree of fluctuation even at a 24 hours chase-period in repeat experiments treated in the same conditions. Since the action of CX halts the translation process, the proteins necessary for cell survival and more importantly those involved in protein degradation may be the limiting factor. We speculate that this influencing factor may be responsible for the large variations we observed.

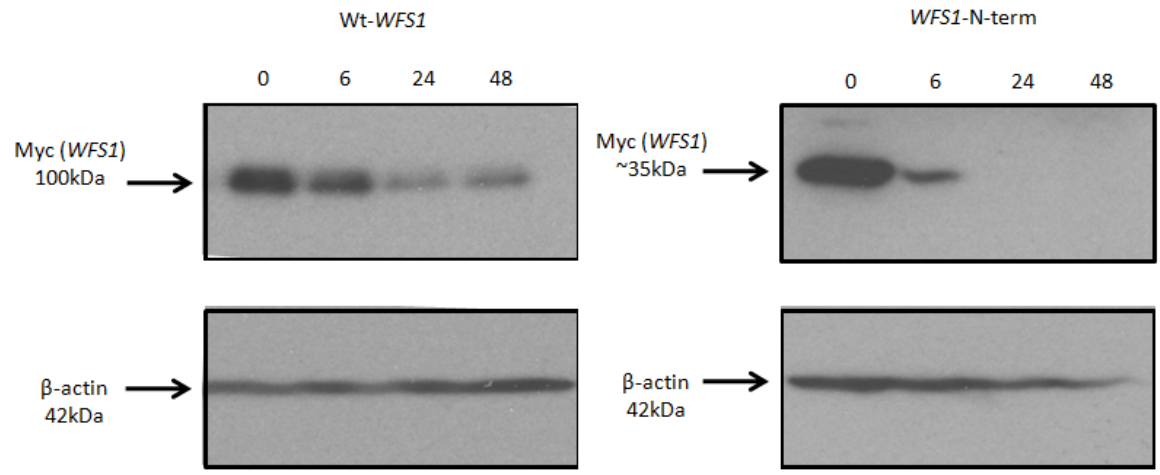


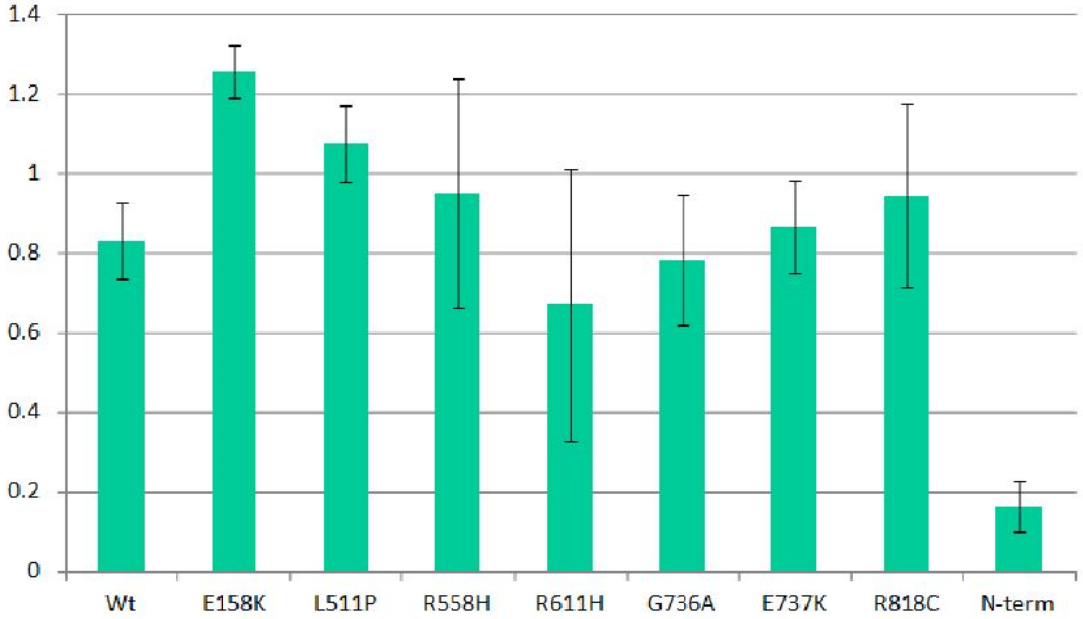
Figure 4.13 Cyclohexamide pilot experiments over a 48 hour chase period. Example immunoblots of myc-tagged wild-type and truncated *WFS1* showing complete degradation of *WFS1*-N-term after 6 hours.

These inconsistencies in our pilot data generated over a 48 hour chase-period suggested that a CX assay over this length of time is unsuitable for our system. We have successfully repeated control experiments up to 6 hours however, observing consistent stability of wild-type wolframin and degradation of our highly unstable negative control products. Further investigation of *WFS1* variants was therefore conducted over a 6 hour time course where we could compare the relative stability of our missense variants to that of an unstable truncated mutant (*WFS1*-N-term). The earlier work of Hofmann *et al* (2006), who reported

significant differences in the stability of missense wolframin variants at a 6 hour chase-period, further supported the suitability of our option for a 6 hour time-course.

Pilot assays including the full range of our wolframin variants were carried out although in light of the 'normal' levels of steady-state expression we observed in previous experiments we expected most of the mutants to display stability similar to that of wild-type wolframin. Figure 4.14 (table 4.8) shows the relative stability of wolframin variants (n=2/3) over a 6 hour chase-period observed in our preliminary experiments (see appendix 17 for raw data). None of these variants appeared to match the high degree of instability seen in the negative control (*WFS1*-N-term) or that of previous mutants examined in earlier studies (Hofmann, 2006). These observations, combined with the steady-state data we have generated, initially indicate that none of our selected wolframin variants result in production of highly unstable proteins. Even with an assay conducted over a limited chase-period we can at least report no protein instability of the magnitude measured in previously investigated wolframin missense variants.

Unlike the mutations studied by Hofmann *et al* (2006), all of our selected *WFS1* mutants except *WFS1*-L511P, appeared to confer 'normal' steady-state levels of protein expression, and perhaps as expected, our preliminary experiments revealed no significant differences in stability of these variants over 6 hours either. Significantly lower levels of *WFS1*-L511P steady-state expression were observed in our earlier experiments (section 4.3.0) so we considered this mutation the most likely candidate to cause protein instability. We therefore performed more detailed analysis of *WFS1*-L511P over a 6 hour time course to confirm our pilot data and included *WFS1*-R611H also due to the relevance of this variant to our future experiments and inconsistency of preliminary data.



WFS1 variant	average	Std error
Wt	0.829159	0.09685
E158K	1.255031	0.06511
L511P	1.074774	0.095871
R558H	0.948864	0.287295
R611H	0.66893	0.340765
G736A	0.783345	0.163916
E737K	0.865641	0.115613
R818C	0.943971	0.228704
N-term	0.16322	0.063567

Figure 4.14 and Table 4.8 Pilot data (n=2/3) of the relative stability of wolframin variants after 6 hours. HEK293 cells transfected with *WFS1* constructs were treated with 100µM cyclohexamide and harvested at time intervals of 0 hours and 6 hours. Following SDS-PAGE and western blot analysis the relative stability of each variant was calculated by comparing protein levels (measured in arbitrary units) at 0 hours and 6 hours post-treatment.

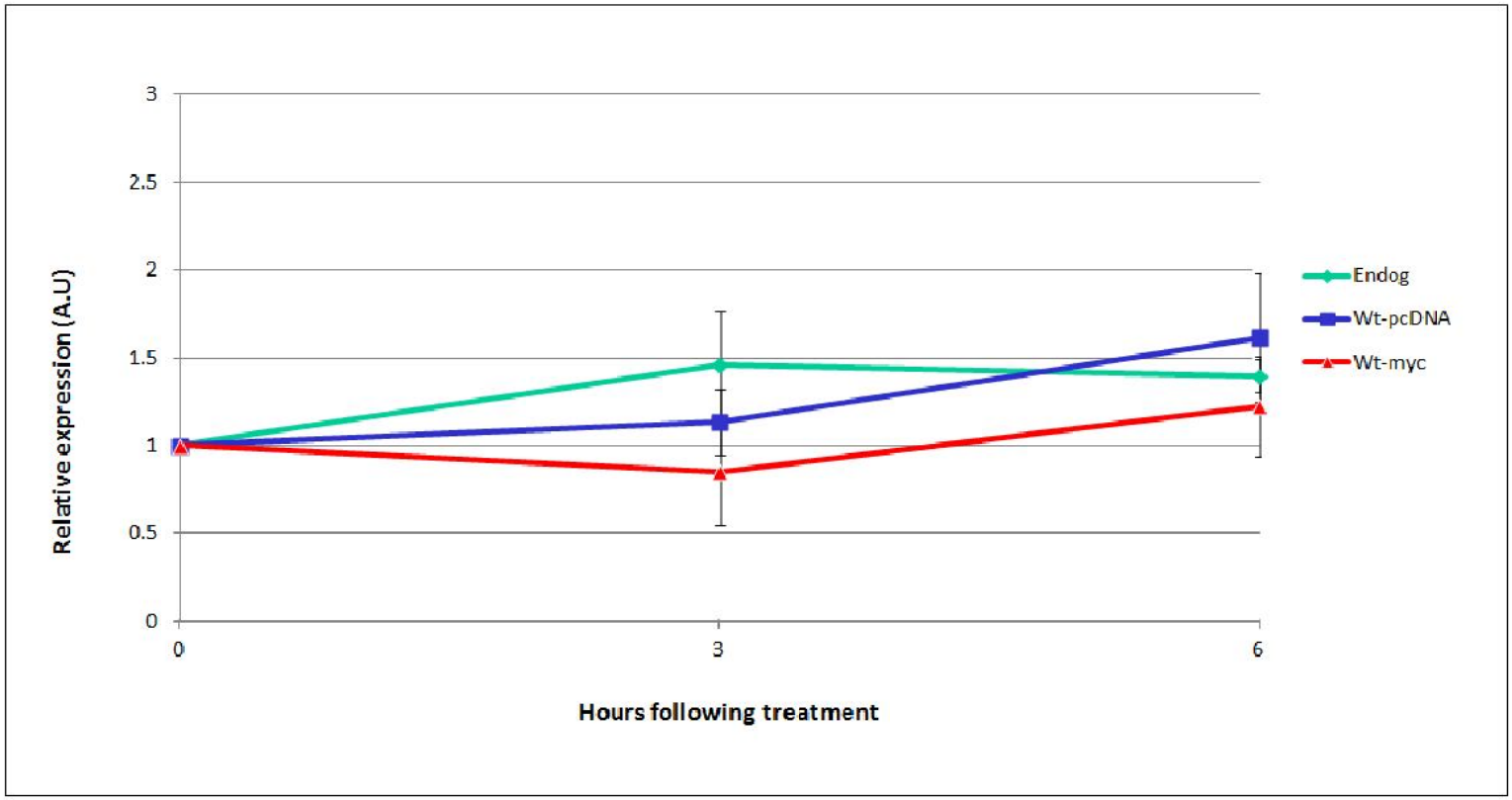
4.5.0 Relative stability of wolframin variants

4.5.1 There are no significant differences in stability between wild-type wolframin, *WFS1*-L511P and *WFS1*-R611H

The first sets of control experiments confirmed that there were no significant differences in stability between endogenous wolframin and over-expressed wild-type wolframin (tagged and untagged) (figs 4.15, 4.18, 4.19). We did not observe any degradation up to 6 hours, and unexpectedly measured slight increases in wolframin. We ruled out over-expression as a cause for the observed increases, since we also measured endogenous wolframin in the same range. These measurements were therefore accepted as reasonable experimental variation.

Significant differences (p.=0.02) were observed between wild-type wolframin and our negative control, which represented a nonsense mutation (*WFS1*-N-term). *WFS1*-N-term was reduced to 16%±6% after a 6 hour chase-period (figs 4.16, 4.18, 4.19) correlating with measurements reported for other *WFS1* nonsense and some missense variants (Hofmann, 2006).

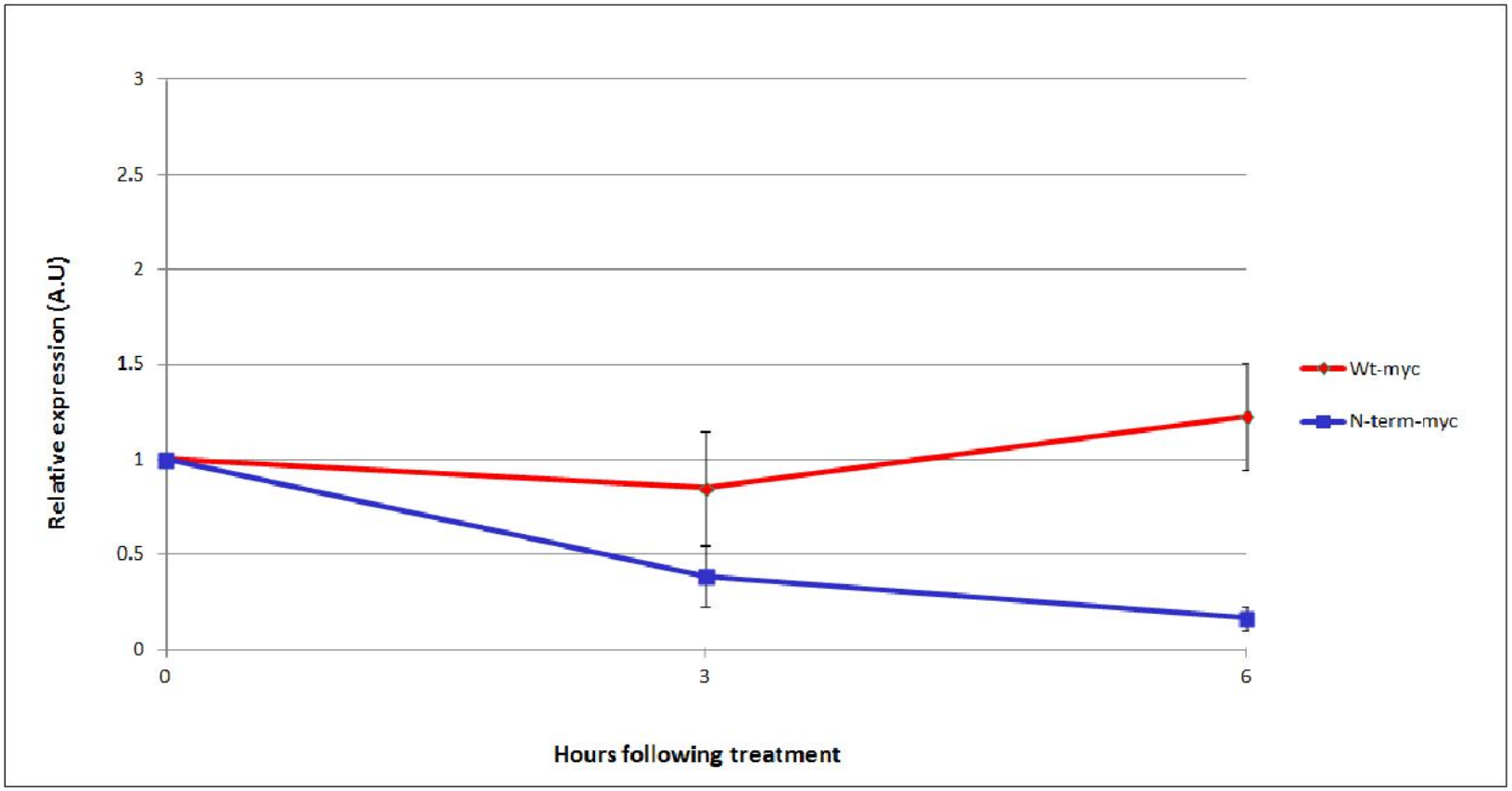
Following treatment with CX, the experimental samples *WFS1*-L511P and *WFS1*-R611H remained stable up to 6 hours chase, and no significant differences were observed between these missense variants and wild-type wolframin (figs 4.17-4.19). Corresponding immunoblots and a summary of the relative protein levels of *WFS1* variants after 6 hours chase period are illustrated in figures 4.18 and 4.19 (see appendices 18-21 for raw data).



	Relative expression (AU)/Time		
WFS1 variants	0	3	6
Endog	1	1.464151	1.396115
Wt-pcDNA	1	1.136026	1.614572
Wt-myc	1	0.848146	1.226077

	STD error/Time		
WFS1 variants	0	3	6
Endog	-	0.30368	0.094119
Wt-pcDNA	-	0.189179	0.363891
Wt-myc	-	0.302071	0.283325

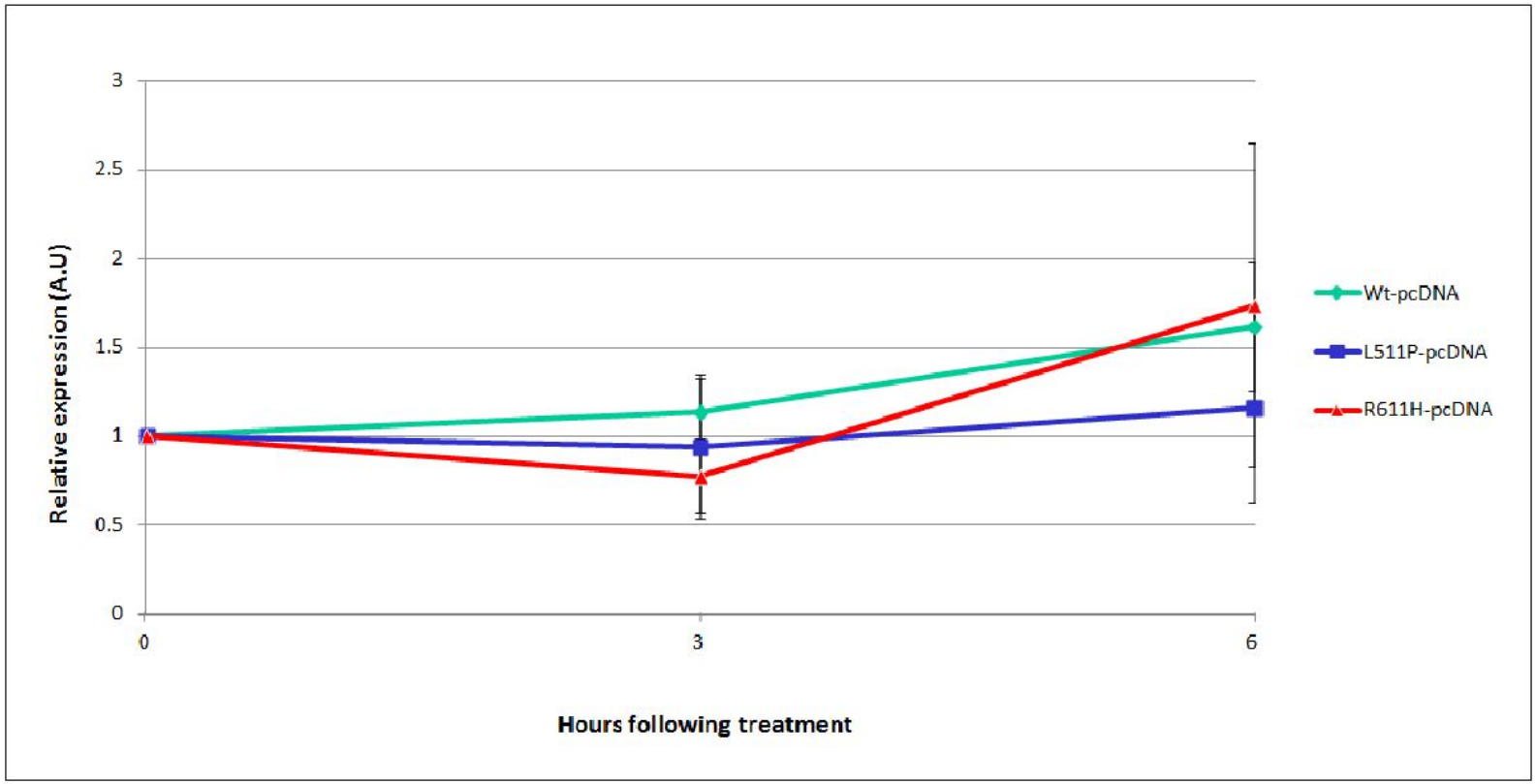
Figure 4.15 Relative stability of endogenous and over-expressed tagged and untagged wolframin (n=3). Immunoblot analysis of wolframin in HEK293 cells over a 6 hour chase period following 100µM cyclohexamide treatment. Expression measured in arbitrary units (AU) relative to time 0 (designated 1).



	Relative expression (AU)/Time		
WFS1 variants	0	3	6
Wt-myc	1	0.848146	1.226077
N-term-myc	1	0.383618	0.16322

	STD error/Time		
WFS1 variants	0	3	6
Wt-myc	-	0.302071	0.283325
N-term-myc	-	0.157414	0.063567

Figure 4.16 Relative stability of over-expressed myc-tagged wild-type wolframin and *WFS1*-N term only mutant (n=4). Immunoblot analysis of wolframin in HEK293 cells over a 6 hour chase period following 100μM cyclohexamide treatment. Expression measured in arbitrary units (AU) relative to time 0 (designated 1).



	Relative expression (AU)/Time		
WFS1 variants	0	3	6
Wt-pcDNA	1	1.136026	1.614572
L511P-pcDNA	1	0.941269	1.157739
R611H-pcDNA	1	0.774277	1.737775

	STD error/Time		
WFS1 variants	0	3	6
Wt-pcDNA	-	0.189179	0.363891
L511P-pcDNA	-	0.29277	0.480335
R611H-pcDNA	-	0.210457	0.910242

Figure 4.17 Relative stability of over-expressed wild-type wolframin and wolframin variants (n=3). Immunoblot analysis of wolframin in HEK293 cells over a 6 hour chase period following 100µM cyclohexamide treatment. Expression measured in arbitrary units (AU) relative to time 0 (designated 1).

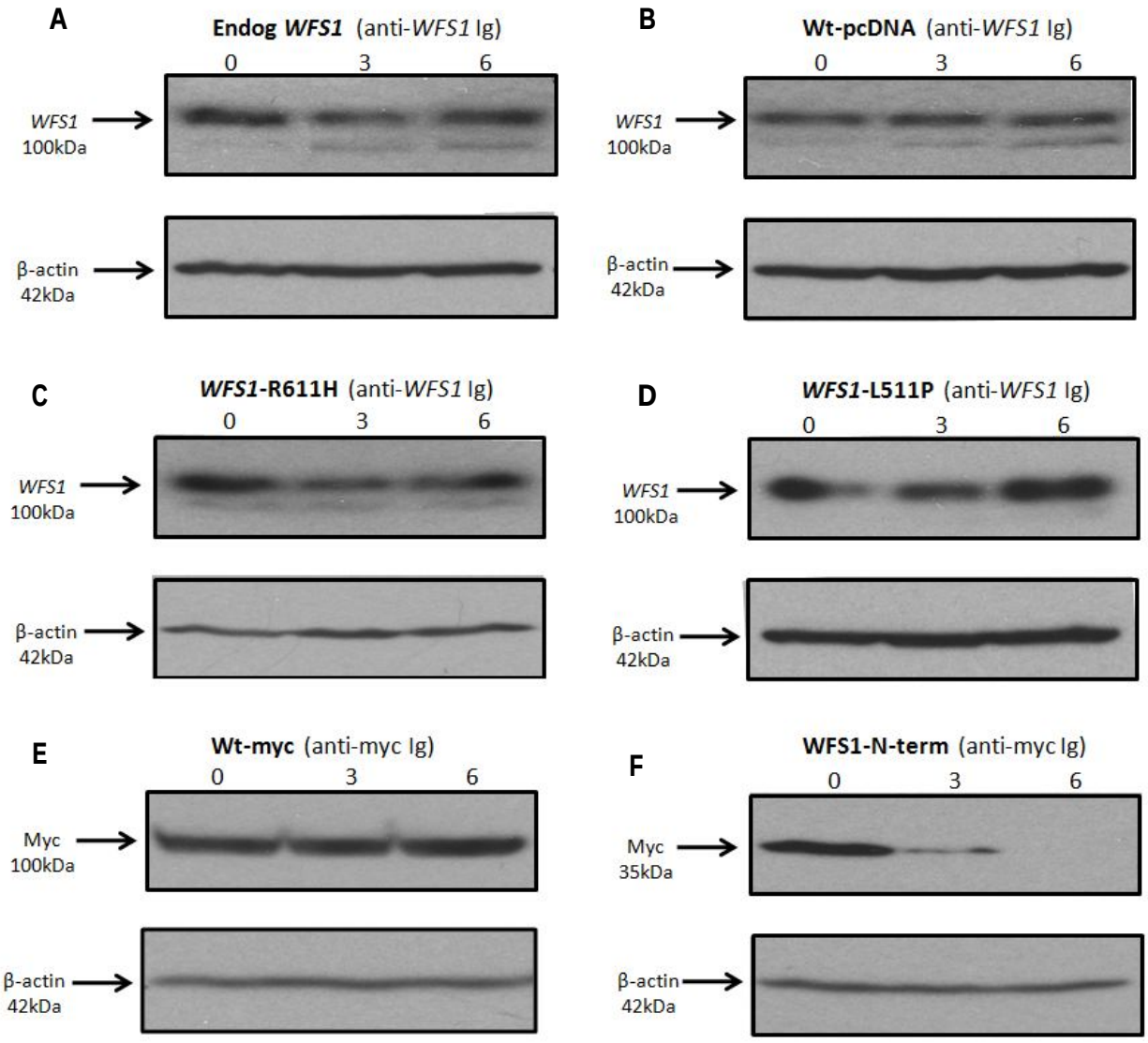
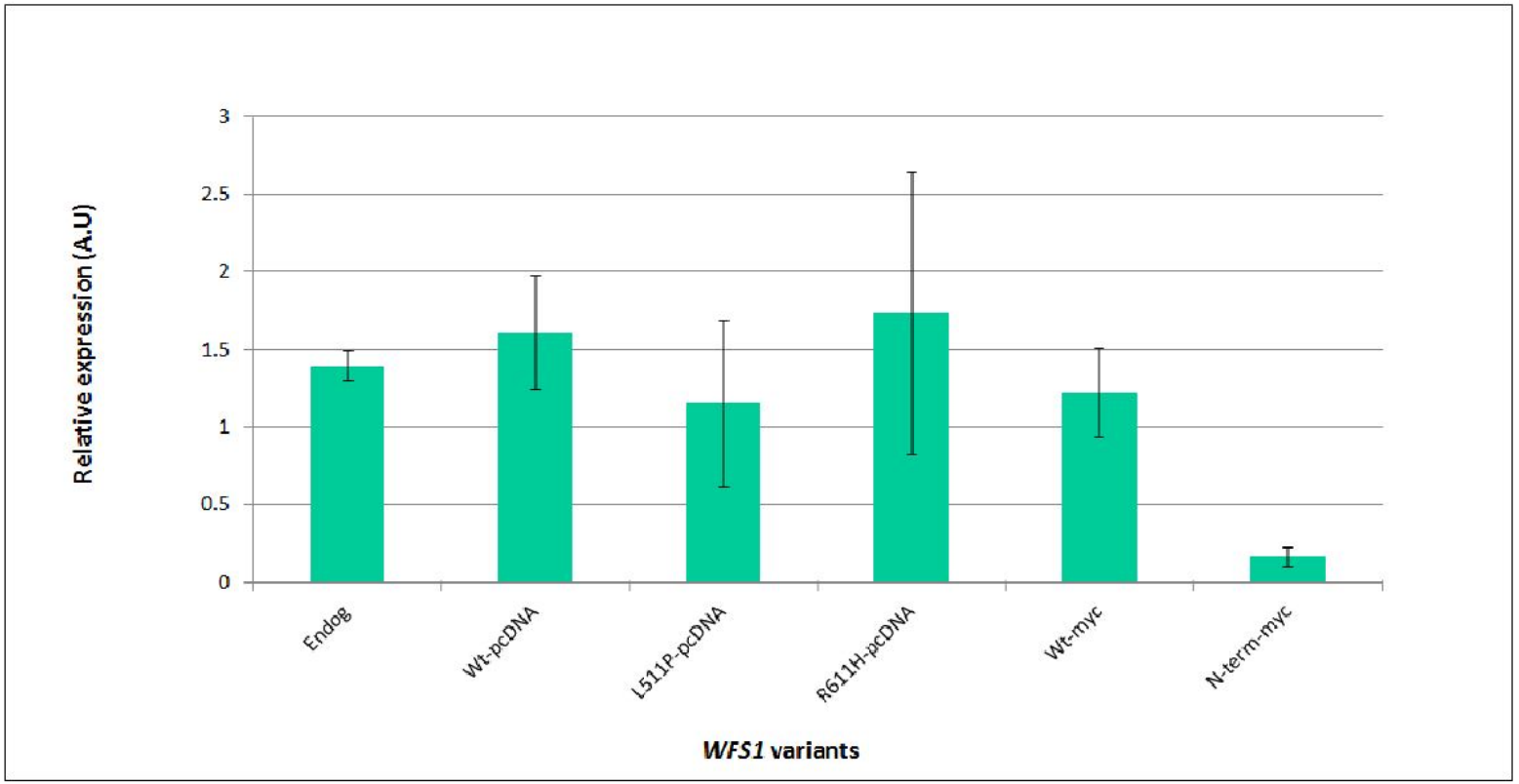


Figure 4.18. Stability of wild-type and mutant wolframin. Immunoblot analysis of wolframin in HEK293 cells over a 6 hour chase period following 100µM cyclohexamide treatment. (**A**) Untransfected cells demonstrating stability of endogenous wolframin after 6 hour (blot developed with prolonged film exposure). (**B-D**) Cells transfected with untagged Wt-*WFS1*, *WFS1*-R611H and *WFS1*-L511P constructs. (**E**) Myc-tagged Wt-*WFS1* detected with anti-myc Ig (**F**) Cells transfected with a myc-tagged negative control construct coding for the N-terminal domain (residues 1-321) of wolframin only, detected with anti-myc Ig.



WFS1 variant	Average relative	STD error
Endog	1.3961153	0.0941192
Wt-pcDNA	1.6145718	0.36389108
L511P-pcDNA	1.1577394	0.5382366
R611H-pcDNA	1.7377745	0.91024166
Wt-myc	1.2260773	0.28332519
N-term-myc	0.1632203	0.06356709

T-test (P.value)	Comparisons
0.542720412	Endog v Wt-pcDNA
0.366456755	Wt-pcDNA v Wt-myc
0.443946025	Wt-pcDNA v L511P
0.888728592	Wt-pcDNA v R611H
0.019984106	Wt-myc v N-term-myc

Figure 4.19 Relative expression of wolframin variants 6 hours following cyclohexamide treatment (n=3/4). Expression measured in arbitrary units (AU) relative to time 0 (designated 1). T tests compare the differences in expression between variants 6 hours post treatment.

4.5.2 Summary and discussion

We have developed an over-expression system in HEK293 cell lines where we can examine the relative stability of wolframin variants over a 6 hour period. Over-expressed wild-type wolframin proved to be stable over 6 hours, in agreement with previous pulse-chase (Yamaguchi, 2004; Hofmann, 2006), and CX experiments (Fonseca, 2010). Also correlating with earlier findings, we demonstrated the relative instability of a severely truncated *WFS1* product (*WFS1*-N-term), representing a nonsense mutation. This 'negative control' protein was reduced to 16% ±6% after 6 hours, significantly different to the relative stability of wild-type *WFS1* after the same chase-period (p.=0.02).

The data generated from our CX assays reveal that the stability of *WFS1*-L511P and *WFS1*-R611H does not differ significantly from that of wild-type. Contrasting with earlier studies of *WFS1* mutants (Yamaguchi, 2004; Hofmann, 2006), our results indicate that some *WFS1* missense variants code for stable wolframin protein products.

This study has enabled us to compare the stability of wolframin missense variants to that of severely unstable protein. Unfortunately we were unable to examine the fate of these variants over extended periods using our system of CX treatment, and therefore cannot comment on any differences in stability up to 24 hours, which may have been revealed. We have however, demonstrated the capacity for some variants (*WFS1*-L511P and *WFS1*-R611H) to remain stable for periods in excess of 6 hours, which is in contrast to the behaviour of truncated (nonsense) products, or of other reported missense variants.

Our choice of a CX assay conducted in a human cell line (HEK293) differs from the pulse-chase methods in primate derived cell lines (COS7) used in previous study of wolframin

stability. It is consequently difficult to make direct comparison of results. Since we found our negative control protein (*WFS1*-N-term) produced corresponding patterns of degradation to nonsense mutants examined with the pulse-chase method, we felt that making general comparisons between the results generated by these two methods was reasonable.

Consideration must be given, as for any chosen assay, to the effect that introduced chemicals such as CX or transfection reagents might have on the folding capacity of the experimental proteins, which would also vary between protocols. We are assuming here that the effects of mutation would cause loss of molecular interactions and not gain, anticipating that proteins with substitutions are more likely to be 'unfolded' and exposed to chemical denaturation, as opposed to more tightly folded and potentially long-lived. Furthermore, based on our physio-chemical data we expect different amino acid substitutions to have varying effects on overall protein conformation; some 'milder' changes in protein folding may affect degradation less markedly, necessitating very sensitive assays covering a longer time-scale.

Detailed analysis of the stability of our mutants over a longer time-course could potentially be carried out by radio-labelling and pulse-chase techniques in order to investigate both of these eventualities. The pulse-chase method has the advantage that translation is not 'switched off' as it is in CX assays, meaning that the cellular machinery required for normal activities such as protein degradation are not exhausted once translation stops. This in theory allows cells to live longer, assuming radioactivity exposure is not sufficient to cause premature cell death, so we can observe activity over a longer period.

The accuracy of our findings could also be further improved with the introduction and measurement of reporter genes such as GFP (green fluorescent protein) at the transfection

stage to correct measurements of protein at each time interval for variation in transfection efficiency (section 4.3.8).

Since only one wolframin variant in our panel demonstrated lower steady-state expression, we anticipated the majority of our mutants to be relatively stable. Surprisingly, although expressed at a reduced steady-state, *WFS1*-L511P remained stable up to 6 hours. This could imply that any relative differences in protein stability caused by this variant may only be detectable beyond 6 hours. Alternatively, reduced protein expression could result from aberration at the RNA level.

This work provides further information about the characteristics of our chosen wolframin variants, suggesting that not all missense mutations reduce wolframin protein stability. Apart from *WFS1*-L511P, corresponding to a severe phenotype, our panel of *WFS1* mutations have shown expression characteristics *in vitro* correlating to those of wild-type, increasing the likelihood of some protein functionality. Our data support the notion that loss of wolframin function in WS could be due to several mechanisms, not all of which relate to protein instability and low expression.

4.6.0 Comparative mRNA levels of wild-type-*WFS1* and *WFS1*-L511P

4.6.1 Background and aims

Previous study of *WFS1* missense mutants revealed a reduction in steady-state wolframin expression levels, which was attributed to protein instability (Hofmann, 2006). We have also found a novel *WFS1* missense mutation (L511P) that induces low steady-state protein expression *in vitro*, but does not however appear to be markedly unstable. To explain the possible causes for low expression, we postulate that this mutation may instead cause perturbation at the stage of RNA synthesis or processing. To explore this possibility we compared RNA levels of over-expressed wild-type-*WFS1* and *WFS1*-L511P. If the differences we observed in protein steady-state were a consequence of aberration at the RNA level, we expected to see relatively lower RNA in samples transfected with *WFS1*-L511P relative to those transfected with equal amounts of wild-type *WFS1*.

4.6.2 Specific methods and approach

Real-time PCR (Q-PCR; section 2.1.3) was used to quantify and compare RNA extracted from HEK293 cells transiently transfected with wild-type *WFS1* or *WFS1*-L511P. The Taqman based detection system (Applied Biosystems) was the method of choice due to a number of technical advantages over alternative Q-PCR protocols such as SYBR Green. The Taqman system uses fluorogenic probes that detect specific amplification products as they accumulate during PCR. This serves to minimise the generation of false-positive signals, a major limitation of traditional protocols using intercalator dyes, which bind any double-stranded DNA.

Applied Biosystems offer a range of pre-designed Taqman assay probes, which have been optimised and validated. We selected FAM-labelled human *WFS1* primers, specific for a sequence in the region of exon 4, which amplify 81bp products (Applied Biosystems, ref. Hs00903605_m1). We deliberately avoided probes covering exon 8, since the substitution relating to codon 511 (nucleotide 1532) is situated in this area, and could interfere with the reaction, skewing results.

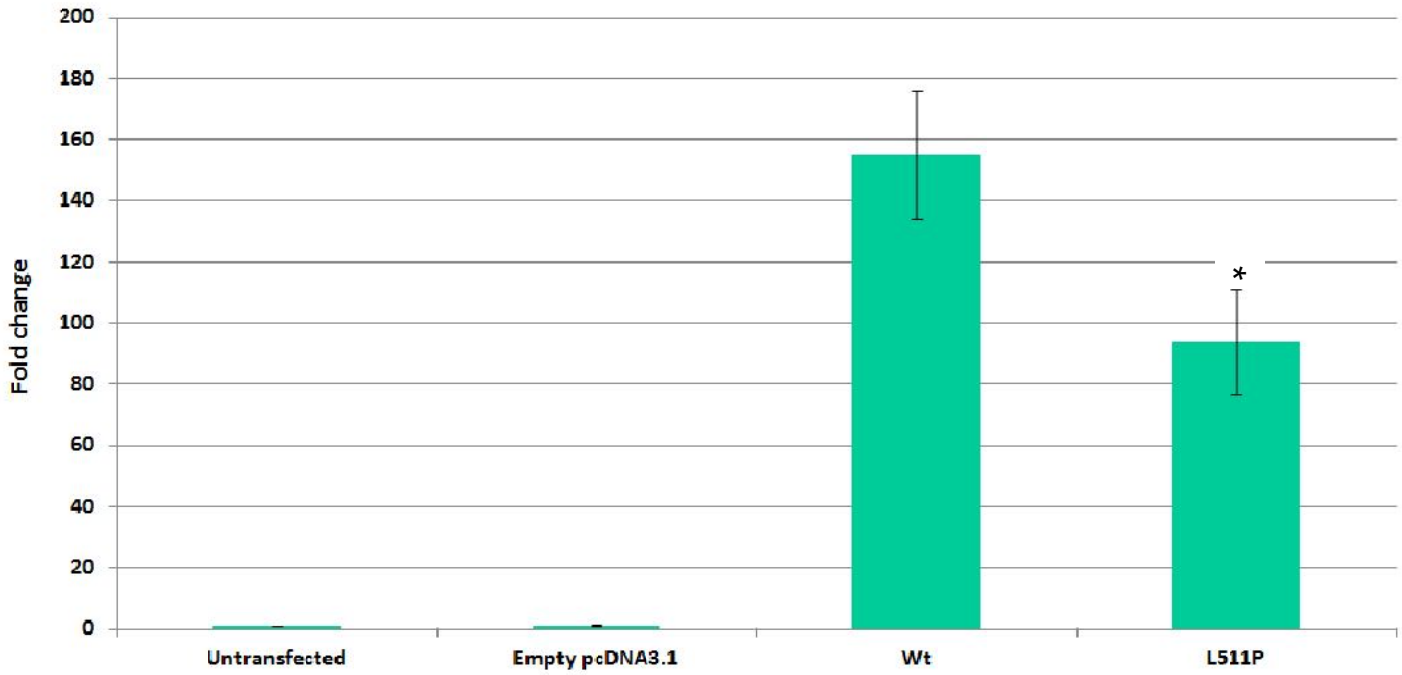
HEK293 cells were grown to 80-90% confluency and transfected (FugeneHD, Roche; section 2.9.5) with either wild-type *WFS1* (pcDNA3.1), *WFS1*-L511P (pcDNA3.1), or an empty pcDNA3.1 control vector. After a 24 hour post-transfection incubation period, total RNA was harvested with TRIzol (Invitrogen; section 2.10.2), reverse transcribed to cDNA (section 2.1.2), and measured (section 2.12.0). Equal amounts (8.75ng) of cDNA, mixed with Taqman Q-PCR mastermix and *WFS1* probes were then cycled on a 7500 Real Time PCR machine and software (Applied Biosystems). Relative measurements of *WFS1* RNA were calculated according to the delta-delta C_t method (section 2.1.3).

4.6.3 Results

To minimise the effects of transfection efficiency the experiment was repeated 6 times including independently generated plasmids before average fold-changes between different sets of conditions were calculated (appendix 24-26). Internal controls were also made within replicates, where several sets of RNA were harvested per transfection, two independent sets of cDNA were synthesised from each sample of RNA, and two 'identical' samples from each cDNA prep were simultaneously run in the real-time reaction. We were able to take mean measurements from across these controls, minimising experimental error at each stage of the procedure. By analysing several corresponding samples we could gauge the

accuracy of RNA/cDNA measurements and pipetting, giving an indication of how reliable our end results were.

Mean fold-changes between different pairs of sample sets were calculated from average delta ct values over 6 transfections (2x cDNA for each). Also for each condition, all twelve sets of delta ct values were applied in pairs to a two-tailed t-test to compare the differences (fig. 4.20). As expected, we found the mean expression of *WFS1* RNA in samples transfected with a control empty vector (pcDNA3.1) was not significantly different from endogenous levels of *WFS1* in untransfected controls, with a fold change of 1.074±0.10. Also the samples over-expressing both wild-type *WFS1* and *WFS1*-L511P had significantly higher *WFS1* RNA compared to background levels, with mean fold changes of 155.20±20.88 and 94.03±17.27 respectively. Analysis of the experimental samples revealed significantly lower levels of *WFS1* RNA in cells over-expressing *WFS1*-L511P compared to wild-type (p.=0.026), with an average overall reduction of ~37% (fold change=0.63±0.10) (fig. 4.21).



	Fold change	Std error
Untransfected	1	0
Empty pcDNA3.1	1.074039932	0.098778
Wt	155.1989716	20.8841
L511P	94.02548345	17.27463

Comparisons	t-test p.value
Untransfected/Empty	0.743596358
Untransfected/Wt	1.77396E-13
Untransfected/L511P	1.22345E-11
Wt/L511P	0.025712256

Figure 4.20 Relative mRNA expression of *WFS1* variants (n=12) measured by Real-time PCR. *WFS1* expression in cells transfected with: wild-type *WFS1*, *WFS1*-L511P and an empty control vector (pcDNA3.1) relative to untransfected cells. Two-tailed students t-tests compare sets of 12 del-ct values (from 6 transfections) confirming a significant reduction in *WFS1*-L511P compared to wild-type *WFS1* (p.0.026).

	Fold change	Std error
Wt-WFS1	1	0
WFS1-L511P	0.633928213	0.09596

Figure 4.21 Relative mRNA expression of wild-type and mutated *WFS1* measured by Real-time PCR (n=12). Significantly lower levels of *WFS1* mRNA (p.=0.026) were measured from samples derived from cells carrying *WFS1*-L511P, which showed an average overall reduction in *WFS1* expression of 37% compared to wild-type control samples.

4.6.4 Summary/conclusions

Due to the massive upregulation of gene expression following transfection, we should perhaps expect a significant degree of mRNA variation when measuring levels with sensitive quantitative assays. We did observe fluctuations in *WFS1* RNA both within and between experiments, despite the accuracy of sample preparation and measurements assured by consistency of β-actin measurements. This variation could be explained by differences in transfection efficiency, which we aimed to minimise by conducting several direct repeat experiments with more than one independent preparation of plasmid-DNA. An additional improvement we may consider in future is the analysis of an internal control gene, present on the expression vector (eg. neomycin), which could be used to normalise for transfection efficiency more accurately.

Overall these data suggest that less *WFS1* mRNA is generated from the L511P variant compared to wild-type *WFS1*, and provides a possible explanation as to why the L511P variant has significantly lower steady-state protein expression than wild-type wolframin. The basis for reduced *WFS1*-L511P mRNA could be related to hindrance of the transcription process resulting in relative inefficiency and synthesis of fewer mRNA molecules per unit of time. Alternatively the transcript itself may be less stable than wild-type, the mutation conferring increased susceptibility to RNase degradation, a feature common to the development of other diseases such as cancer, inflammatory and autoimmunity disorders (Knapinska, 2005).

This study provides the first evidence that missense variants may reduce normal *WFS1* mRNA levels, proposing this mechanism as a feature of WS pathogenesis.

4.7.0 Localisation of wolframin variants

4.7.1 Background and aims

Immunofluorescence microscopy analysis (Hoffman, 2003; Fonseca, 2005; Takeda 2001; Osman, 2003) and sub-cellular fractionation studies of mouse brain (Philbrook, 2005) and human fibroblasts (Takeda, 2001) reveal that wolframin is predominantly situated on the endoplasmic reticulum (ER) membrane. Although the precise function is still unclear, wolframin is implicated in a number of pathways including protein trafficking, calcium homeostasis, and the ER stress response. In any given role we can reasonably assume correct positioning of the protein is crucial for the propagation of normal cell signal transduction and molecular interactions. Confirmation of appropriate sub-cellular localisation of wolframin variants is therefore an important prerequisite to future functional studies.

Earlier work examining localisation of selected wolframin variants showed that some mutations (*WFS1*-P724L and *WFS1*-G695V) appear to cause a degree of protein aggregation, and form relatively insoluble, higher-molecular weight complexes compared to wild-type wolframin (Fonseca, 2005). Interestingly these variants both involve changes to proline and glycine residues, which as mentioned earlier, are thought to be conformationally important amino acids. Contrasting results were reported by another group (Yamaguchi, 2004), who found their mutated *WFS1* constructs (N663D and N748D*) produced proteins with a 'normal' ER pattern of expression seen under the microscope. Taken together, these studies suggest the fate of mutated *WFS1* products may vary depending on genotype.

The aim of this section was to investigate the sub-cellular localisation of our panel of wolframin variants. We expect any proteins with at least partial function to localise correctly, and if this can be demonstrated, we may have grounds to explore the impairment/loss of function in more detail.

Our general hypothesis states that some missense *WFS1* mutations may allow partial wolframin function. From the spectrum of reported mutations, those associated with milder WS phenotypes (E158K and R558H), or non-syndromic diabetes susceptibility (R611H) are most likely to retain some functionality so consequently, we anticipate these variants are the most likely candidates to localise normally.

4.7.2 Specific methods

We used immunocytochemical analyses to look at cells overexpressing our variants of interest (section 2.15.0). Cell lines were transfected with *WFS1*-pcDNA or *WFS1*-myc plasmids containing the sequence variants, incubated for 24 hours, then fixed onto glass microscope slides with 4% paraformaldehyde. The localisation of wolframin was determined with a combination of anti-*WFS1*, anti-myc, and anti-PDI (ER marker) antibodies (table 2.2).

Results

4.7.3 Pilot experiments

A series of preliminary studies were conducted to ascertain appropriate experimental conditions. These included choosing suitable cell lines, antibody optimisation, and determining acceptable conditions for cell growth, transfection and fixation.

4.7.4 Choice of cell line

The obvious choice of cell lines for our localisation experiments were HEK293 and COS7. Both of these cell lines are easily grown and manipulated in the laboratory. HEK293 cells have already been successfully transfected with our *WFS1*-pcDNA constructs in earlier experiments and being derived from humans, have the advantage of providing perhaps a more accurate model for studying human *WFS1*. Earlier wolframin immunocytochemical studies (Yamaguchi, 2004; Fonseca, 2005; Hoffman, 2006) alternatively used COS7 cell lines, which originate from monkeys (*Chlorocebus pygerythrus*). COS7 cells are larger and rounder than HEK293 cells (approx 30 μm and 13μm respectively) making it comparably easier to distinguish their substructures under the microscope. These cells are also readily transfected and are most appropriate for making direct comparisons between our findings and those previously published by other groups.

Initially, HEK293 and COS7 cell lines were cultured, fixed, probed for wolframin and protein-disulphide isomerise (PDI) independently, and viewed under the fluorescent microscope. A range of antibody concentrations were used, selected as per manufacturers recommendation (Proteintech, Stressgen). We aimed here to detect endogenous wolframin, and to detect endogenous PDI, which is a ubiquitously expressed ER resident protein intended for use as an organelle marker. We also wished to determine which of the cell lines was most fitting for subsequent experiments.

A rabbit polyclonal anti-*WFS1* Ig (Proteintech) was used as a primary antibody for the detection of wolframin, followed by an anti-rabbit secondary Ig conjugated to Alexa-Fluor 488 (green). The presence of PDI was detected with a mouse monoclonal anti-PDI Ig (Stressgen), followed by an anti-mouse secondary Ig conjugated to Alexa Fluor-594 (red).

Examples of the wolframin and PDI staining patterns are shown in figure 4.22 (x40mag). Endogenous levels of both proteins were detected in COS7 cells (left panels) and HEK293 cells (right panels), correlating well with expected staining patterns detailed in the literature. The differences in cell morphology were also immediately apparent (fig 4.23). The COS7 cells grew in a more uniform monolayer making them easier to distinguish from each other, whereas the HEK293 cells had a tendency to form clumps. The larger size and circular, flattened shape of the COS7 cells also enabled better visualisation of the sub-structures, compared with the small, angular HEK293 cells with less obvious features. For these reasons, COS7 cells were the cell line of choice used for all our subsequent localisation experiments.

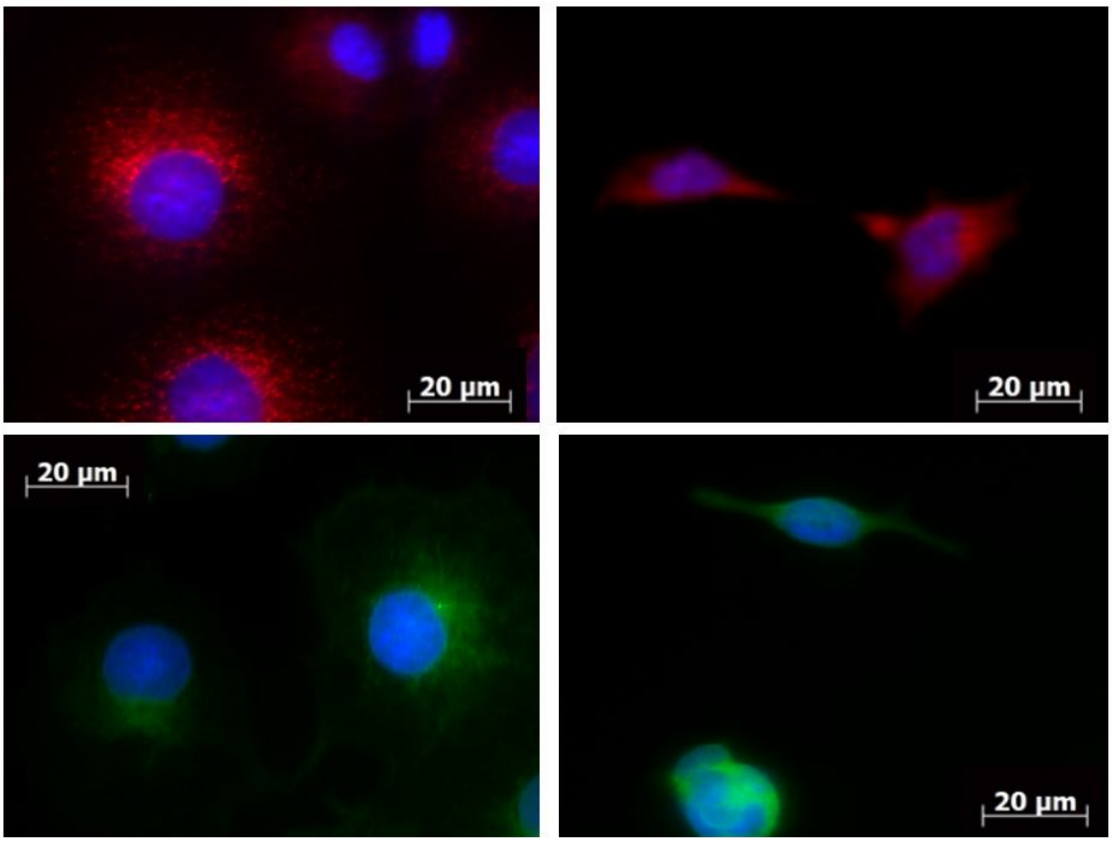


Figure 4.22 Cells fixed and stained with anti-PDI or anti-*WFS1* Ig (x40 mag). TOP PANELS: COS7 (left) and HEK293 (right) stained with anti-PDI Ig demonstrating detection of endoplasmic reticulum. LOWER PANELS: COS7 (left) and HEK293 (right) stained with anti-*WFS1* Ig demonstrating detection of wolframin.

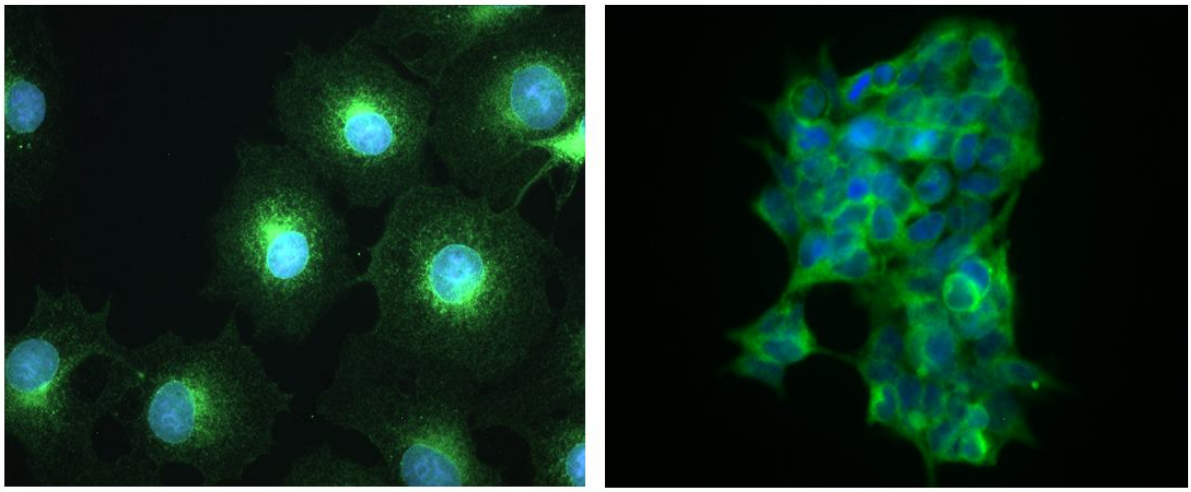


Figure 4.23 COS7 and HEK293 cell lines fixed and stained with anti-WFS1 Ig (x40 mag), contrast adjusted. LEFT PANEL: Evenly distributed COS7 cells stained with anti-*WFS1* Ig demonstrating a typical ER pattern of staining with signal most concentrated in an area to one side of the nucleus (blue) and spanning out to the periphery of the cell. RIGHT PANEL: HEK293 cells stained with anti-*WFS1* Ig; Cells forming a clump, with indistinguishable cells.

4.7.5 Antibody optimisation and transfection

Fixed COS7 cells were stained with dilution series of anti-*WFS1* and anti-PDI Igs to establish optimum concentrations. Three dilutions of each primary antibody were made in increments from the upper, middle, and lower range recommended by the manufacturers. Anti-*WFS1* Ig was applied to slides at concentrations of 1:200, 1:100, or 1:50 (appendix 27). Anti-PDI Ig was applied to slides at concentrations of 1:500, 1:350, or 1:200 (appendix 28). Optimal concentrations of antibodies were determined with consideration to specificity and signal to noise ratio.

The optimum cell confluency for efficient transfection was 80-90%. Following a 24 hour incubation period in these conditions the high density of tightly packed cells were difficult to distinguish from each other to capture a clear image. A compromise was therefore made by lowering the confluency at transfection to ~60-70%, which reduced the transfection efficiency but enabled visualisation of single cells after incubation and staining.

4.7.6 Detection of endogenous and over-expressed wolframin

A range of transfection experiments were conducted using wild-type *WFS1*-pcDNA. The first aim was to detect cells over-expressing *WFS1*, and distinguish these from untransfected cells, with endogenous levels of *WFS1* expression. This exercise was important groundwork for later experiments using the *WFS1* mutants, to ensure any observed effects could be confidently attributed to the mutations and not complicated by endogenous *WFS1* signals.

We also aimed to make comparisons between untransfected and transfected cells to verify normal localisation of over-expressed wild-type *WFS1*, ruling out any potential effects of 'over-expression' itself.

Following transfection, cells over-expressing *WFS1* were distinctly apparent when viewed under the fluorescent microscope. The green (Alexa-Fluor 488) signal intensity corresponding to wolframin was significantly higher than that produced by endogenous *WFS1* expression alone, as seen in untransfected cells (fig 4.24). To confirm the initial observations we were able to capture images under the (green) filter of endogenous wolframin then reduce the exposure to just below detection levels. Using the same exposure settings we could confirm the 'over-expression' of *WFS1* by detecting a clear signal in transfected cells, thus enabling distinction between endogenous and over-expressed *WFS1*.

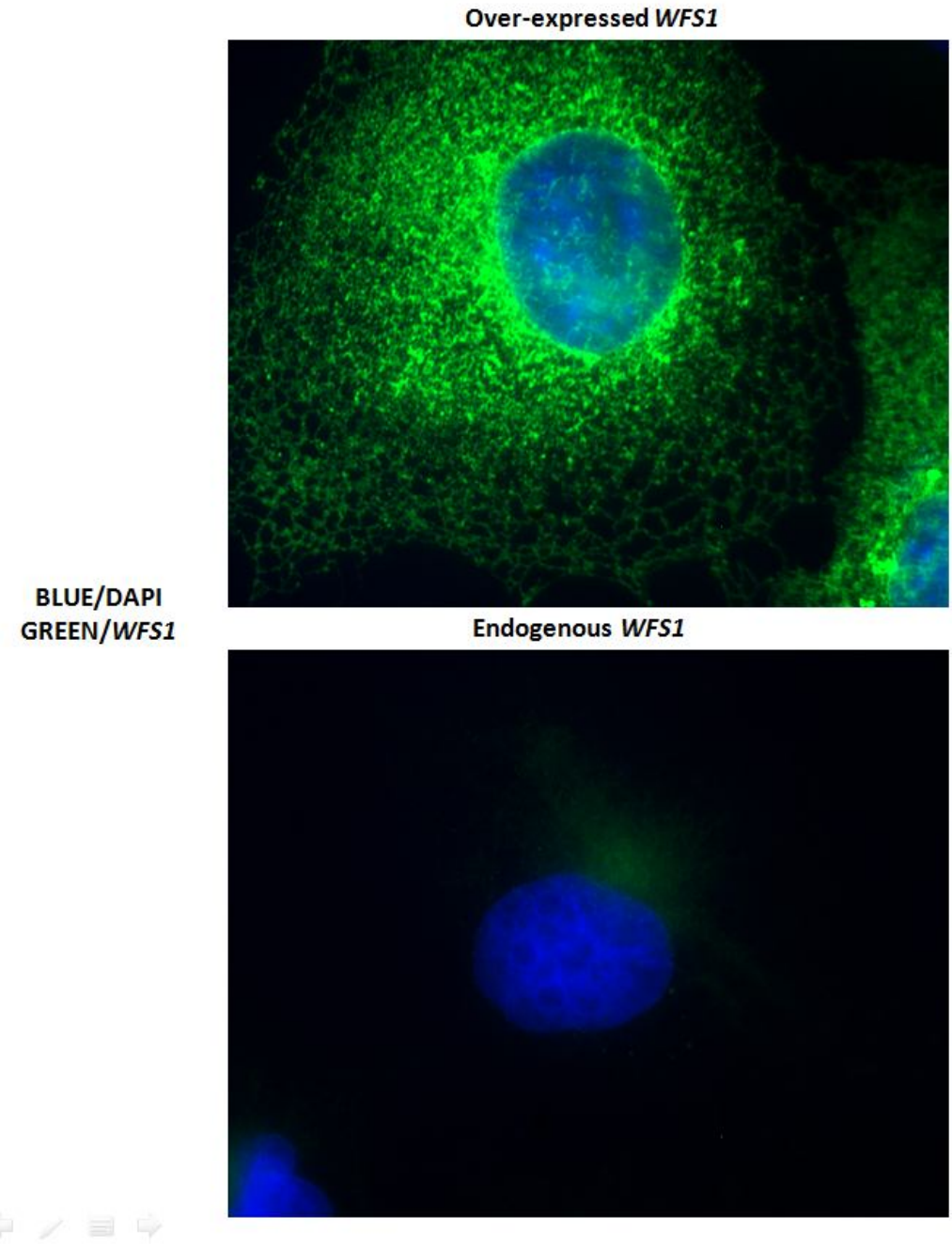


Figure 4.24 COS7 cells stained with anti-WFS1 Ig (1:100). Images captured at same exposure settings demonstrating the clear differences between over-expressing cells and untransfected cells. In all subsequent experiments untransfected cells were identified, then exposure times were reduced until the signal became non-visible. This enabled clear determination of transfected cells, minimising any confounding signals from endogenous wolframin.

The distribution of over-expressed wild-type *WFS1*-pcDNA was consistent with endogenous wolframin, co-localising with PDI in the ER to produce an orange overlapping signal (fig 4.25), validating our wild-type *WFS1*-pcDNA positive control. An obvious reticular distribution, characteristic of the ER was observed, predominantly to one side of the nucleus and fanning out the cell membrane. Obtaining comparable signal intensity of wolframin (green) and PDI (red) at the same exposure was not possible in cells over-expressing *WFS1* as the abundance of wolframin protein produced a disproportionately high signal 'swamping' that of PDI, which was only present at endogenous levels. For this reason exposure of red channel was prolonged to enable clearer representation, justified by our qualitative experimental aims to simply observe a phenomenon.

4.7.7 Negative control experiments

Severely truncated *WFS1* products do not localise correctly to the ER membrane

We have now established a system in COS7 cells where over-expressed wild-type *WFS1* (pcDNA3.1) can be shown to localise correctly. Before introducing *WFS1*-mutants and interpreting our findings, we carried out a series of negative control experiments to demonstrate the possibility of identifying mis-localisation using our method. We predicted that severely truncated wolframin was potentially incapable of localising to the ER membrane correctly. We used a number of human *WFS1* constructs with various deletions (provided courtesy of Dr Malgosia Zatyka) to investigate their suitability as negative controls. The vectors contained: *WFS1* N-terminal domain only (amino acids 1-321); *WFS1* C-terminal domain only (amino acids 652-890), and *WFS1* minus the N-terminal domain (amino acids 322-890). To enable antibody detection, these sequences were cloned into myc-tagged pCMV expression vectors.

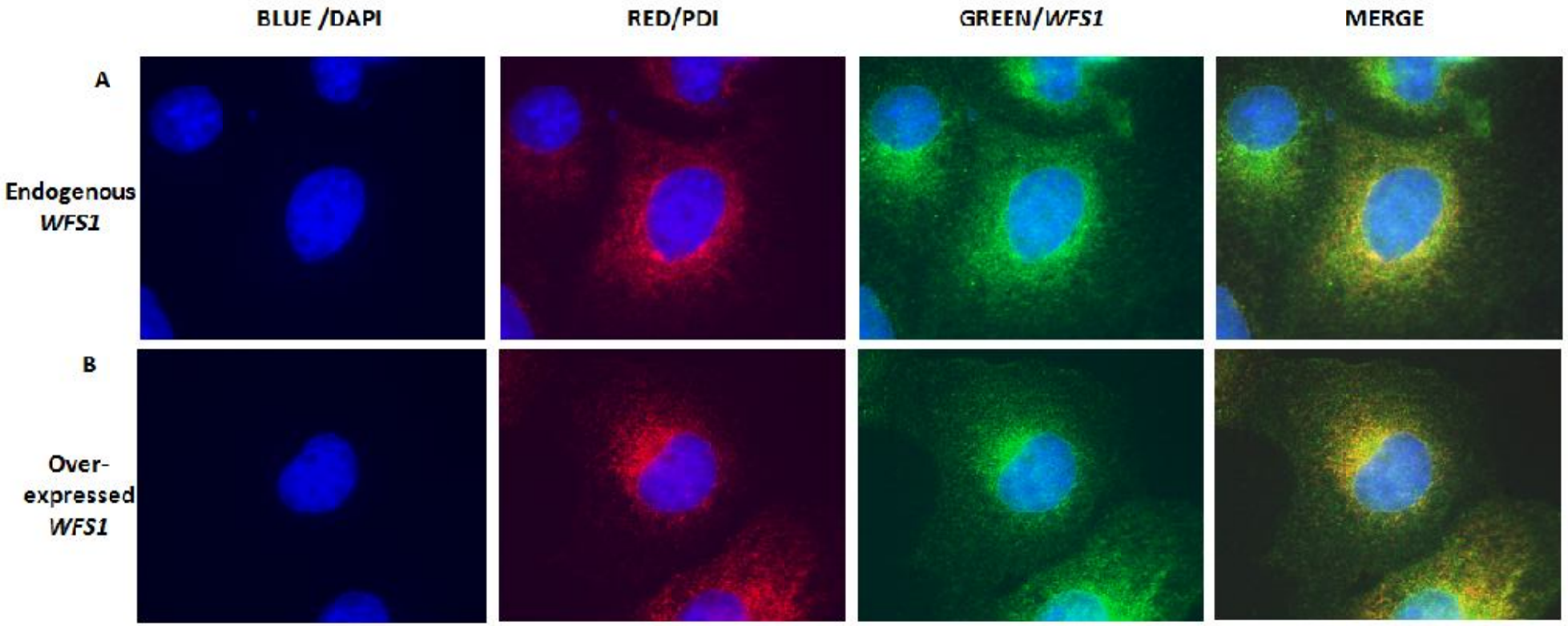


Figure 4.25 Co-staining of COS7 cells with anti-*WFS1* and anti-PDI Igs (x100mag), contrast adjusted. (A) Untransfected cells demonstrating a typical ER pattern of wolframin localisation. **(B)** Over-expressed wild-type *WFS1* produces an identical pattern of distribution.

Full length wild-type *WFS1*-myc was firstly transfected and over-expressed in COS7 cells, and compared to wild-type *WFS1*-pcDNA. We observed consistent ER localisation between both tagged and untagged wolframin (fig 4.26), validating the use of myc-tagged proteins and anti-myc Igs for our control experiments.

The *WFS1* 'deletion' constructs, previously shown to induce over-expression of the correctly sized products by western blotting (Zatyka, 2008a) were then transfected into COS7 cells for immunostaining (section 2.15) with mouse anti-myc primary Ig (table 2.2), followed by anti-mouse secondary Alexa-594 (red; table 2.2). Figures 4.27A-B illustrates the typical patterns of localisation exhibited by each of the negative controls and wild-type *WFS1*-myc. The results confirmed our hypothesis that these severely truncated wolframin proteins did not localise correctly.

Previous data indicate the C-terminal domain of wolframin lies on the luminal side of the ER membrane and the N-terminal domain on the cytosolic side. Based on our knowledge about the structure of wolframin we assume our corresponding truncated mutants do not contain transmembrane domains and are therefore unlikely to anchor to the ER membrane.

Over-expressed *WFS1*-C-term products appear to be concentrated in clusters around the nucleus, perhaps forming aggregates within the ER, the main bulk of which lies in this region. The *WFS1*-N-term products look homogenously spread throughout the body of the cell, possibly indicating dispersal throughout the cytoplasm. The mutant containing the transmembrane domain and C-terminus (*WFS1*-del-N-term) produces a pattern primarily of ER co-localisation, but with additional punctuate wolframin staining suggestive of aggregation. This pattern is similar to the staining observed by Fonseca *et al* (2005) when investigating the missense mutations P724L and G695V.

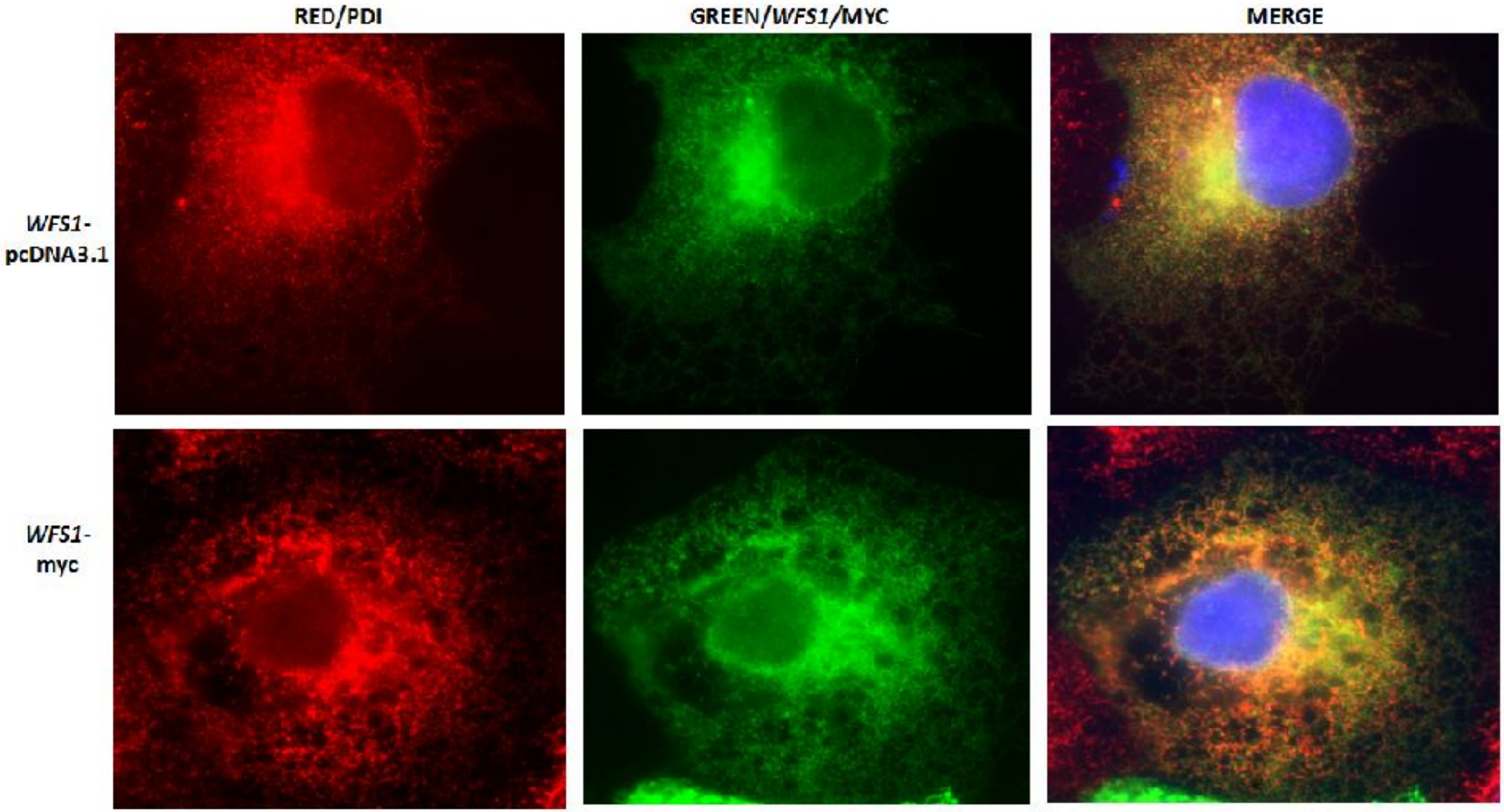


Figure 4.26 COS7 cells over-expressing tagged (myc) or untagged *WFS1* (x100 mag), contrast adjusted. Co-immunostaining for ER with anti-PDI Ig and *WFS1* with anti-*WFS1* or anti-myc Igs. Appropriate localisation of tagged wolframin proteins to the ER validates negative control experiments using tagged mutant proteins.

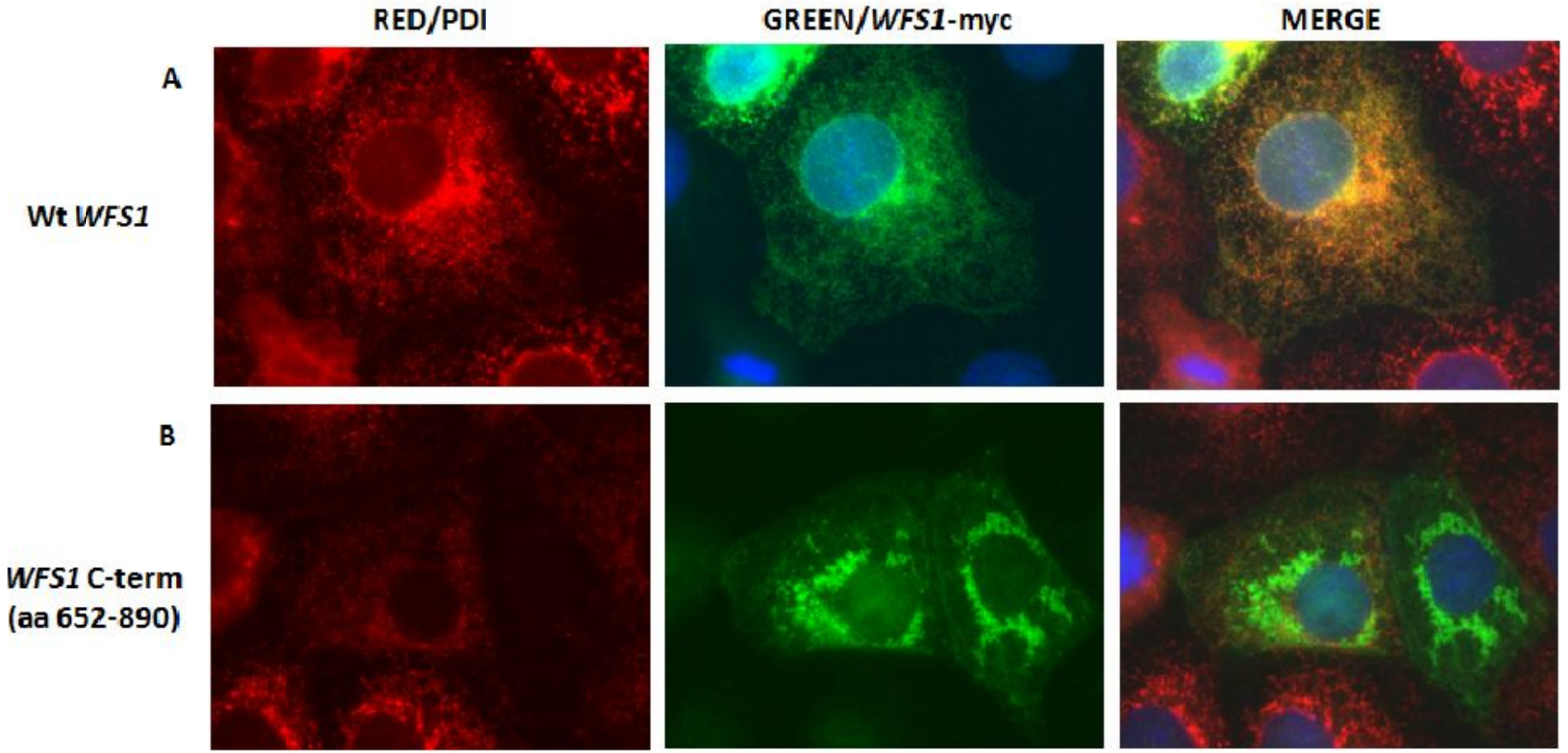


Figure 4.27A COS7 cells fixed and co-stained with anti-*WFS1* and anti-myc Igs (x100mag), contrast adjusted. **(A)** Over-expressed wild-type *WFS1* demonstrating a typical reticular ER pattern of wolframin localisation. **(B)** Over-expressed WFS1-C-terminal mutants appear to cluster around the nucleus, in the region of the main body of the ER, possibly within the ER lumen.

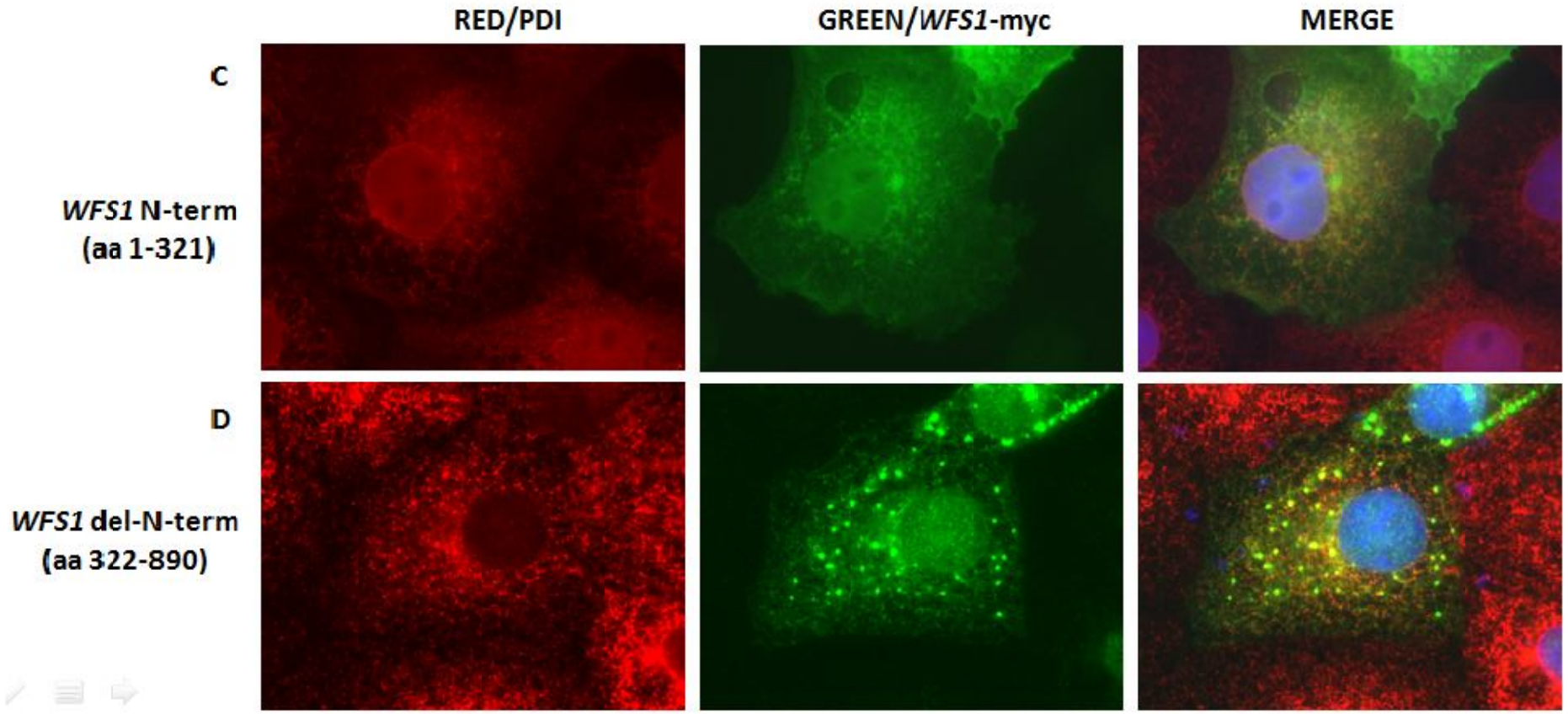


Figure 4.27B COS7 cells fixed and co-stained with anti-*WFS1* and anti-myc Igs (x100mag), contrast adjusted. **(C)** Over-expressed *WFS1*-N-term mutants produce a diffuse, homogenous signal suggestive of localisation throughout the cytosol. **(D)** Over-expressed WFS1-Del-N-term mutants do demonstrate a predominantly reticular pattern of ER localisation, but with punctuate staining characteristic of protein aggregation.

We have speculated about the localisation of our negative controls, although further investigation would be necessary to confirm these suggestions. With respect to the current study aims, our negative control data confirm our system as an effective means of detecting mis-localisation of wolframin.

4.7.8 Localisation of wolframin variants

The *WFS1*-L511P variant appears to cause a degree of protein aggregation *in vitro* whereas other selected wolframin missense variants have the capacity to co-localise correctly to the ER.

Each of the *WFS1*-variants was transfected and over-expressed in COS7 cells for immunostaining. We repeated the transfections three times, using two independent DNA preparations. For each repeat we observed at least two separate slides of fixed and stained cells under the fluorescent microscope for each condition. The slides were scanned by eye systematically and representative images of typical cells were captured.

We observed apparently normal ER localisation of all our missense wolframin variants except for *WFS1*-L511P. The majority of the variants exhibited clear co-localisation with PDI (fig 4.28-31), illustrated by the orange staining (merged images), and obvious reticular 'net-like' pattern, which can be seen more clearly in larger images. The *WFS1*-L511P proteins were noticeably different from the others, and despite co-localising with PDI to some extent, did have the propensity to produce punctuate staining patterns (fig 4.32, appendices 29 and 30). Again, this may suggest the formation of protein aggregates.

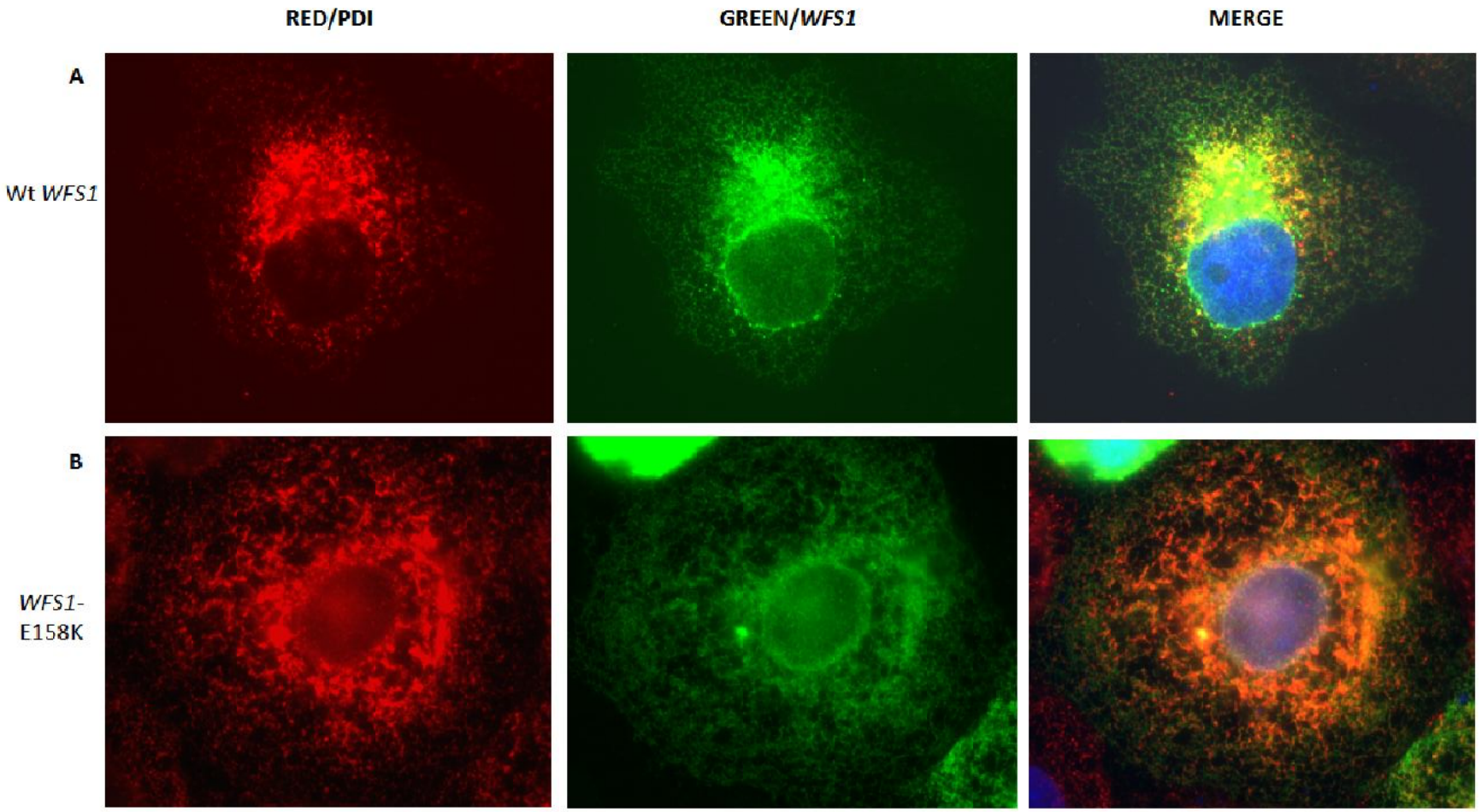


Figure 4.28 COS7 cells fixed and co-stained with anti-*WFS1* and anti-PDI Igs (x100), contrast adjusted. (A) Over-expressed wild-type *WFS1* **(B)** Over-expressed *WFS1*-E158K.

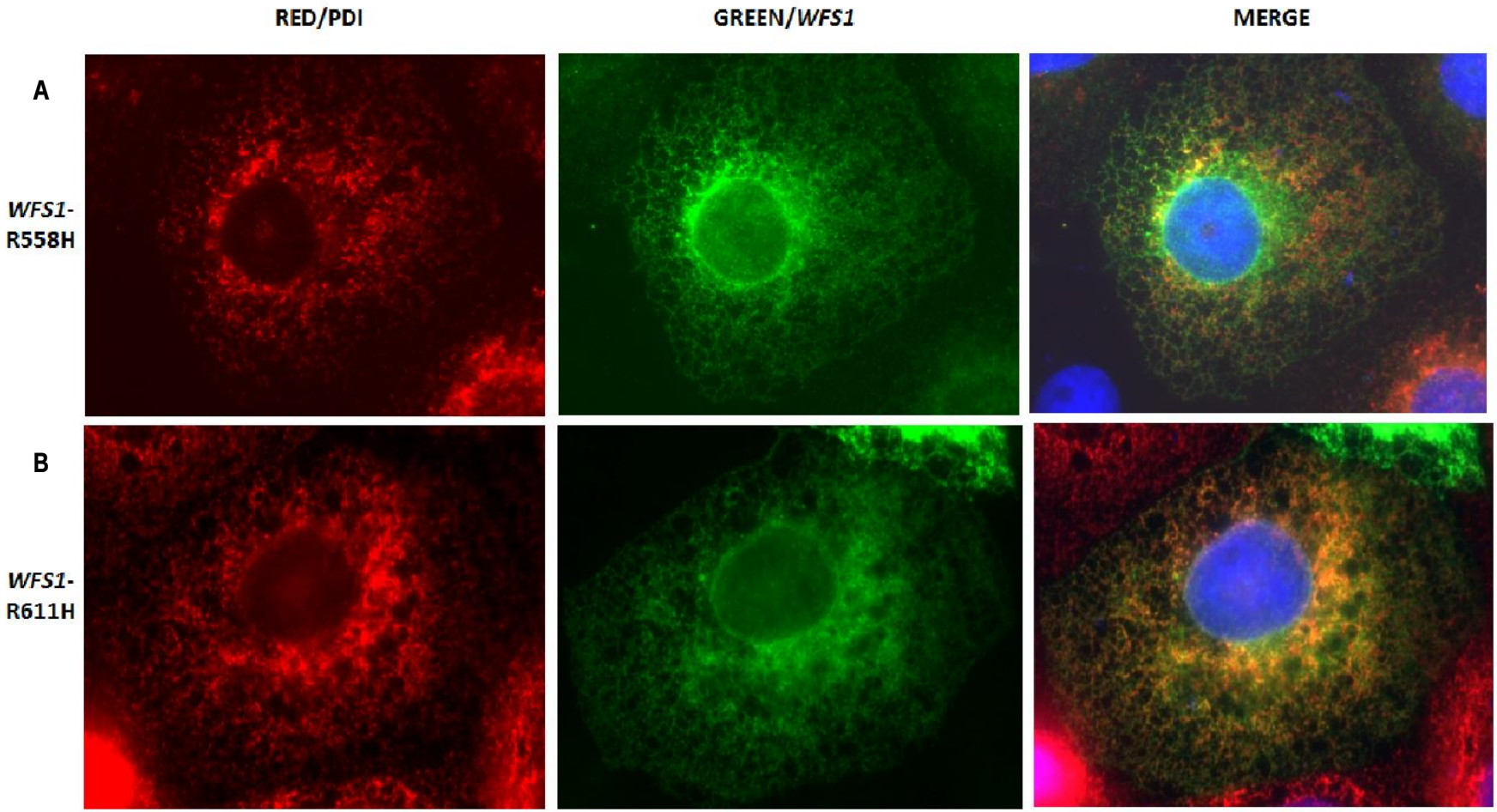


Figure 4.29 COS7 cells fixed and co-stained with anti-*WFS1* and anti-PDI Igs (x100), contrast adjusted. (A) Over-expressed *WFS1*-R558H **(B)** Over-expressed *WFS1*-R611H.

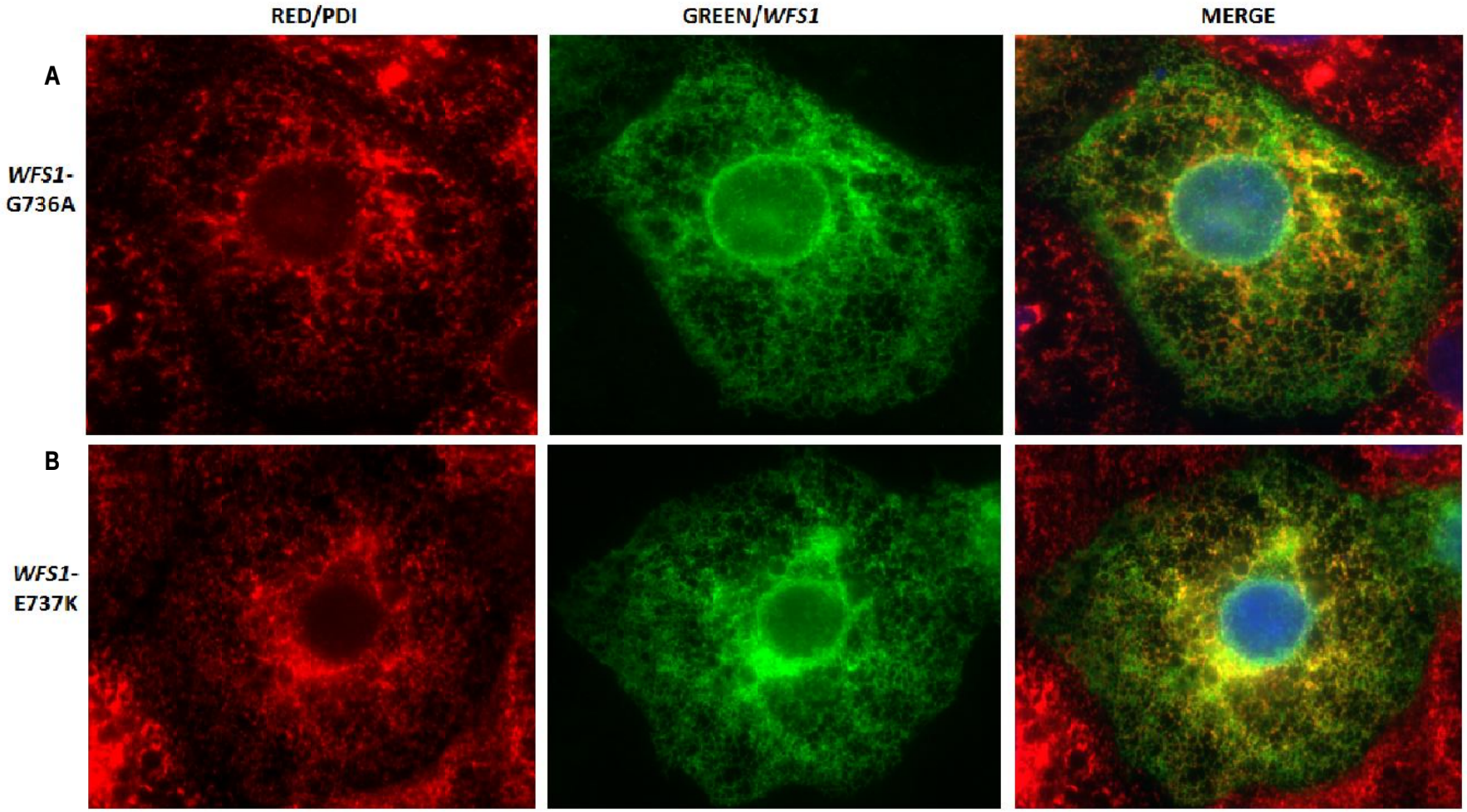


Figure 4.30 COS7 cells fixed and co-stained with anti-*WFS1* and anti-PDI Igs (x100), contrast adjusted. (A) Over-expressed *WFS1*-G736A **(B)** Over-expressed *WFS1*-E737K.

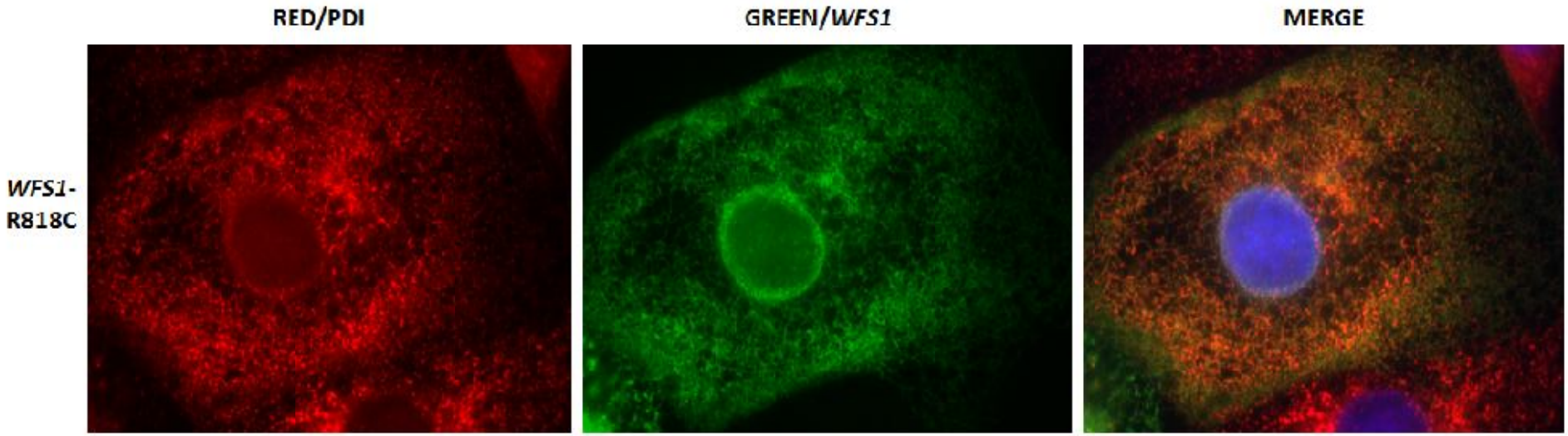


Figure 4.31. COS7 cells fixed and co-stained with anti-*WFS1* and anti-PDI Igs (x100), contrast adjusted. Over-expressed *WFS1*-R818C.

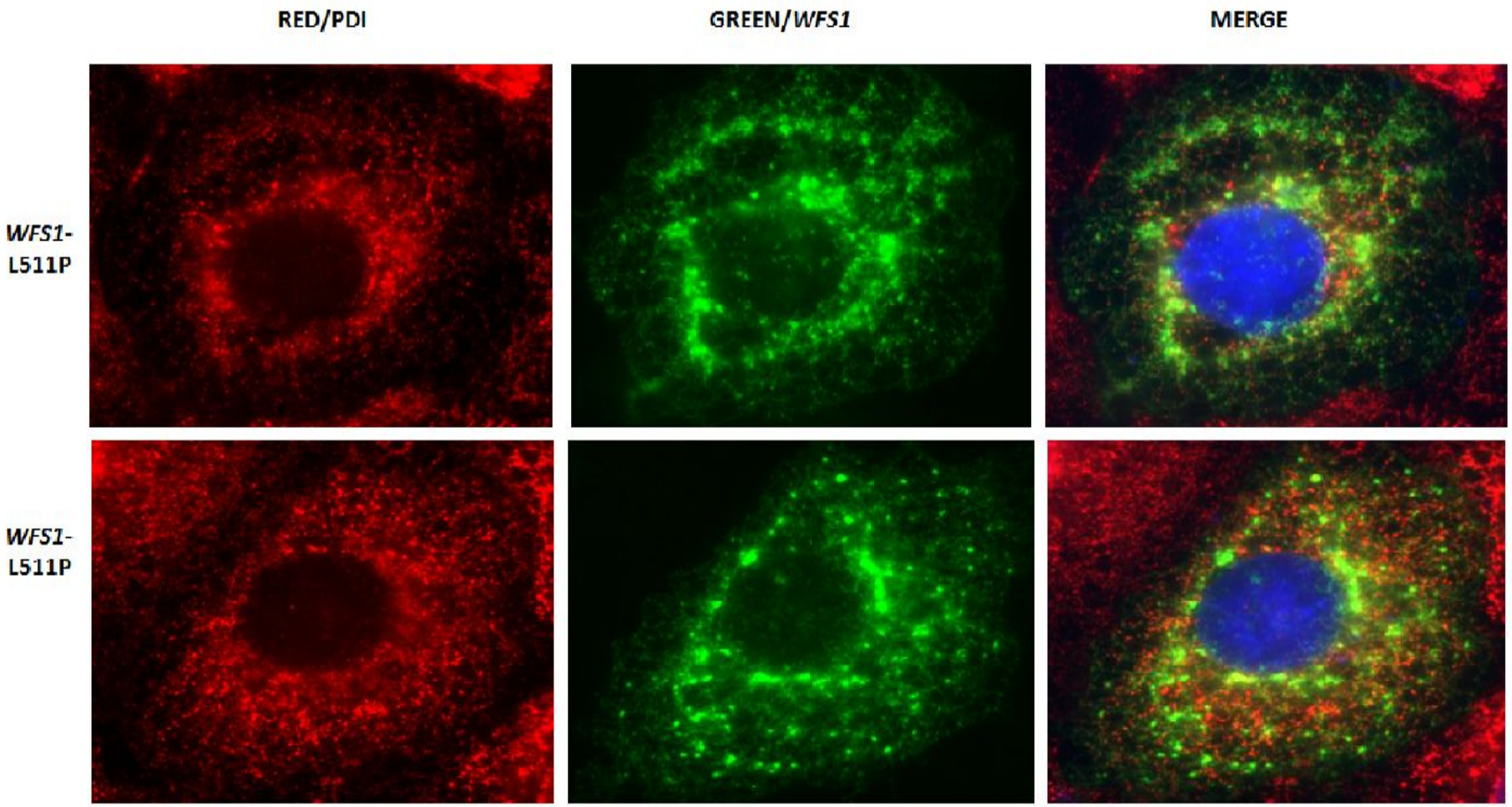


Figure 4.32 COS7 cells fixed and co-stained with anti-*WFS1* and anti-PDI Igs (x100), contrast adjusted. Over-expressed *WFS1*-L511P produces reticular, punctuate staining suggesting localisation to the ER with tendency for aggregation.

4.7.9 Summary and discussion

We have demonstrated *in vitro* that large deletions severely disrupt the normal ER distribution of wolframin as expected. In agreement with an earlier study (Fonseca, 2005) our data also suggest that some *WFS1* missense mutations may alter the sub-cellular localisation of wolframin potentially causing aggregates as illustrated by the distribution of *WFS1*-L511P (see appendix 29 for comparisons). The majority of wolframin variants on our selected panel did however co-localise correctly to the ER, suggesting that loss of wolframin function caused by missense mutation is not necessarily a consequence of protein mis-localisation. In an over-expression system we provide evidence that a number of wolframin disease-associated missense variants have the potential to properly localise to the ER *in vitro*.

Interestingly, although we have observed the tendency for *WFS1*-L511P to form what looks like aggregates, it is not clear from our data whether this may occur due to the solubility of the proteins or to a propensity for these particular variants to form multi-homodimers. No significant differences in solubility were observed when we performed our earlier steady-state experiments (section 4.3.6), although more thorough examinations of relative solubility would be necessary to explore this in more detail. Furthermore, conditions *in vivo* may in fact result in differences not readily detectable in lysed and homogenised cells.

4.8.0 Chapter summary

The overall aim of this investigation was to explore WS pathogenesis caused by *WFS1* missense mutation, attempting to further define the pathomechanisms leading to loss of function. This study was approached with consideration to phenotype and sought firstly to characterise a panel of *WFS1* missense variants, associated with both mild and severe phenotype.

A number of *WFS1* missense mutations were selected, cloned and characterised in terms of protein expression, stability and sub-cellular localisation. From a panel of seven variants, only one mutant (*WFS1*-L511P), representing severe WS, demonstrated the potential to significantly reduce protein expression at the transcription level, and cause mis-localisation. No significant differences were observed between wild-type wolframin and the remaining *WFS1* variants, including those associated with severe WS.

These findings support our hypothesis that loss of wolframin function in WS could be due to several different mechanisms, depending on genotype. In contrast to previous investigations where all WS-associated missense variants studied showed significant 'physiochemical' abnormalities (Hofmann, 2003 and 2006; Yamaguchi, 2004; Fonseca, 2005), we provide the first evidence that some wolframin missense variants have expression and localisation characteristics comparable to wild-type protein. We postulate that some missense variants, particularly those corresponding to a relatively mild WS phenotype, may have residual function and the incentive to explore this idea further is strengthened by our data, since the majority of our selected wolframin variants display normal physiochemical characteristics. Our evidence suggests that the underlying loss of function caused by some missense *WFS1* mutations may arise from impairment of more specific molecular

interactions or functions, which is a possibility we intend to address with more detailed functional analyses.

4.8.1 Functional analysis of wild-type and mutant wolframin

Building on our characterisation data we can now begin to investigate the impact these missense variants have on specific wolframin functions. The common biochemical feature associated with WS is ER-stress mediated apoptosis. Wolframin has been shown to play an important role in a controlled UPR, but is also implicated in a range of other biological systems including calcium homeostasis, insulin secretion, and protein folding and maturation. It is still unclear whether the perturbations in ER-stress caused by *WFS1* mutation are only directly related to the transduction of UPR signalling, or if disruption of other functions such as protein processing could indirectly result in the same outcome.

A top down approach is proposed for future functional analysis of our missense variants, examining firstly the overall effect our selected mutants have on ER stress and apoptosis, before considering a step back to more specific molecular mechanisms including insulin secretion and chaperone activity (see general discussion).

An important role for wolframin in ER stress management is the negative regulation of ATF6α (Fonseca, 2010). This particular interaction, mediated by HRD1, was shown to be impaired in selected variants (*WFS1*-P742L, *WFS1*-G695V), mutations associated with protein aggregation. We intend to investigate the potential for our panel of wolframin mutants to perform this function, given the differences we have discovered during our characterisation experiments. In accordance with previous studies, we anticipate that *WFS1*-

L511P, which shows a propensity to aggregate, may have an impaired interaction with ATF6α. The other variants in our panel, which demonstrate normal 'characteristics,' may have the capacity to carry out this function more efficiently.

In addition to ATF6α, comparing the binding capacity of wild-type and mutant wolframin to other molecular partners could provide more specific information about biochemical differences associated with particular mutations that result in loss of wolframin function. Certainly if all WS-associated mutations prove to have little or no function, we still hope to learn more about the R611H variant and provide the first functional data about how the diabetes 'risk' allele H611, relates to normal and pathological β-cell conditions.

CHAPTER 5

IDENTIFYING WOLFRAMIN MOLECULAR PARTNERS

5.0 Introduction

Complex networks of biochemical interactions are the fundamental components of any living system. Virtually every cellular process including DNA replication, gene expression, protein modification and secretion, signal transduction, cell cycle control and apoptosis is coordinated by very specific protein-protein interactions (PPIs).

The nature of PPIs are diverse, and largely governed by the location and relative concentrations of the interacting participants within the cell at a given time, in combination with the structural and physio-chemical properties of the interacting proteins. Associations between protein partners may be stable, forming core complexes, or transient, as is often seen in regulatory interactions (Ruffner, 2007). These are determined in part by the relative binding strength or specificity of the constituent proteins, and the location of the interacting segments. Many protein complexes are necessary for function, and therefore disruption of these vital associations is an important contributor to the onset of disease.

Much research effort is focused on identifying and understanding PPIs and gathering information about these associations is an important step for understanding disease pathogenesis and developing new therapeutic approaches.

Depending on experimental aims, identifying and understanding PPIs can be approached in a number of ways. Some methods are principally concerned with discovering new interactions, and others may be directed at understanding the dissociation kinetics between known interacting pairs. Well established high throughput screening methods for identifying potential PPIs include surface-anchored protein arrays and two-hybrid assays. Once identified, protein pairs can be examined more closely using affinity purification techniques, or fluorescence resonance energy transfer (FRET) analysis.

5.0.1 The yeast two-hybrid system

GAL4 is a transcriptional regulator that mediates expression of most genes involved in the uptake and metabolism of galactose in yeast (Traven, 2006). It has two distinct functional domains that must be in physical proximity for the factor to function, one that promotes DNA binding and another that activates transcription.

The yeast two-hybrid system exploits the modular properties of transcription factors such as GAL4 for the detection of unknown molecular interactions. Genes regulated by GAL4 for example, contain upstream activating sequences (UAS) in their promoter regions to which GAL4 binds (Bram, 1986), and these UAS are artificially introduced upstream of reporter genes in the yeast two-hybrid system.

Yeast cells can be transformed with selected plasmids conferring expression of a protein of interest fused to a GAL4 transcription factor DNA binding (DBD) domain and a library of cDNAs for unknown proteins which are fused to the complementary GAL4 activation domain (AD) (fig 5.0A). The known protein is used as ‘bait’ to screen for potential partners

from the library 'prey' proteins. When the partners associate and bring the two transcriptional domains together, expression of a target gene is triggered which acts as a reporter (fig 5.0B). A functional GAL4 transcription factor for example, activates expression of the *MEL1/LacZ* gene (α/β-galactosidase), which enables transformed yeast/bacteria to metabolise colourless X-Gal and produce a blue insoluble product (fig 5.0C).

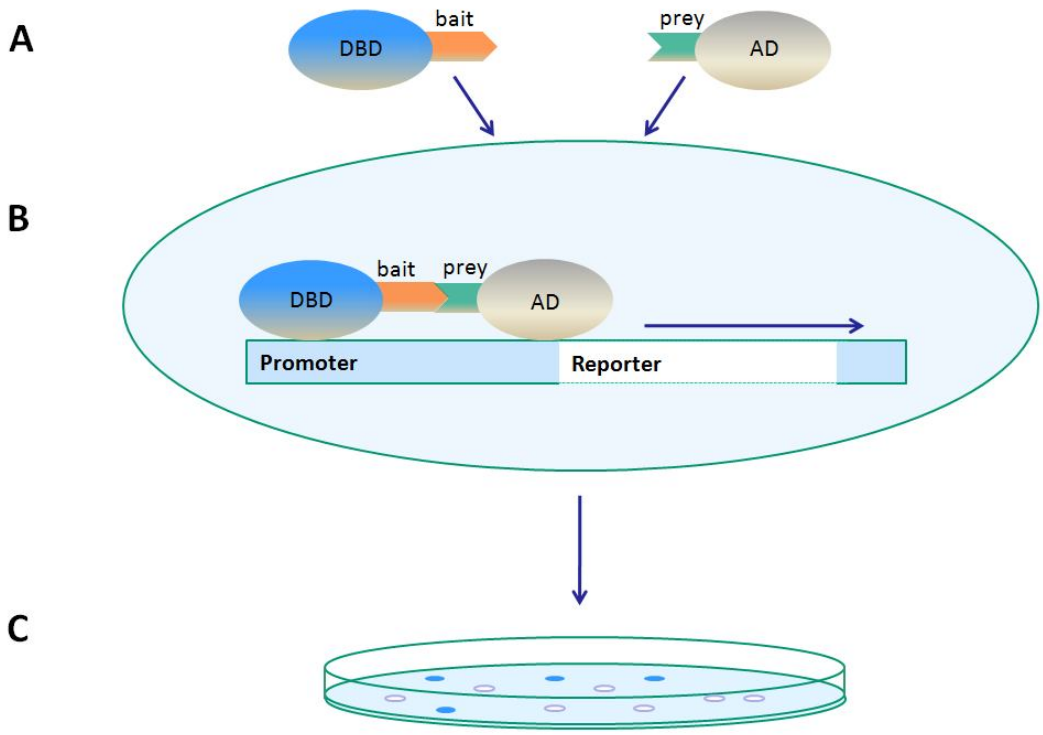


Figure 5.0. **Principles of the yeast two-hybrid assay**. (**A**) Yeast cells are co-transformed with vectors containing a known gene of interest fused to a transcription factor DNA binding domain (DBD), and a library of potential molecular partners fused to the same transcription factor activation domain (AD). (**B**) When binding partners associate, the transcription factor is activated by close proximity of the functional domains, triggering expression of a reporter gene (e.g. *MEL1*). (**C**) Reporters such as *MEL1* allow 'blue-white' selection.

Due to significant amplification of reporter gene expression, the yeast two-hybrid assay is useful for detecting transient or weak interactions, although this sensitivity does result in a higher number of 'false-positive' results, which need to be examined further and discounted. For investigation of mammalian protein interactions, yeast also have the advantage of providing a host background environment more closely resembling that of a higher eukaryote compared with bacterial or *in vitro* assays. This method is however limited when protein-protein interactions require post-translational modification such as glycosylation or phosphorylation, which do not occur in yeast systems.

Overall a relatively large number of potential protein-partners can be screened *in vivo* using this technique to isolate a number of novel candidates for further investigation. Questions arising about the biological relevance of any protein associations detected in yeast can then be addressed by more detailed analysis in mammalian systems.

5.0.2 Molecular partners of wolframin

A number of wolframin molecular partners have been identified in previous studies, helping to improve our understanding of wolframin function.

Calmodulin (CaM) was determined as a wolframin partner-protein by affinity chromatography purification (Yurimoto, 2009) and is an important modulator in a large number of Ca^{2+} dependent signal transduction pathways. Intracellular Ca^{2+} fluctuations affect a wide range of processes including protein secretion, gene expression and cell division all of which share links with proposed wolframin function.

In the Birmingham laboratory, the wolframin (WF) -C-terminal domain was used as bait in a yeast two-hybrid screen with a human brain cDNA library (Zatyka, 2008a). The β1 subunit of Na^+/K^+ ATPase (*ATP1B1*) was isolated, and confirmed as a molecular partner of wolframin by endogenous co-immunoprecipitation. This enzyme, also known as the Na^+/K^+ pump, is situated on the plasma membrane and assists in membrane transport and maintenance of cell volume as well as functioning as a signal transducer, regulator of mitochondrial reactive oxygen species production and intracellular calcium homeostasis; all of these pathways have connections with observed sub-cellular features in Wolfram syndrome models.

GRP94, an ER molecular chaperone from the heat shock protein 90 (HSP90) family, has also been shown to associate with wolframin (Kakiuchi, 2009). Interesting links can be made between this interaction and clinical manifestations of Wolfram syndrome; *GRP94* is strongly linked to bipolar disorder in the Japanese population (Kakiuchi, 2007) and the complex between wolframin and *GRP94* is modulated in neurones by valproate, a standard drug treatment for mood stabilisation.

The identification of other wolframin-partners including *ATF6α*, *Hrd1* (Fonseca, 2010) and *Smurf1* (Guo, 2011) define more clearly the role wolframin plays in UPR signalling. A functional complex between wolframin and the E3 ligase *Hrd1*, is thought to target *ATF6α* for proteasomal degradation, so having a negative regulatory effect on the ER stress response. The E3 ligase *Smurf1* has been shown to interact with, and target wolframin for ubiquitination and degradation, which demonstrates at least one mechanism of *WFS1* regulation. Selected WF-mutants expressed ectopically also seem to exhibit resistance to

the action of *Smurf1* and could potentially result in pathological inhibition of *ATF6α* and disruption of the UPR.

5.1 Specific aims and methods

The specific aim of this section was to isolate novel wolframin molecular partners to extend current knowledge about the functional pathways of wolframin, in particular interactions taking place in the pancreas, which have special relevance to *WFS1* and its involvement in diabetogenesis. Identifying new molecular partners would additionally increase the scope and novelty of our proposed future functional work with WF-variants (chapter 4).

A yeast two-hybrid assay performed by Dr Christopher Ricketts, Dept. Medical and Molecular Genetics, University of Birmingham, formed the foundation for this work. A Human Pancreas Matchmaker cDNA library (Clontech) was screened for potential interactions with the WF-N-terminal domain (amino acids 1-321) using the MATCHMAKER GAL4 Two-hybrid System 3 (Clontech). *WFS1*-N-terminus, cloned into pGBKT7 vectors (*EcoR1*/Xho1) was transformed into yeast (AH109) and checked for protein expression. Positive yeast clones (containing pGBKT7-*WFS1*-N) were then co-transformed with mixed pACT2 plasmids containing a pancreatic cDNA library and grown on selective media ((QDO + X-α-Gal) section 2.8.2) to isolate potential interactors (fig 5.2).

As described in section 5.0.1, the association of the bait and prey proteins can be detected by expression of *MEL1* and X-α-Gal metabolism (blue/white screening). Selection stringency for interaction is increased further still in the AH109 yeast strain, minimising false-positive results by growing co-transformed cells on media lacking adenine and histidine, known as

quadruple drop out (*-TRP/-LEU/-ADE/-HIS*). The nutritional reporters *ADE2* and *HIS3*, are, like *MEL1*, under the control of distinct GAL4 upstream activating sequences, and only transcribed if bait and prey proteins are in close proximity.

Each positive clone isolated and identified from the two-hybrid screen resulted in growth of blue AH109 colonies. Colonies were evaluated according to strength of blue product (X-α-Gal metabolites), which was hypothesised to reflect strength of interaction. Putative positive clones were selected, numbered and stored (-80°C).

This work was continued firstly by extraction of plasmids and identification of 'prey' proteins from positive yeast clones. Plasmids were purified from yeast using the Yeastmaker Yeast Plasmid Isolation Kit (Clontech; section 2.8.5), and T7/3' DNA-AD vector primers (appendix 2.7) used to PCR amplify (section 2.1.0) across the region containing the 'unknown' library gene. PCR products were electrophoresed (section 2.3.0), purified from agarose gel (section 2.3.1) and sequenced (section 2.4.0). By applying results to a BLAST search tool (NCBI), the identity of captured 'prey' proteins from the library was determined.

Following sequence analysis, positive clones were verified by 'one-to-one' co-transformation of yeast (section 2.8.4) with both bait (WF-N-terminus) and known prey plasmids to confirm expression of *MEL1*, thus signifying potential protein interaction. Selected candidates were then investigated further in mammalian cells by over-expression and co-immunoprecipitation (fig 5.1).

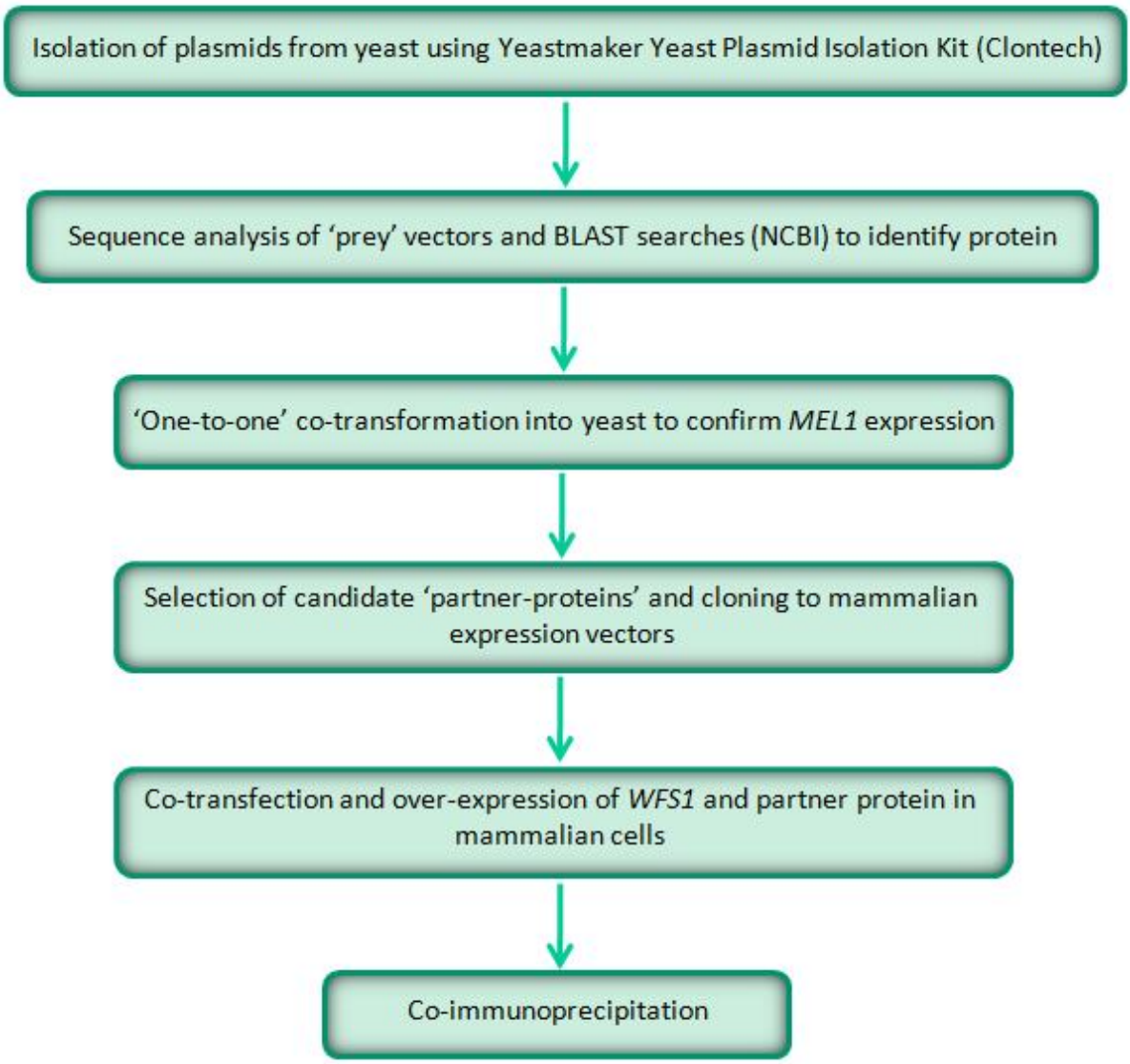


Figure 5.1. **Methods workflow for identifying and verifying protein-protein interactions following a yeast two-hybrid screen.** Plasmids containing unknown library gene sequences were extracted from 'positive' yeast clones and sequenced. Genes of interest, coding for potential wolframin partner proteins, were identified using the BLAST sequence alignment search tool (NCBI). Plasmids containing positively identified 'partners' were then reintroduced to yeast with *WFS1* to confirm initial findings. Finally, each of the candidate gene sequences and *WFS1* were cloned into mammalian expression vectors for co-transfection and immunoprecipitation experiments in human cell lines.

Results

5.2 Isolation and identification of DNA from positive yeast two-hybrid clones

Following sequencing analysis, a number of putative wolframin partner-proteins were identified from the pancreatic cDNA library screen: protease serine 2, pancreatic lipase and cofactor colipase, regenerating islet-derived 1 alpha (REG) protein, alpha amylases 2A/2B, and carboxypeptidase B1 (table 5.0).

These secretory proteins are components of the pancreatic juice, produced and exported from the exocrine pancreas and with the exception of the REG protein, most are thought to be principally involved in digestion.

Protease serine 2

Protease serine 2 (*PRSS2*) or trypsinogen, is a 247 amino acid zymogen secreted by acinar cells (Emi, 1986). In the small intestine cleavage by enteropeptidase activates the enzyme (trypsin, 223 amino acids), which stimulates autocatalysis. Trypsin functions to hydrolyse peptide bonds, reducing proteins to their constituent amino acids for efficient absorption by the body.

Pancreatic lipase

Pancreatic lipase (*PNLIP*) is a carboxyesterase that catalyses hydrolysis of intraduodenal long chain fatty acids and glycerol esters, which are then able to enter the circulation through the lymphatic system (Lowe, 1989). Secreted in its mature form together with a signal

sequence (465 amino acids), pancreatic lipase only becomes active when associated with colipase (*CLPS*) in the small intestine.

PNLIP belongs to a lipase protein superfamily related by tertiary structure (Wong, 1992). Other members of the lipase family can be found in a variety of tissue types as well as the pancreas, functioning amongst other things in the metabolism of circulating lipoproteins and tissue supply of fatty acids (Olivecrona, 1993).

Colipase

Colipase (*CLPS*) is secreted in coordination with pancreatic lipase and is a vital cofactor for lipase activity. Proteolytic cleavage of procolipase in the gut lumen activates this co-enzyme, enabling it to bind the C-terminal domain of lipase. This interaction prevents the inhibitory effects of bile salts on lipase, thus initiating lipase activity. The immature colipase prepropeptide is comprised of 112 amino acids (Lowe, 1990), 1-17 forming the signal sequence, and 18-22 the activation sequence, otherwise known as enterostatin. In addition to the pancreas, the procolipase gene is also expressed in the stomach, duodenal mucosa (Okada, 1993), and regions of the brain (Lin, 2002).

Enterostatin appears in the circulatory system and lymph following food intake and has been shown to relay satiety signals to the brain (Erlanson-Albertsson, 1997). The metabolic effects of enterostatin are associated with reduction in fat intake, maintenance of bodyweight (Silvestre, 1996), and decreased insulin secretion (Park, 2009).

Clone	Gene	Protein function
Protease, serine 2 (trypsin 2)	5645 *PRSS2*	protease – hydrolyses proteins into smaller peptides/amino acids
Pancreatic colipase	1208 *CLPS*	lipase cofactor – assisting digestion of dietary fats
Pancreatic lipase	5406 *PNLIP*	digestion of dietary fats
Regenerating islet-derived 1 alpha (pancreatic stone protein)	5967 *REG1A*	associated with islet cell regeneration and diabetogenesis may be involved in pancreatic lithogenesis
Pancreatic amylase alpha 2A/2B	279 *AMY2A* 280 *AMY2B*	requires Ca^{2+} to function breaks down long-chain carbohydrates acts at random locations along the starch chain
Carboxypeptidase B1	1360 *CPB1*	hydrolyzes the carboxy-terminal (C-terminal) end of a peptide bond. digestion, protein maturation, regulation

Table 5.0. Clones identified from initial yeast-two hybrid screen using wolframin N-terminus as bait in a pancreatic cDNA library.

Regenerating protein

The REG protein (*REG1A*) also referred to as; the pancreatic stone protein, lithostathine, thread-protein and protein X, is a non-enzymatic secretory glycoprotein (Terazono, 1988). The precursor protein of 166 amino acids consists of a 22 amino acid signal peptide, susceptible to proteolytic cleavage. The mature C-terminal product (133 amino acids) displays a tendency to precipitate at physiological pH and forms the predominant component of pancreatic stones and fibrils (De Caro, 1979).

REG1A belongs to the *REG* gene family, which includes *REG1B*, *REGL* and *PAP*. These genes are found tandemly clustered on chromosome 2p12 (Miyashita, 1995), possibly arising by duplication of a common ancestral gene. Upregulated expression of these genes occurs in response to damage, inflammation and stress in a variety of tissues (Ashcroft, 2004). Some members of the *REG* family have been associated with motor neuron regeneration (Livesey, 1997) and Alzheimer's disease (Ozturk, 1989).

More specifically, *REG1A* has been shown to act as a growth factor, playing a role in the inflammatory response in gastroenterological organs (Dieckgraefe, 2000; Lawrance, 2001), and is expressed in regenerating islet β-cells (Terazono, 1988; Watanabe, 1994).

Amylase 2A/2B

Pancreatic alpha amylases (*AMY2A/AMY2B*) begin the process of polysaccharide digestion by acting on the α-1-4 glycosidic bonds in dietary starch and sugars. The full protein is 511 amino acids including the signal sequence and requires one essential calcium ion to form

the correct tertiary structure (Steer, 1974) and chloride ions to stimulate activity (Levitzki, 1973).

Carboxypeptidase B

Carboxypeptidases are a widely distributed family of hydrolytic enzymes that remove C-terminal residues from proteins. Performing a variety of functions throughout the body carboxypeptidases are associated with protein maturation and secretion (Fricker, 1988), digestion (Beck, 1973), coagulation and inflammatory pathways (Song, 2011), and development (Reznik, 2001). A particularly interesting member of this family with respect to Wolfram syndrome pathology is carboxypeptidase E (*CPE*), which shares 15-20% sequence homology with *CPB1* (Skidgel, 1988). This protein has been shown to provide an important link between diet-induced hyperlipidemia, ER stress-mediated islet β-cell apoptosis and type 2 diabetes (Kharroubi, 2004; Jeffrey, 2008).

The pancreatic exocrine carboxypeptidases (A1, A2 and B) are protease enzymes that cleave amino acids from proteins during digestion. Like other digestive enzymes carboxypeptidase B1 (*CPB1*) is also secreted as an inactive propeptide to protect the pancreatic tissue from damage (417 amino acids), and activated in the gut lumen by trypsin to a mature, active protein of 307 residues.

Summary of initial screen

Unexpectedly we did not isolate any potential WF-interacting partners typically expressed in β-cells, which would be the most relevant to WS pathology, although the vast majority (~90%) of pancreatic tissue is exocrine. The proteins identified from a yeast two-hybrid

screen are predominantly associated with digestion but some of these also have additional functions in the brain and links with insulin secretion and β-cell apoptosis, which relates them more specifically with systems primarily affected by WS. There are no gastrointestinal symptoms reported in *WFS1*-associated WS but cells producing digestive enzymes are overburdened with protein synthesis and susceptible to ER stress. The role of wolframin in the UPR may suggest a function for this protein in the exocrine pancreas and perhaps the absence of disease manifestation in these cell-types is due to pathway redundancy?

5.3 Confirming protein interactions in yeast

Positive interactions were re-tested in yeast cells. Electrocompetent *E.coli* were used to produce large volumes of purified bait and prey clones in preparation for re-transformation. Yeast-extracted plasmids were electroporated individually into bacterial cells and cultured (section 2.7.4). Plasmids were harvested using maxiprep kits (QIAgen; section 2.7.5) and sequence verified (section 2.4.0). Bait vectors (pGBKT7-*WFS1*-N) together with pACT2 containing a *known* partner protein sequence (from table 5.0) were then co-transformed (section 2.8.4) on a 'one-to-one' basis into competent yeast cells and grown on selective medium (fig 5.2).

According to the same basic principles as the initial screen in yeast (section 5.0.1), selective markers were used for verifying the presence of each plasmid and the interaction of bait and prey. Blue colonies produced from co-transformations plated on a QDO-X-α-Gal substrate signified an active GAL4 transcription factor indicative of bait-prey interaction.

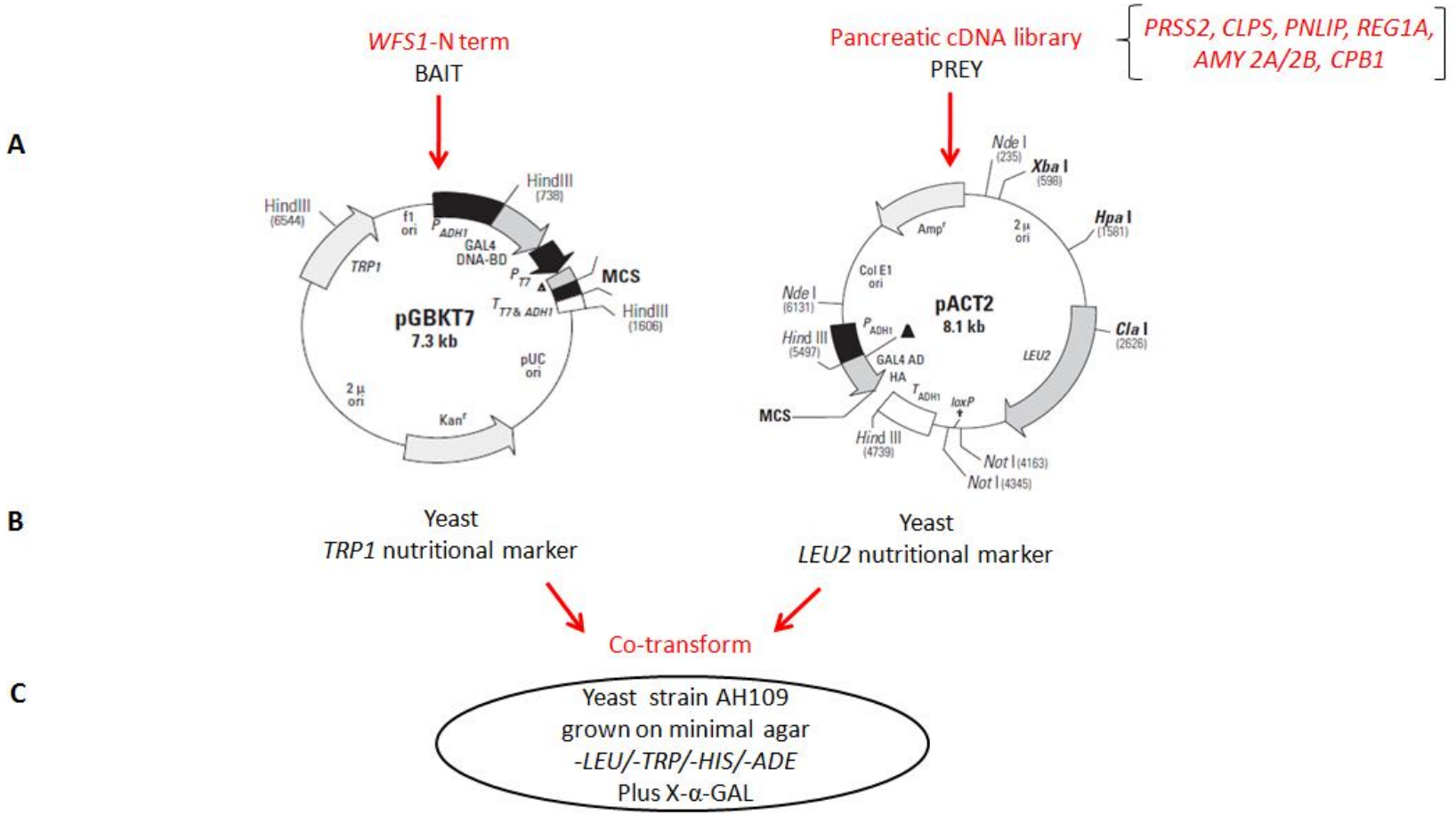


Figure. 5.2. **MATCHMAKER GAL4 Two-Hybrid System 3.** (**A**) Fusion vectors containing *WFS1*-N term as bait in pGBKT7 and a pancreatic cDNA library in pACT2 are prepared. (**B**) Each vector contains unique selection markers. (**C**) Following co-transformation, yeast are able to grow on minimal media; the presence of *WFS1* is associated with growth on *–TRP1,* the presence of a library protein is associated with growth on *–LEU2*. An interaction between the two proteins produces a functional GAL4 transcription factor, which induces expression of *ADE2* and *HIS2* for more stringent nutritional selection in addition to *MEL1* expression enabling blue/white screening on X-α-Gal substrate.

Control samples of cells transformed with a single bait or prey vector did not result in colony formation on QDO-X-α-Gal agar (data not shown). However, each positive clone (pACT2-prey) identified from the two-hybrid screen (table 5.0) did result in growth of blue yeast colonies on QDO-X-α-Gal when independently co-transformed with *WFS1*-N term/pGBKT7 (fig. 5.3). These results confirm our initial findings and justify further investigation of these candidate WF-partners.

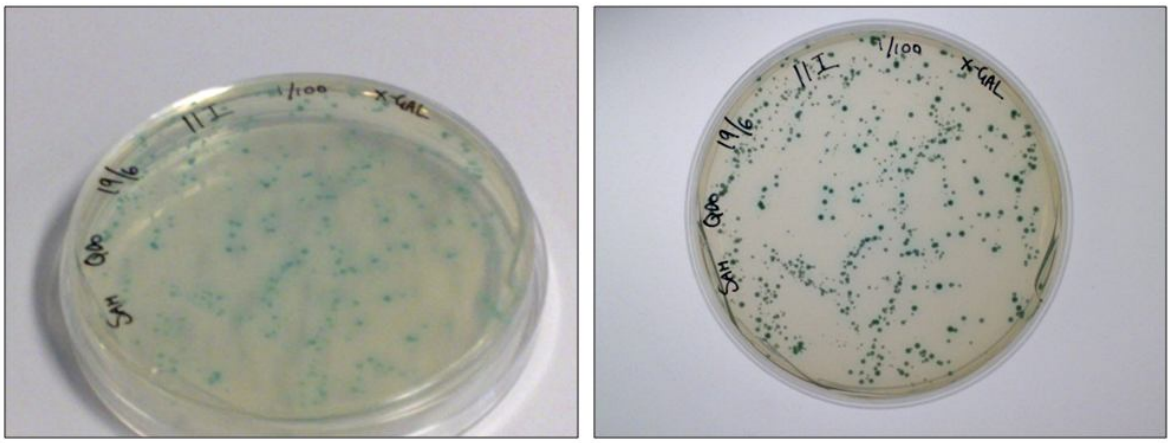


Figure 5.3. **Re-testing positive interactions in yeast.** Co-transformation of competent AH109 yeast with *WFS1*-N term/pGBKT7 and each of: *PRSS1*, *CLPS*, *PNLIP*, *REG1A*, *AMY 2A*/*2B* and *CPB1* in pACT2 produced blue colonies when grown on quadruple drop-out yeast minimal medium with X-α-gal, confirming positive selection from an initial library yeast two-hybrid assay.

5.4.0 Expression of *WFS1* in exocrine tissue

Consideration must be given to the biological significance of interactions detected between foreign fusion-proteins introduced into yeast. False-positive results may occur as a consequence of incorrect protein folding and altered stability, and other molecular associations may be dependent on formation of larger complexes involving multiple proteins. Whilst these possibilities can be largely investigated in mammalian cells, it is important to firstly ensure the protein partners in question are likely to co-exist both spatially and temporally in an endogenous environment.

Although interactions with wolframin were initially detected from a pancreatic library, it is feasible that in reality they may occur in one or several other locations, perhaps other sites specific to the activities of our candidate proteins such as the intestine, brain (*CLPS*) or islet cells (*REG1A*). To justify further analysis of our proposed protein-couples we decided to examine *WFS1* expression in exocrine acinar cells as a starting point since each of our selected 'partner' proteins, in the pancreas, are predominantly expressed and secreted from this cell type.

WFS1 expression in acinar cells has not previously been demonstrated

Earlier studies in fact suggest that *WFS1* is not expressed in exocrine tissue (Ishihara, 2004; Philbrook, 2005; Xu, 2009). The absence of *WFS1* however remains somewhat inconclusive since none of the aforementioned studies were primarily focused on exocrine tissue and did not seek to confirm their observations with suitable controls. We suggest it may be possible

that wolframin could be present, but in much lower concentrations than in islets, where *WFS1* is known to be strongly expressed (Philbrook, 2004; Ueda, 2005).

Two complementary approaches were taken to determine *WFS1* expression in exocrine tissue: examination of *WFS1* mRNA and protein levels in a pancreatic exocrine cell line, and immunohistochemical analysis of whole rat pancreatic tissue sections

5.4.1 Analysis of *WFS1* mRNA and protein in rat acinar cells

A rat pancreatic AR42J acinar cell line was purchased from the European Collection of Cell Cultures (HPA, 2011) and cultured according to supplier recommendations (table 2.0, chapter 2).

The RNeasy kit (Promega) was used to extract RNA from samples of $\leq 1 \times 10^7$ cells (section 2.10.2). Sample concentrations were measured on a Nanodrop 2000 spectrophotometer (Thermo Scientific; section 2.12.0) and 1μg of DNase treated RNA subsequently reverse transcribed with random primers and Reverse Transcription System reagents from Promega (section 2.1.2). Successful cDNA synthesis was verified on agarose gel (section 2.3.0), before standard PCR amplification with selected primers (section 2.1.0).

Primers spanning exons 6 and 7 of rat *WFS1* were designed to produce 241bp products from mRNA (cDNA) if present. A positive control house-keeping gene, rat-*ACTB,* was also analysed with primers spanning a 380bp region of cDNA (see appendix 2.13 for primer sequences) plus negative controls containing no reverse transcriptase. The results (fig 5.4), analysed by

agarose gel electrophoresis (section 2.3.0), confirmed the presence of *WFS1* expression in Acinar AR42J cells.

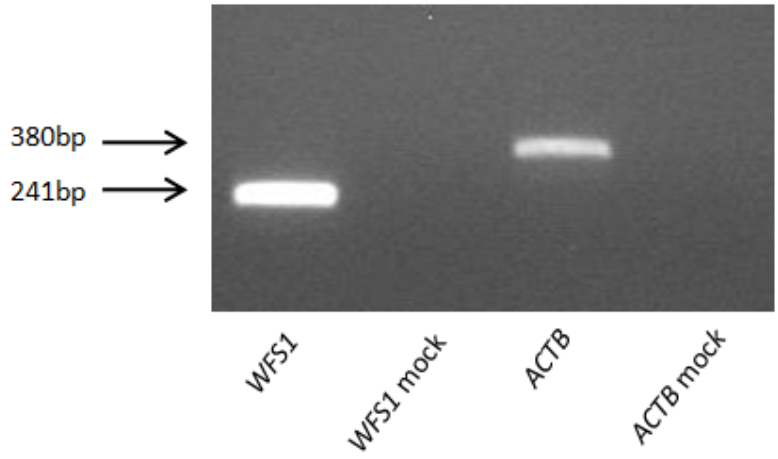


Figure 5.4. **Expression of *WFS1* RNA in exocrine cells**. Reverse-transcription PCR products derived from rat Acinar AR42J RNA demonstrating expression of *WFS1* and positive control gene *ACTB*. Clear lanes for negative control samples, *WFS1*-mock and *ACTB*-mock, cycled without reverse-transcriptase, indicate the absence of genomic DNA or other amplicon contamination.

Protein was harvested from AR42J Acinar cells in RIPA lysis buffer (section 2.10.1) for SDS-PAGE and immunoblotting (section 2.13.0). Positive control whole cell lysates were taken from both: MIN6 cells, a mouse pancreatic β-cell line known to express *WFS1*, and transfected COS7 cells over-expressing *WFS1* (section 2.9.5).

Qualitative analysis using anti-wolframin antibody (Hofmann, 2003; table 2.2) revealed that wolframin protein is present in rat exocrine AR42J cells (fig 5.5). Re-probing blots with anti-β-actin (*ACTB*) antibody (table 2.2), a ubiquitous house-keeping gene, provided a crude indication of total protein loaded.

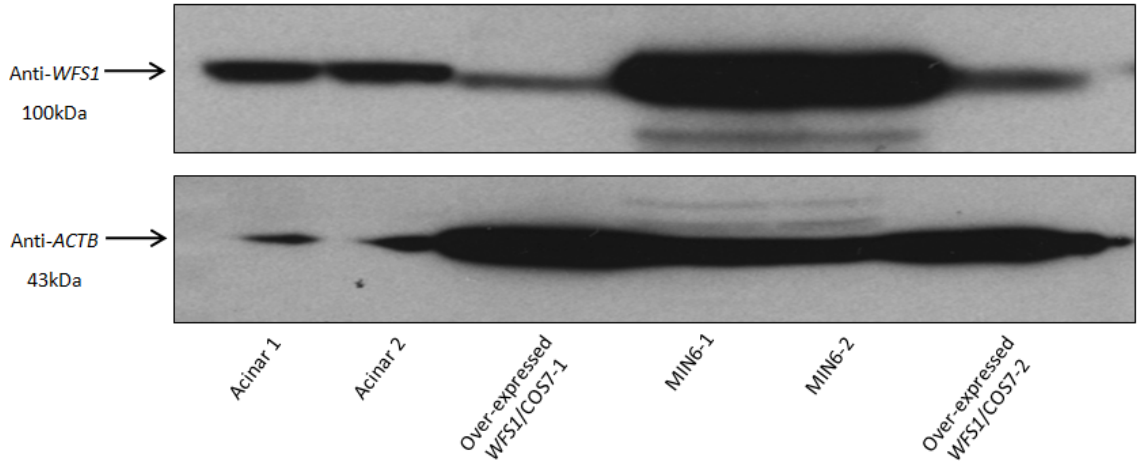


Figure 5.5. Qualitative analysis of *WFS1* expression in exocrine cells. TOP PANEL: Western blot demonstrating detection of wolframin in rat Acinar AR42J cells alongside positive control samples derived from: COS-7 cells over-expressing *WFS1* and untransfected MIN6. BOTTOM PANEL: Signals from blots re-probed for β-actin (*ACTB*) provide an indication of total protein loaded for crude comparison.

Due to the potential differences in antibody specificity to proteins derived from different species a direct comparison of *WFS1* expression levels between cell lines cannot be made. These differences perhaps explain the relatively weak *WFS1* signal in COS7 cells 'over-expressing' *WFS1*. A rat pancreatic β-cell line, such as *INS-1* may provide a useful comparison for semi-quantitative analysis of WF protein expression between endocrine and exocrine tissue.

5.4.2 Immunohistochemical analysis of rat pancreata

Direct analysis of tissue samples is perhaps a more pragmatic approach for investigating gene and protein expression as primary cells have not been subject to immortalisation or prolonged culture and handling in the laboratory, which may influence the expression profile of cells. Therefore, to substantiate our findings in AR42J cells, immunohistochemical analyses were used to assess WF protein expression in rat pancreata.

Whole rat pancreata were provided courtesy of Dr Steven Jacques, Molecular Neuroscience Group, University of Birmingham. Tissue was excised from two animals immediately following CO_2 narcosis (for other purposes), embedded, sectioned and immunostained (section 2.16.0). Anti-*WFS1* antibody (Proteintech; table 2.2) was used to detect the presence of wolframin and anti-insulin antibody (Invitrogen; table 2.2) was used to confirm the location and morphology of islets (β-cells). Fluorescently-labelled secondary antibodies conjugated to different fluorophors (table 2.2) enabled stained sections to be viewed on an Axioplan 2 microscope with Zeiss Neofluar lenses and Axiovision software (Carl Zeiss Ltd).

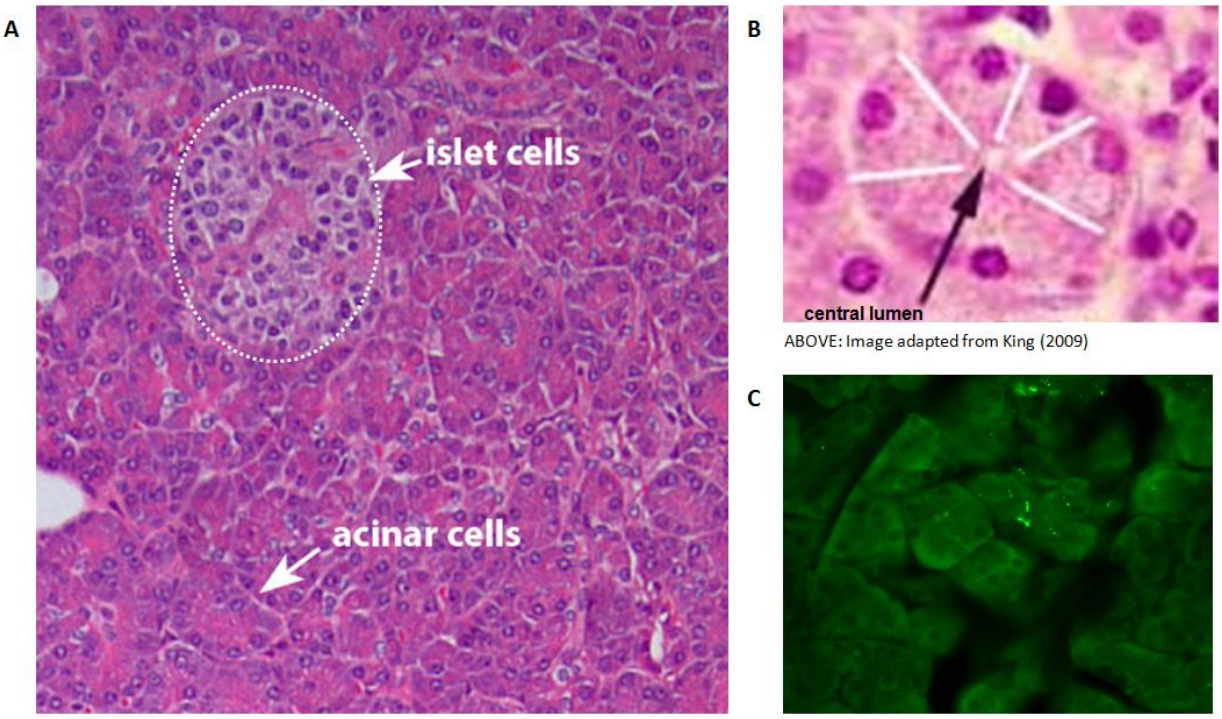
Experimental optimisation

A number of tissue sections were used in preliminary experiments to optimise and verify the antibody specificity and concentration. Negative control samples were prepared according to protocol (section 2.16.0) with omission of primary antibodies. Images were captured from the microscope using a series of exposure settings to assess background, non-specific staining with secondary Alexa Fluor antibodies. An antibody dilution series was also applied

to tissue sections to optimise concentrations for minimal signal to noise ratio. (appendices 31-34).

Histology of the pancreas

The bulk of the pancreas (~90%) is composed of exocrine tissue, from which secretions flow into a network of ducts leading ultimately to the duodenum. The exocrine secretory units, acini, are pyramid-shaped cells with basally located nuclei arranged into tubules to allow protein secretions passage directly from cells to ducts (fig 5.6A-C). Dispersed within the exocrine tissue are Islets of Langerhans, distinctive clusters of hormone-secreting endocrine cells including α, β, δ and PP (pancreatic polypeptide) cells (fig 5.6A).



ABOVE: Image adapted from Whitcomb (2009)

ABOVE: Image adapted from King (2009)

Figure 5.6. **Histology of the pancreas**. (**A**) Hematoxylin and eosin (H&E) stained human pancreas illustrating the arrangement and morphology of exocrine and endocrine tissue. DNA/RNA is stained purple and cytoplasmic structures pink (**B**) A magnified image of H & E stained acini, a small ball of pyramid-shaped epithelial cells with a tiny central lumen (white labelled cell boundaries). (**C**) Section of rat pancreas (x100 mag) stained with anti-*WFS1* and Alexa-Fluor 488 antibodies showing the characteristic clusters of acini.

Detection of wolframin in exocrine cells

Rat pancreas sections stained with anti-insulin (*INS*) antibody (red) clearly demonstrated the location of islets (fig 5.7A). As expected, islets also gave rise to a strong signal (green) when probed for the presence of wolframin, known to be abundant in pancreatic β-cells (fig 5.7B and 5.8). The surrounding exocrine tissue also produced a less pronounced, but evident green signal (fig 5.8), suggesting that wolframin may additionally be expressed in acini. It is unclear at this stage whether the signal observed is specific, relatively weak expression of *WFS1* in acini, or if in fact the signal is only due to background staining.

Confirming these observations requires a suitable blocking peptide as a primary antibody absorption control and a *WFS1*-null negative control, both of which were not available at this time.

Immunoblot analysis of primary pancreatic tissue is an alternative method that could be used to substantiate *WFS1* expression patterns, but does require accurate separation of islets and acini to avoid cross contamination of cell-types. A number of techniques such as fluorescence-activated cell sorting (FACS) or laser-capture microdissection (LCM) could be employed for this purpose, but do pose a risk of chemical or mechanical injury to the cells, perhaps causing changes to normal characteristics and gene expression.

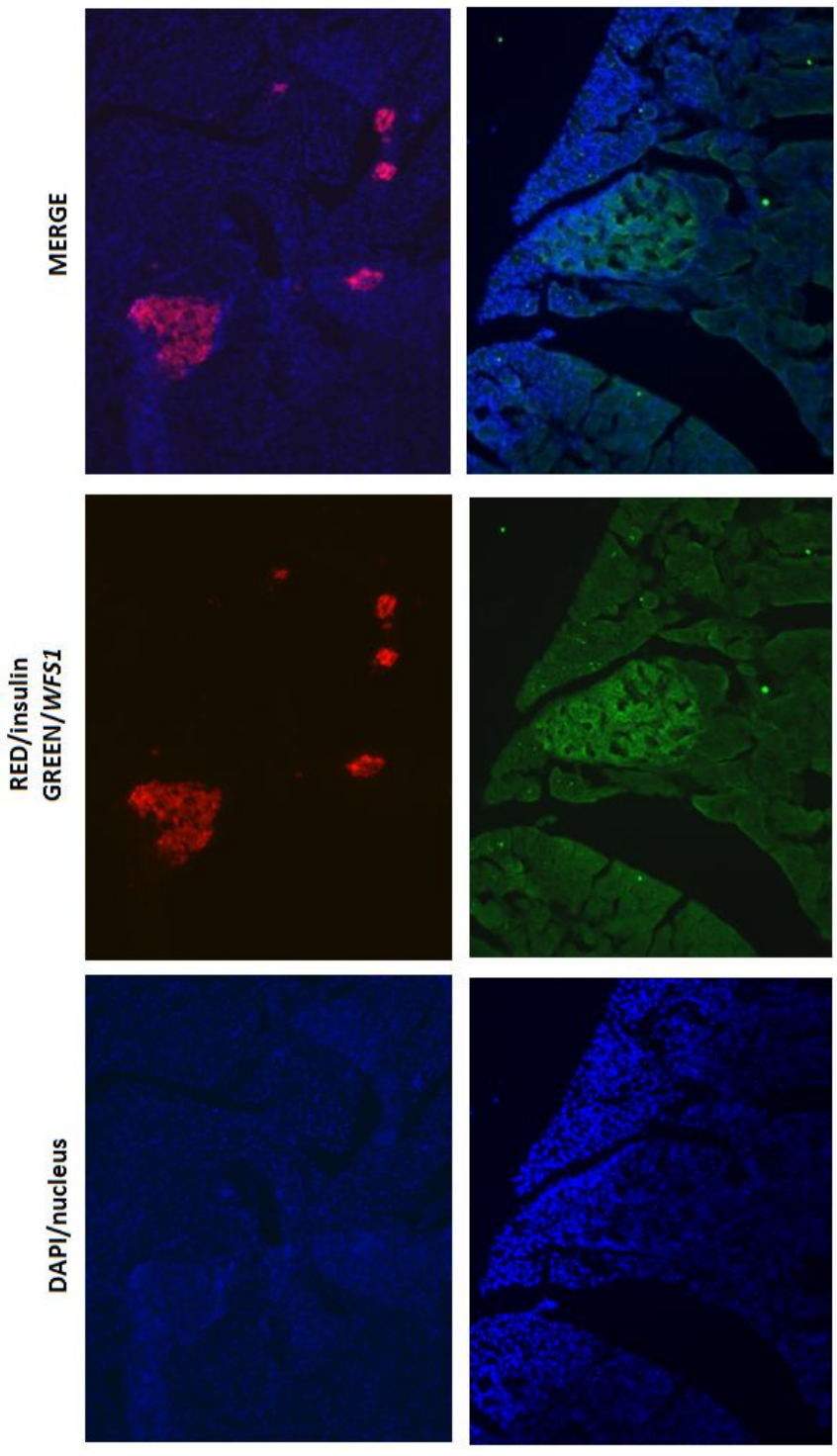


Figure 5.7. **Immunohistochemical analysis of rat pancreas sections (x100mag).** (**A**) Anti-insulin Ig (red) demonstrates the location of islet β–cells (**B**) Anti-*WFS1* Ig (green) as expected produced an intense signal in islets, but also appears to detect a weaker signal in the surrounding exocrine tissue.

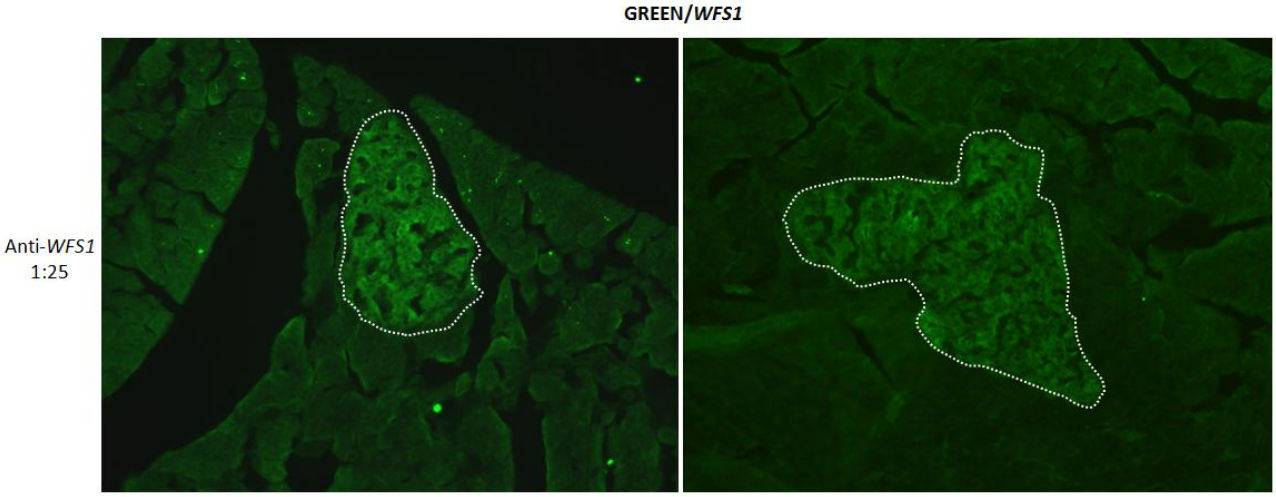


Figure 5.8. **Immunohistochemical analysis of rat pancreas sections (x100 mag).** A closer look at anti-*WFS1* detection in islets (outlined in white) and the staining pattern observed in the surrounding acinar cells.

Although our immunostaining data remain inconclusive at this stage, they do not contradict results obtained in rat AR42J acinar cells showing *WFS1* expression at RNA and protein levels. Overall our findings indicate that *WFS1* is expressed in exocrine tissue, supporting the notion that physical proximity of digestive zymogens and wolframin is possible in this tissue. Further investigation of yeast-isolated candidate proteins (table 5.0) as molecular partners to wolframin is therefore justified.

5.5.0 Cloning genes for proposed wolframin partners to mammalian expression vectors

To confirm WF interactions primarily identified in yeast, more detailed analysis was carried out in mammalian cells. From those isolated in the two-hybrid screen (table 5.0), several candidate 'partner' proteins were initially selected: lipase, colipase, REG protein and carboxypeptidase B1.

Sequence verified IMAGE clones were purchased from the European Collection of Cell Cultures (ECACC). These vectors containing cDNA sequences for the specified genes were used for cloning to HA-tagged mammalian expression vectors, pCMV-HA (Clontech), using a range of restriction sites and intermediate vectors as appropriate (section 2.19.0). Table 5.1 lists the chosen genes/proteins and their relating unique IMAGE clone reference ID number, the cloning vectors used and MCSs.

Details of the final expression vectors are illustrated in figure 5.9. *WFS1* cloned into a myc-tagged expression vector (pCMV-myc) was provided courtesy of Dr Malgosia Zatyka, Dept. Medical and Molecular Genetics, University of Birmingham.

Gene	protein	IMAGE ID	intermediate sub-cloning vector	CS for pCMV-HA ligation
CLPS	colipase	3950152	pGEM-T	*EcoR1/Xho1*
PNLIP	pancreatic lipase	4042766	pGEM-T	*EcoR1/Not1*
REG1A	pancreatic stone protein	3950349	pGEM-T	*Sal1/Not1*
CPB1	carboxypeptidase B1	4422620	pJET1.2	*EcoR1/Xho1*

Table 5.1. **Cloning methods for selected *WFS1*-partner candidates**. IMAGE clone ID (Geneservice), intermediate cloning vector and cloning sites used for ligation to pCMV-HA expression vectors.

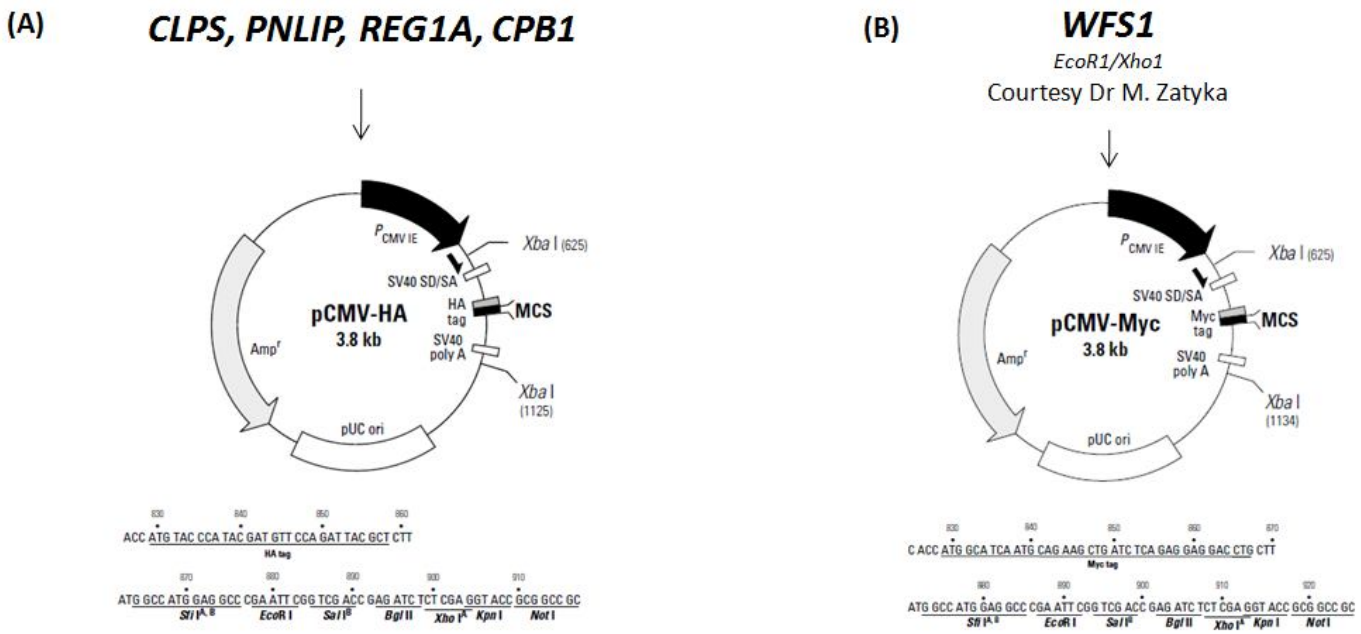


Figure 5.9. **Cloning methods for selected *WFS1*-partner candidates.** N-tagged expression vectors for cloning/transfection of *WFS1* and proposed partners: (**A**) colipase, pancreatic lipase, stone protein and carboxypeptidase were cloned into pCMV-HA . (**B**) *WFS1* (courtesy of Dr M Zatyka) was cloned into pCMV-Myc

5.6.0 Transfection and over-expression of wolframin and candidate partner-proteins in mammalian cells

Plasmids containing HA-tagged; *CLPS*, *PNLIP*, *REG1A, CPB1,* and Myc-tagged *WFS1* were transfected and co-transfected into mammalian cells (section 2.9.5). Experiments were aimed at over-expressing each protein for the purposes of co-immunoprecipitation. A series of directions were taken for protocol optimisation, summarised in figure 5.10, encompassing transfection, sample preparation and detection.

5.6.1 Method Development

Pilot

Preliminary experiments began with transfection of COS7 cells with 1µg plasmid DNA and Fugene6 transfection reagent (Roche). Transfection mixes containing each of the chosen 'partner' genes and *WFS1* respectively were prepared according to standard protocol (section 2.9.5) and incubated for 15 minutes before adding to cells at 70-80% confluency. Protein was harvested in RIPA buffer after a 24 hour incubation period (section 2.10.0), and samples were prepared for SDS-PAGE by sonication (2x 10 seconds), centrifugation for 30 minutes (14,000rpm), and collection of supernatant. Samples were mixed with x2 sample loading buffer, electrophoresed on SDS-polyacrylamide gel and immunoblotted (section 2.13.0) using anti-HA/anti-Myc antibodies (Sigma; table 2.2).

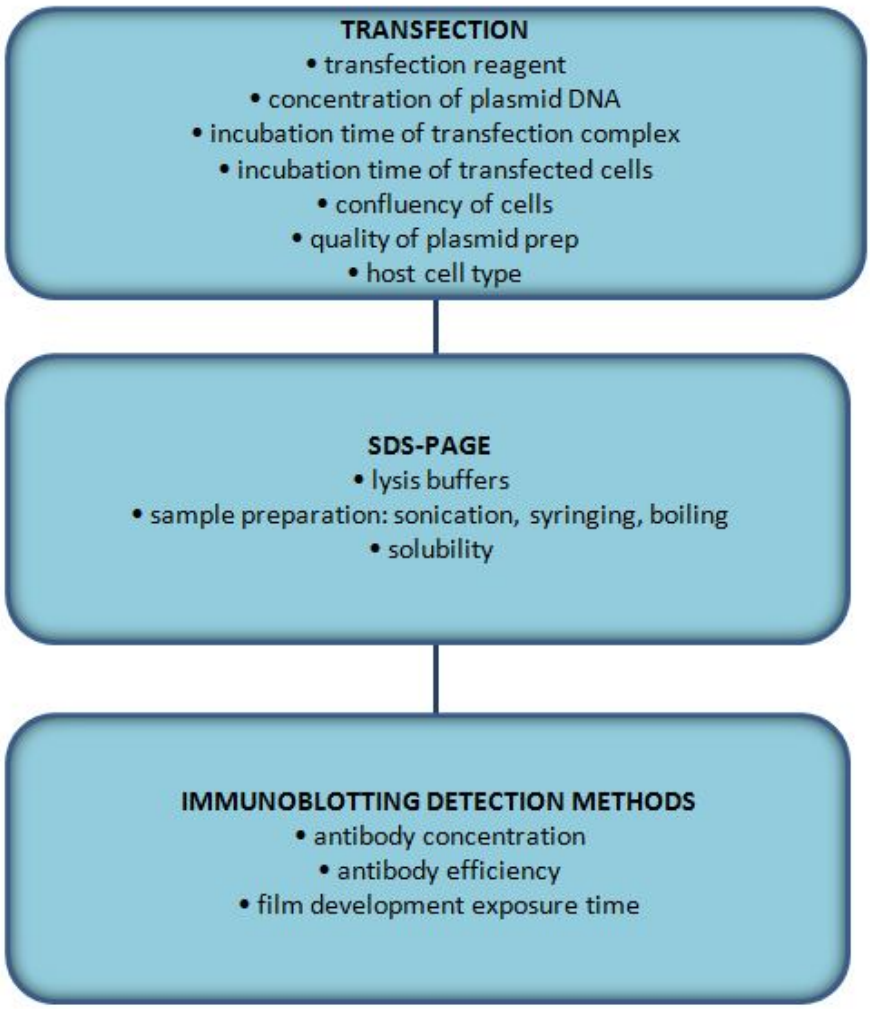


Figure 5.10. **Method development.** Three primary approaches for optimisation of *CLPS*, *PNLIP*, *REG1A*, and *CPB1* over-expression relating to: transfection, protein harvest and sample preparation for western blotting, and signal detection.

Transfection

Transfection conditions were the first set of parameters optimised for detection of protein over-expression. Cell lines routinely used to support reliable over-expression of a wide range of proteins were used. Both COS7 and HEK293 cells were transfected at a range of densities and manipulations were made with: DNA concentration, and DNA to Fugene6 ratio, incubation times for transfection mix and for cells post-transfection. FugeneHD

(Roche), a 'next generation' transfection reagent was also incorporated into the protocol to improve transfection efficiency. The quality of plasmid DNA was verified with fresh preparations, which were measured and independently transfected.

SDS-PAGE

The next stage of method development was concerned with protein harvest and preparation for polyacrylamide gel electrophoresis. Transfected cells were harvested in a range of lysis buffers (table 2.1) and subject to various pre-treatments prior to western blotting including: sonication and syringing, sample boiling, resuspension and attempted homogenisation of pelleted debris following centrifugation. A more concentrated sample loading buffer (x5) was also included enabling larger sample volumes to be loaded on gels.

Detection

Finally, the effectiveness of protein/signal detection was explored using alternate antibodies (table 2.2) and dilution series. Signal detection was maximised with high sensitivity ECL plus detection reagents (Amersham), reflective developing cassettes with high sensitivity X-ray film, and increasing film exposure times.

5.6.2 Over-expression of *WFS1* and proposed molecular partners (*CLPS*, *PNLIP*, *REG1A*, *CPB1*)

We were able to demonstrate over-expression of *WFS1*-Myc and HA-tagged; *CLPS*, *PNLIP*, *REG1A* and *CPB1* in mammalian cells following single and co-transfection experiments.

CLPS-HA and *WFS1*-Myc were used in preliminary time-series transfection experiments, where protein was harvested at intervals from 12-72 hours. Results indicated that the highest concentrations of colipase protein could be collected from cells 48 hours post-transfection following our protocol (fig. 5.11A). *WFS1* expression continued to increase at each time interval up to 72 hours post-transfection, although suitably high levels of protein were detectable after 48 hours (fig. 5.11B) suggesting this as the most appropriate time-scale for subsequent co-transfection experiments. HEK293 and COS7 cells were transfected with *WFS1* and *CLPS* and found equally able to support over-expression of both proteins (fig. 5.12).

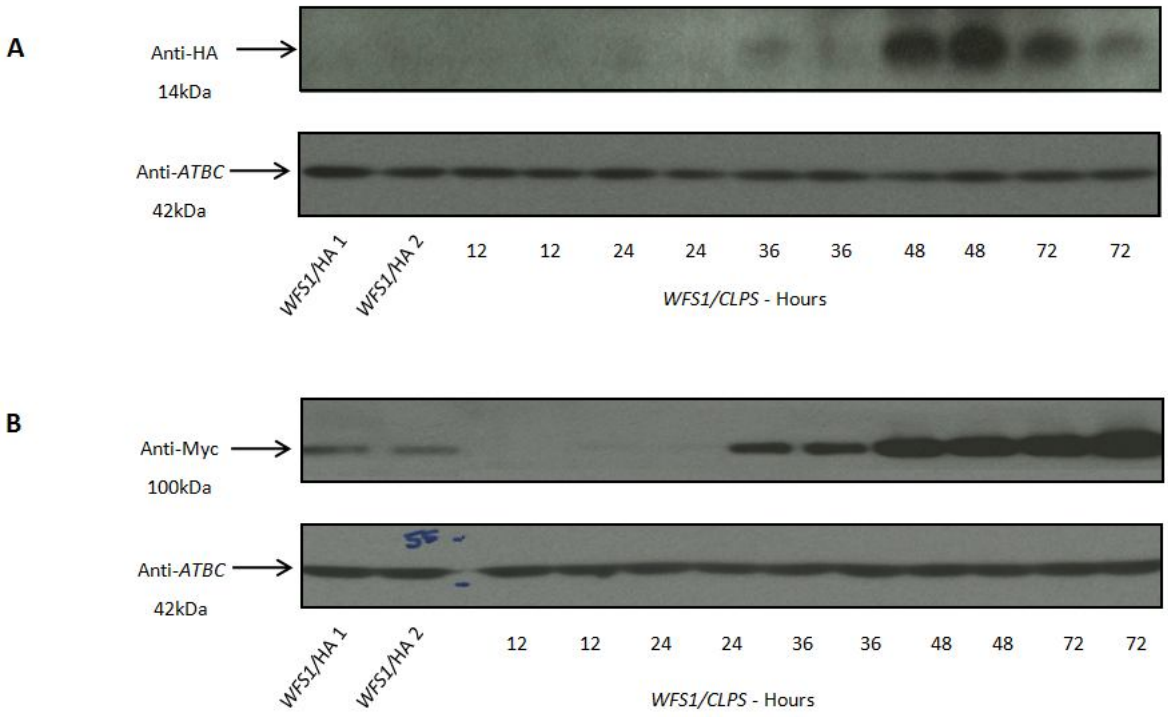
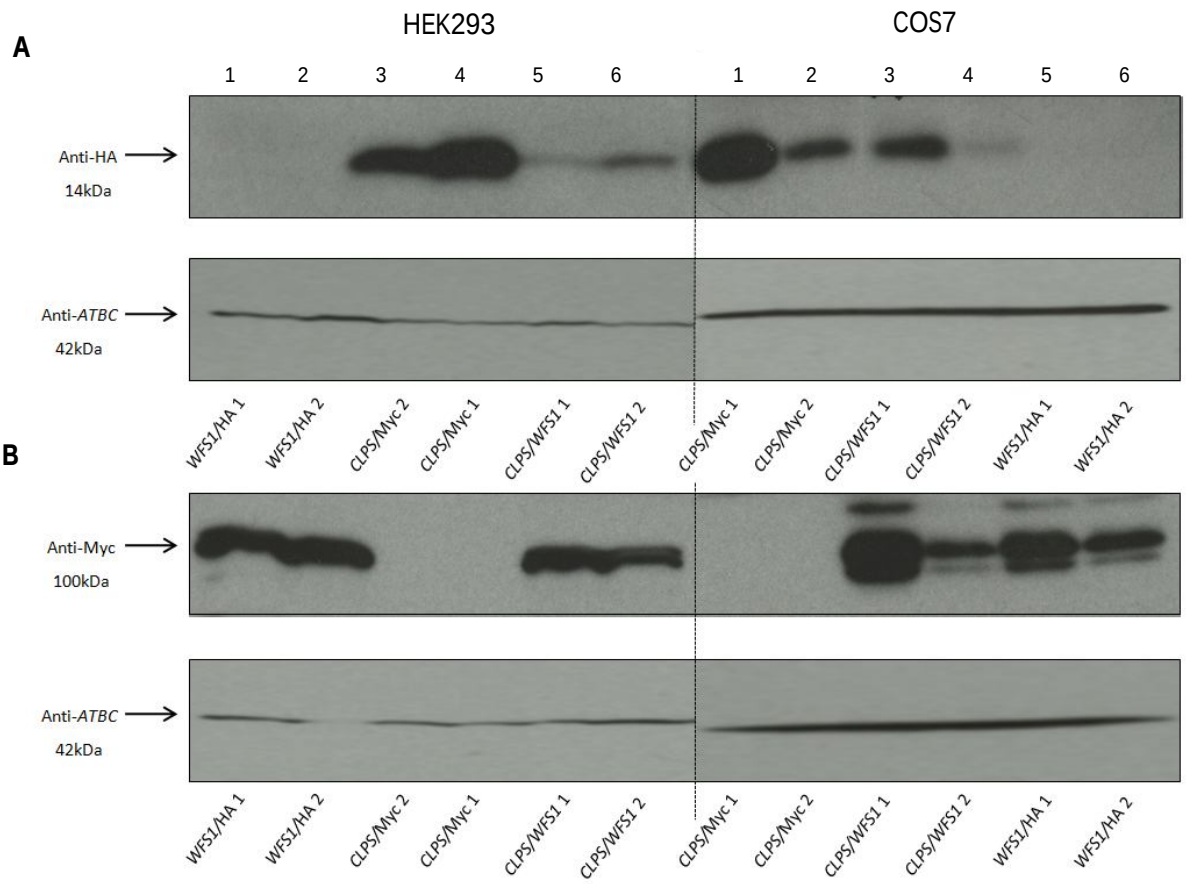


Figure 5.11. **Method development - Testing different post-transfection incubation times.** Western blot analysis of HEK293 cells co-transfected with *CLPS*-HA and *WFS1*-myc harvested at time intervals, showing optimal time for good expression of both at 48hours.



KEY

CLPS *- HA-tagged CLPS,* ***WFS1*** *– Myc-tagged WFS1,* ***HA****- empty HA vector,* ***Myc****- empty Myc vector,* ***ATBC*** - β-actin

Figure 5.12. Method development – Comparison of protein over-expression in host cell lines. Qualitative western blot analysis of *CLPS*-HA and *WFS1*-Myc over-expression in HEK293 and COS7 cells showing both cell lines equally support over-expression of *WFS1* and *CLPS*. Co-transfections of *CLPS* and *WFS1* (*CLPS*/*WFS1*) or each with empty negative control vectors (*WFS1*/HA, *CLPS*/Myc) were carried out in duplicate. Anti-β-actin Ig was used to give an indication of protein loading. (**A**) *CLPS* detected with anti-HA Ig (**B**) *WFS1* detected with anti-Myc Ig.

Colipase (*CLPS*) was the first candidate randomly selected for transfection and immunoblotting experiments. Figure 5.13 demonstrates a good level of *WFS1*-myc expression at ~100kDa and *CLPS*-HA at ~14kDa both singly and co-transfected into HEK293 cells. Protein was harvested for SDS-PAGE in RIPA lysis buffer according to standard protocol (section 5.6.1) 48 hours post-transfection. However, detection of *CLPS*-HA over-expression

following immunoblot was inconsistent despite large numbers of repeat experiments in seemingly identical conditions. We tried a wide range of method modifications for transfection, SDS-PAGE and detection although none of these conditions improved the consistency of the results (data not shown).

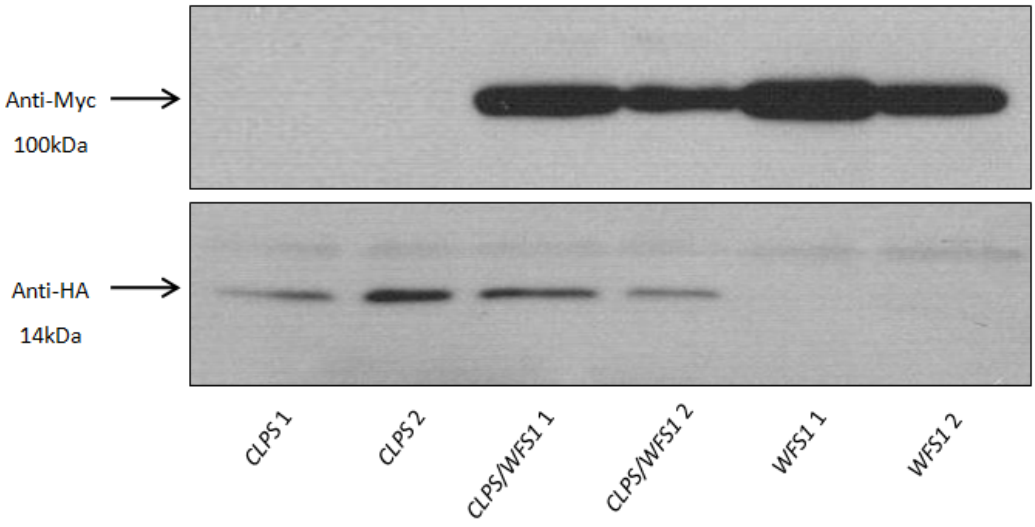


Figure 5.13. **Expression of *WFS1* and *CLPS* in HEK293 cells.** Qualitative western blot analysis of *CLPS*-HA and *WFS1*-myc over-expression from single and co-transfections. Cells were harvested in RIPA buffer after 48 hours and lysate supernatant was collected following sonication.

The remaining three candidates; pancreatic stone protein (*REG1A*), carboxypeptidase B1 (*CPB1*) and pancreatic lipase (*PNLIP*) were also transfected and over-expressed in HEK293 cells. Proteins of appropriate size were observed on western blots probed with anti-HA Ig, but again these results proved difficult to replicate reliably.

Pancreatic stone protein over-expression in HEK293 cells was detected at 19kDa by anti-HA Ig. Samples prepared according to standard protocol in RIPA buffer supernatant are shown in figure 5.14A with corresponding samples probed for *WFS1*-Myc in the lower panel. *REG1A*-HA protein was found in samples derived from all combinations of transfection; *REG1A* alone, with *WFS1*, and with an empty control vector, although co-transfections with *WFS1* produced significantly lower levels of *REG1A*, with signals barely visible even after extended film exposure.

Figure 5.14B shows protein harvested from the same transfection as those in figure 5.14A, but collected from insoluble material following centrifugation (section 2.10.1). Cell debris, usually spun into an insoluble pellet and discarded following sample preparation was resuspended in a small volume of RIPA buffer (table 2.1) and repeatedly pipetted up and down in an effort to homogenise. This material was then immunoblotted with anti-HA Ig to reveal the presence of a substantial amount of *REG1A* protein in this fraction of the lysate. Large amounts of *WFS1*-Myc (lower panels) were also present in the insoluble portion of cell lysates.

After discovering relatively abundant *REG1A* protein in the insoluble fraction, we performed a range of experiments with the aim of solubilising *REG1A*. A range of lysis buffers were used to harvest protein (table 2.1). Other means of releasing *REG1A* from (aggregated) pelleted material were also attempted including further sonication, syringe homogenisation and boiling samples. For the purposes of co-expression with *WFS1*, maintaining optimal conditions for wolframin was found to be a principle limitation to these methods of sample preparation.

Interestingly, we were only able to determine carboxypeptidase B expression in the insoluble fraction of RIPA buffer lysates (fig 5.15). Although transfection, harvesting and immunoblotting conditions were kept consistent between samples we observed little or no *CPB1* expression from the crude 'resuspended' material used for western blot when co-transfected with *WFS1*-Myc, yet clear signals with anti-HA Ig (46kDa) from *CPB1* samples transfected alone or with control vectors. The possible biological significance of this observation remains to be investigated.

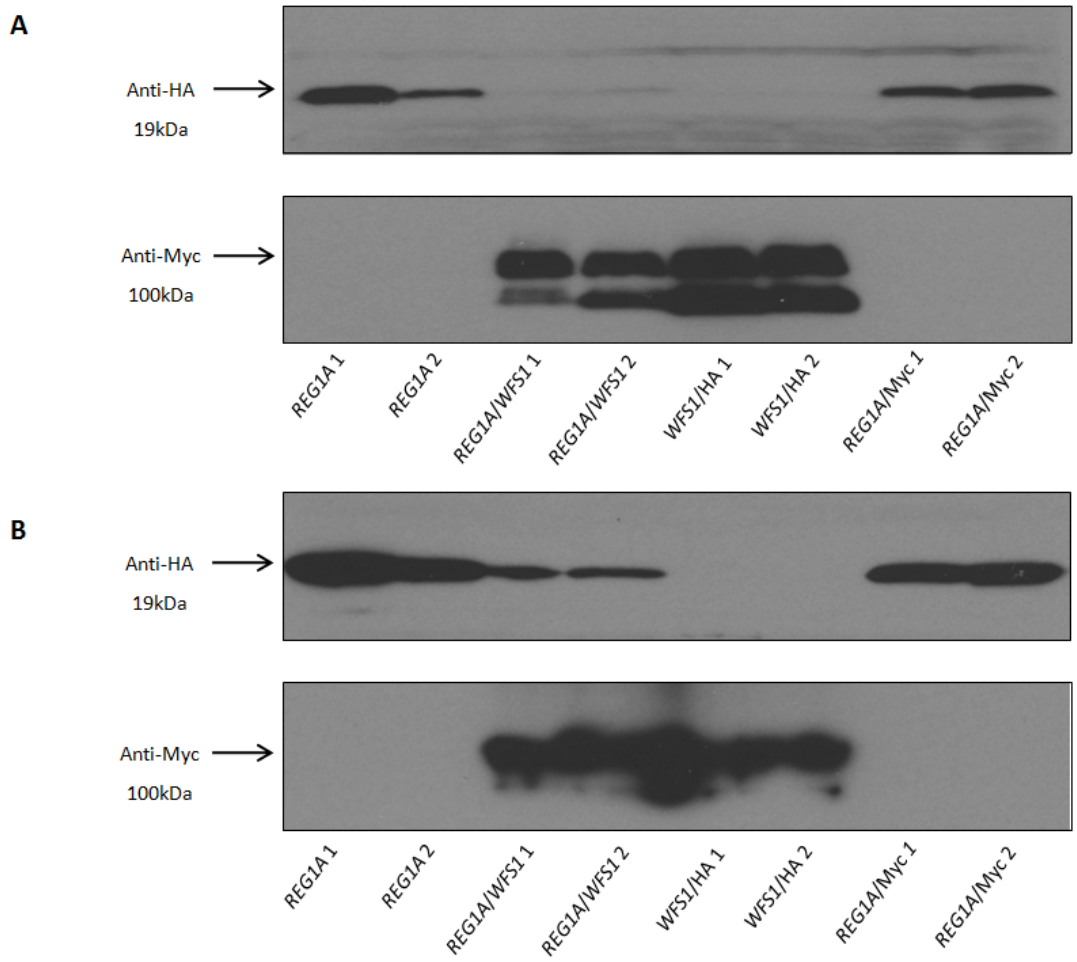


Figure 5.14. **Method development- Comparison of *REG1A* expression in lysate supernatant and pellets.** Qualitative western blot analysis of HEK293 cells transfected and co-transfected with combinations of *REG1A*-HA, *WFS1*-myc and empty control vectors. (**A**) Supernatant from cell lysates in RIPA buffer prepared according to standard protocol. (**B**) Resuspended pellets from insoluble material following lysis in RIPA buffer. More *REG1A* protein was detected in the insoluble fraction than in supernatant.

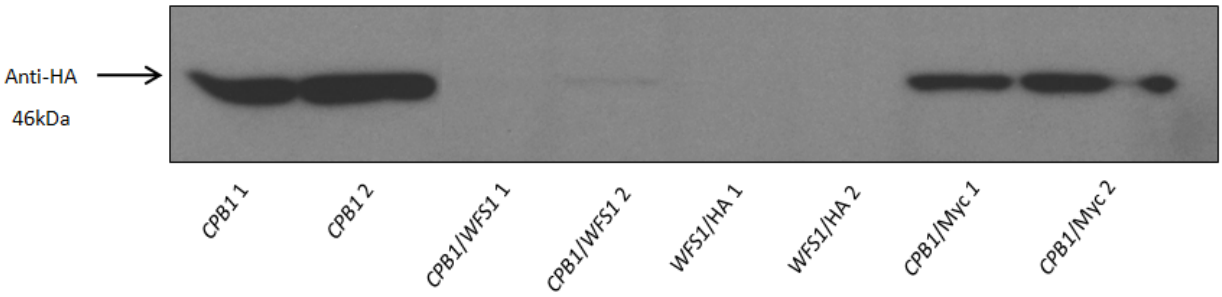


Figure 5.15. **Method development – Expression of *CPB1*-HA in HEK293 cells.** Qualitative western blot analysis of HEK293 cells transfected and co-transfected with combinations of *CPB1*-HA, *WFS1*-myc and empty control vectors. Insoluble pellets resuspended in RIPA buffer, vigorously mixed and boiled in loading buffer. *CPB1*-HA was only detectable in insoluble portion of cell lysates; a good level of expression was observed in cells transfected with *CPB1*-HA alone or with an empty control vector but little or no expression was detectable when co-transfected with *WFS1*-Myc.

Pancreatic lipase was detected at 51kDa, again only detectable in the insoluble fraction of RIPA buffer lysates (fig 5.16). Although samples were not measured quantitatively here, obvious inconsistencies in expression levels were observed between both different and duplicate samples. In this instance, each of the *PNLIP* transfection combinations does produce a low intensity signal although some positive samples (e.g *WFS1/PNLIP* 1) clearly have comparably higher concentrations of *PNLIP*.

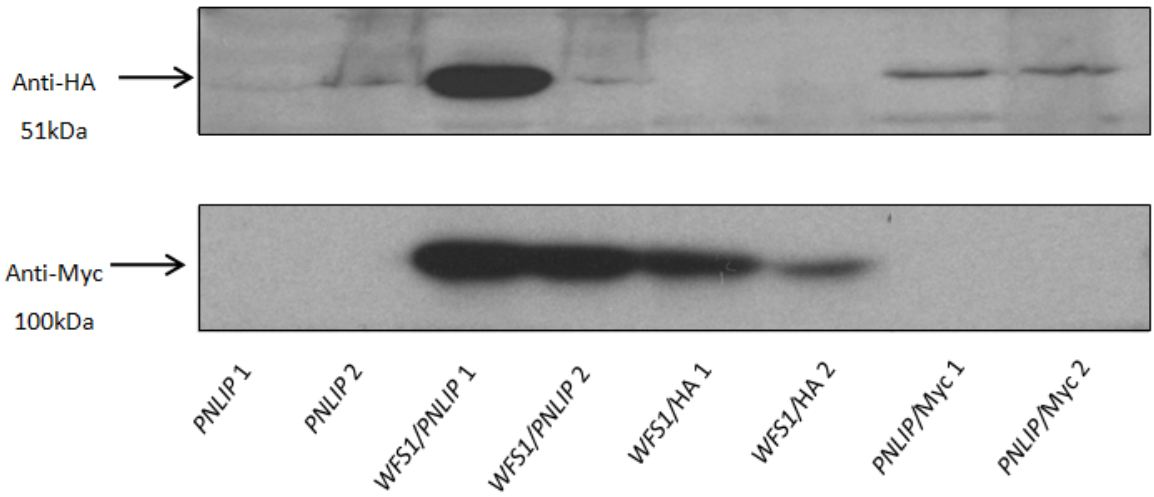


Figure 5.16. **Method development - Expression of *PNLIP*-HA and *WFS1*-Myc in HEK293 cells.** Qualitative western blot analysis of HEK293 cells transfected and co-transfected with combinations of *PNLIP*-HA, *WFS1*-myc and empty control vectors. *PNLIP*-HA was only detectable in insoluble portion of cell lysate; pellets were resuspended in RIPA buffer, vigorously mixed and boiled in loading buffer.

Closer consideration was therefore given to solubility as a factor influencing the inconsistent observations we made of our 'partner' proteins following immunoblotting. Samples harvested in a number of buffers were also subject to a range of pre-treatments prior to SDS-PAGE, and efforts to resuspend material from pelleted cell debris were made. These experiments are ongoing.

Summary of section

We have demonstrated consistent over-expression of *WFS1*-Myc in mammalian cell lines and over-expression of single and co-transfected HA-tagged: *CLPS*, *REG1A*, *CPB1*, and *PNLIP* with variable degrees of success. We have observed good expression of *CLPS*-HA in the

soluble fraction of RIPA buffer cell lysates, but these results are inconsistent between experiments. The remaining proteins (*REG1A*, *CPB1* and *PNLIP*) have been detected in the insoluble fraction only, again with inconsistency between experiments. A notable observation was the expression level of *CPB1*-HA, which was detectable in the absence of *WFS1*-Myc only, although this finding must be verified.

To date, our co-transfection experiments have been based on modifying our established 'optimal' wolframin harvesting and preparation conditions to accommodate the presence of selected 'partner' proteins. It is apparent that optimal conditions for reliable *WFS1* over-expression and immunoblot detection are not comparable to those required for our chosen molecular partners.

Previous studies have successfully demonstrated these proteins using western blotting techniques (Crenon, 1994; Cavard, 2006; Ventura, 1999; Lowe, 1992), furthermore, 'ready to load' prepared protein lysates are commercially available (Abcam, Origene). Consulting relevant literature enables us to ascertain the alternate harvesting and preparation conditions used for application in our own work. We intend to explore and adapt these protocols further with the aim of reaching sufficient, if compromised, conditions for adequate solubilisation of wolframin and 'partner' proteins together. Despite the method development needed for reliable WF-partner co-expression, we took forward the available protein we were able to effectively produce to co-immunoprecipitation experiments.

5.7.0 Co-immunoprecipitation of over-expressed wolframin and proposed molecular partners

Co-immunoprecipitation (Co-IP), a technique used to isolate protein-protein interactions *in vivo*, was used to purify WF-partner complexes from mammalian cell lysates. By targeting wolframin or a partner-protein with specific antibodies, it is possible to physically isolate (pull-down) from a solution not only the targeted antigen, but also any proteins that are bound to it, which can be analysed by SDS-PAGE and western blot.

5.7.1 Specific methods

An indirect capture method was employed using agarose A beads as a solid phase substrate (section 2.14.0). Agarose beads are highly porous, microscopic resin slurries with surfaces suitable for binding antibodies. When specific antibodies are immobilised onto agarose beads, this support scaffold is used to pull-down not only target antigens but entire protein complexes from solution (fig 5.17). The targeted antigen (e.g. wolframin) and its associated partners were firstly incubated with specific antibodies to allow antibody-antigen binding. Agarose beads were then added to the mix to immobilise and 'pull-down' complexes. Beads with attached proteins were then isolated and washed to remove unwanted non-specifically bound molecules before the targeted antigen (e.g. wolframin) and its associated partners were released from the beads and analysed.

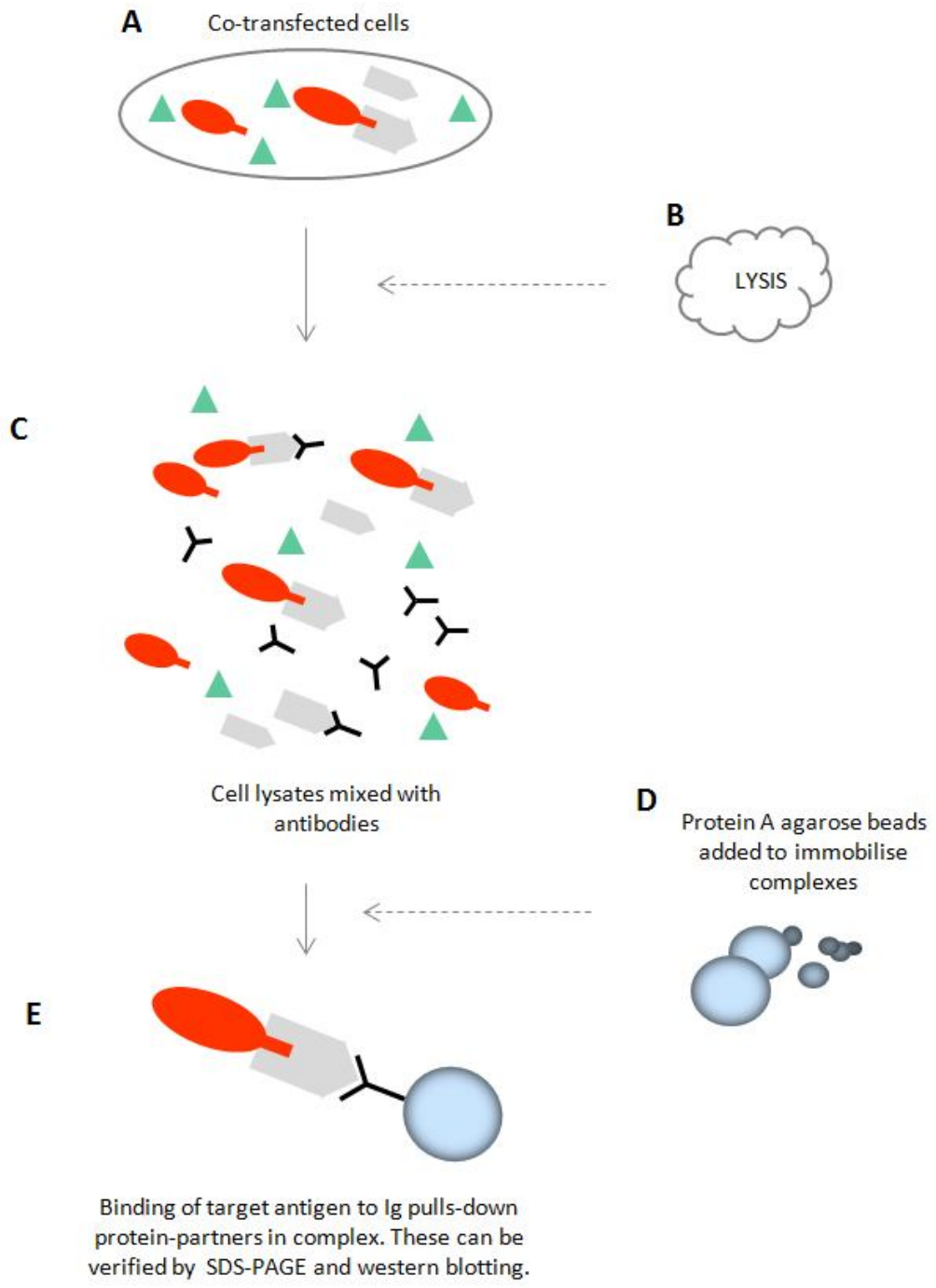


Figure 5.17. Principles of indirect co-immunoprecipitation. (A) Cells are co-transfected with both potential partner proteins. **(B)** Cells are lysed to release proteins. **(C)** Lysates are incubated with antibodies specific to a target antigen, e.g. one of the partners. **(D)** Agarose beads are added to the mixture to provide a solid-phase support for 'pulling-down' the entire complex as antibodies have affinity to the surface of the beads. **(E)** Immobilised complexes are purified from solution ready for analysis by western-blot.

Whole cell lysates, containing over-expressed proteins of interest were used for Co-IP. Experiments were initially performed using anti-Myc Ig for pull-down, which served to target WF-Myc, and anti-HA Ig for detection of any associated HA-tagged partner-proteins. Where material was available, the pull-down was carried out in reverse, using anti-HA antibody for isolation of potential partners, and anti-Myc for detection of any coupled WF-Myc protein.

Following pull-down, agarose beads (with attached proteins) were washed 3 times for 10 minutes in RIPA buffer (table 2.1). The stringency of this step was experimentally modified in attempts to minimise non-specific protein binding to beads and remove background staining of subsequent western blots; wash duration was increased and alternate buffers were applied.

5.7.2 Co-IP of *WFS1-CLPS*

The largest numbers of protein samples were generated from *CLPS* constructs, enabling a series of alternative IP experiments to be performed. Using anti-Myc Ig for targeting wolframin we were able to repeatedly detect the presence of 'pulled-down' *CLPS*-HA at 14kDa with anti-HA Ig (fig 5.18A), suggesting an association between *WFS1-CLPS* could be legitimate. The validity of our findings however, remains in question due to the recurrent 'pull-down' of *CLPS*-HA observed in our negative control, which comprised empty pCMV-Myc vector and *CLPS*-HA. The same experiment was carried out in reverse using anti-HA to target *CLPS* (fig 5.18B). We found that WF-Myc was pulled-down as expected, but again a clear signal was present in the negative control (*WFS1*/empty pCMV-HA).

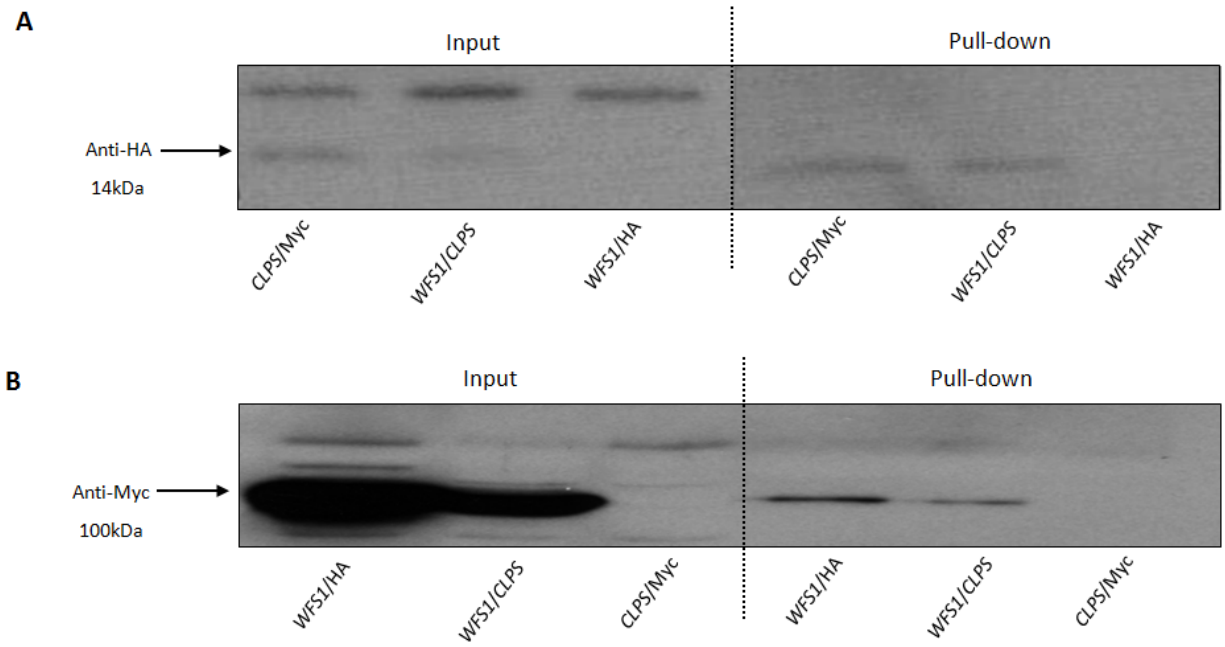


Figure 5.18. Co-immunoprecipitation of tagged wolframin (*WFS1*-Myc) and colipase (*CLPS*-HA), harvested from HEK293 cells in RIPA buffer. Pull-down of protein complexes was performed with (**A**) anti-Myc Ig, specific to *WFS1,* and (**B**) anti-HA Ig, specific to *CLPS*.

Since agarose is able to bind some proteins and cellular components, non-specific interaction and pull-down of off-target molecules directly with beads may be the cause for unwanted background signals. The pre-clearing step where lysates are incubated with beads alone is designed to minimise this outcome by collecting and removing non-target molecules, although increasing the stringency of washes following IP can further lessen this occurrence.

We repeated *WFS1-CLPS* Co-IP experiments, modifying the stringency of washing. Our aim was to maintain binding of desired specific proteins but strip away non-specific interactions. Increased number and duration of washes with RIPA buffer (fig 5.19A-B), and alternate buffers with variable concentrations of salt (NaCl) and detergents were applied (fig 5.20A-B), although detection of protein in the negative control samples persisted.

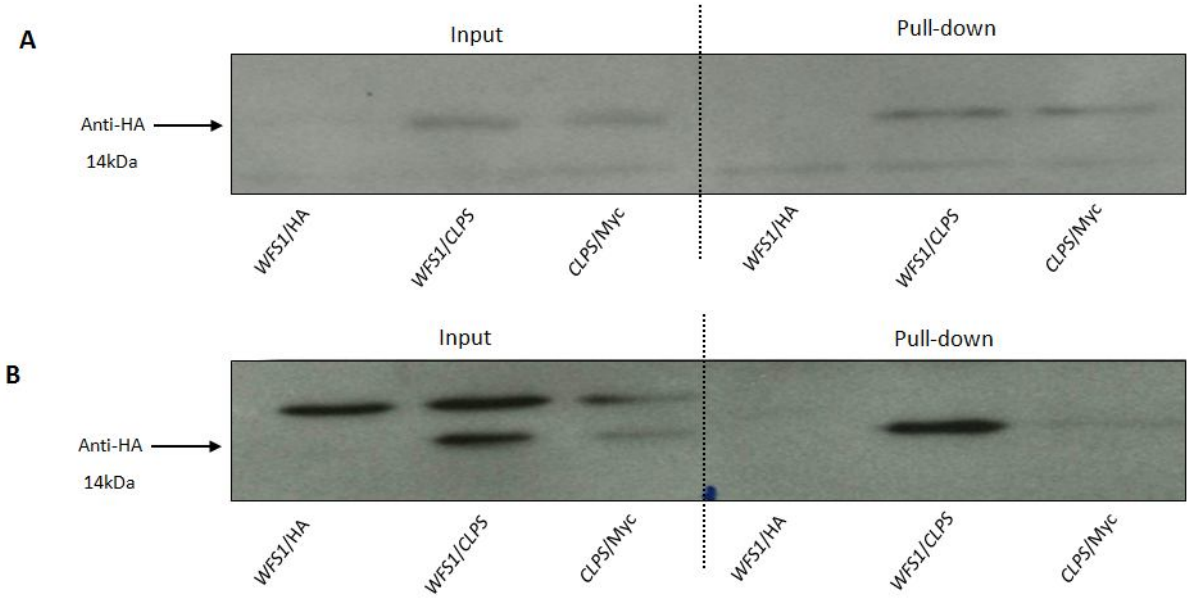


Figure 5.19. **Co-immunoprecipitation of tagged wolframin (*WFS1*-Myc) and colipase (*CLPS*-HA), harvested from HEK293 cells in RIPA buffer.** Pull-down of protein complexes was performed with anti-Myc Ig, specific to *WFS1,* and washed with RIPA buffer for: **(A)** 2 x 20 min **(B)** 3 x 20 min.

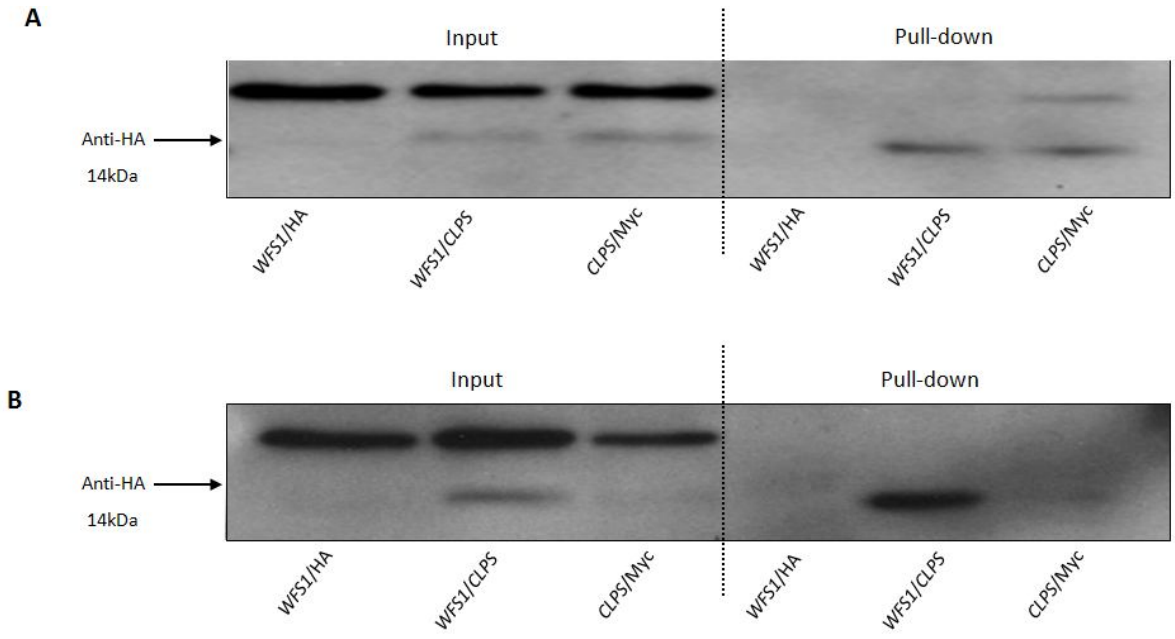


Figure 5.20. **Co-immunoprecipitation of tagged wolframin (*WFS1*-Myc) and colipase (*CLPS*-HA) harvested from HEK293 cells in RIPA buffer.** Pull-down of protein complexes was performed with anti-Myc Ig, specific to *WFS1,* and washed in: **(A)** high salt buffer (table 2.1) **(B)** low detergent buffer (table 2.1).

Generation of additional protein extracts suitable for Co-IP will enable us to continue along this line of investigation. Increasing buffer NaCl concentration to as much as 1M may disrupt any electrostatic or ionic attractions although some physiological interactions can also be lost at salt concentrations above 300mM, higher stringency can also be achieved by changing detergents and raising concentrations. Further reduction of non-specific protein binding may be achieved with more rigorous 'blocking' of beads with an irrelevant protein such as bovine serum albumin (BSA), which serves to reduce the available surface area for attachment of superfluous proteins. We confirmed the specificity of anti-Myc Ig immunoprecipitation by stripping a membrane of anti-HA Ig using 0.2M NaOH (sodium hydroxide) and reprobing it with anti-Myc Ig (fig 5.21).

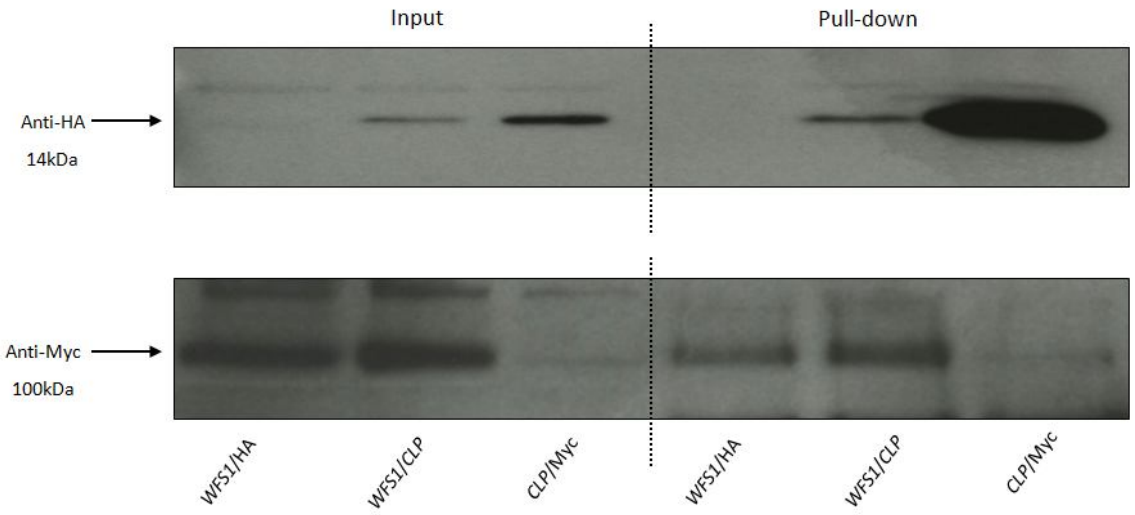


Figure 5.21. **Co-immunoprecipitation of tagged wolframin (*WFS1*-Myc) and colipase (*CLPS*-HA), harvested from HEK293 cells in RIPA buffer.** TOP PANEL: Pull-down of protein complexes was performed with anti-Myc Ig, specific to *WFS1.* BOTTOM PANEL: The membrane was probed with anti-HA Ig, specific to *CLPS*, then stripped with 0.2M NaOH and reprobed with anti-Myc Ig to verify correct binding pattern of *WFS1* constructs and specificity of anti-Myc Ig.

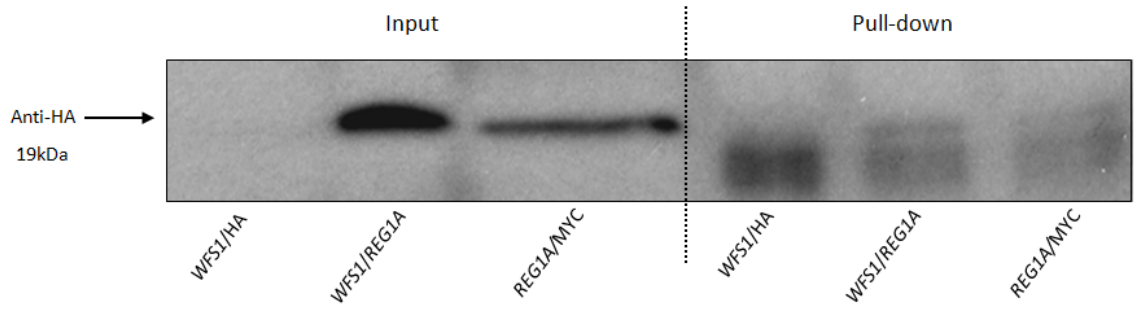
Immobilising antibodies/protein-complexes onto alternative supports may provide another solution to non-specific pull-down. Materials such as acrylamide and cellulose resins, magnetic beads, polystyrene particles and controlled pore glass (Thermo Scientific) are all popular options for IP protocols.

5.7.3 Co-IP of *WFS1-REG1A*

Preliminary Co-IP was also carried out with *WFS1-REG1A* combinations, collected from both soluble and insoluble lysate material. Pull-down of complexes was achieved with anti-Myc Ig, specific to *WFS1*-Myc. We were able to detect a positive *REG1A* signal in each sample set (fig 5.22A-B), but as before a clear signal in negative controls also.

Although providing a useful starting point, the use of material not fully dissolved does interfere somewhat with the precipitation process, where protein aggregates often stick to beads randomly. Proteins of interest may perhaps be isolated in this instance by binding of another unrelated protein, which happens to form part of the composite.

Supernatant (A)



Pellets (B)

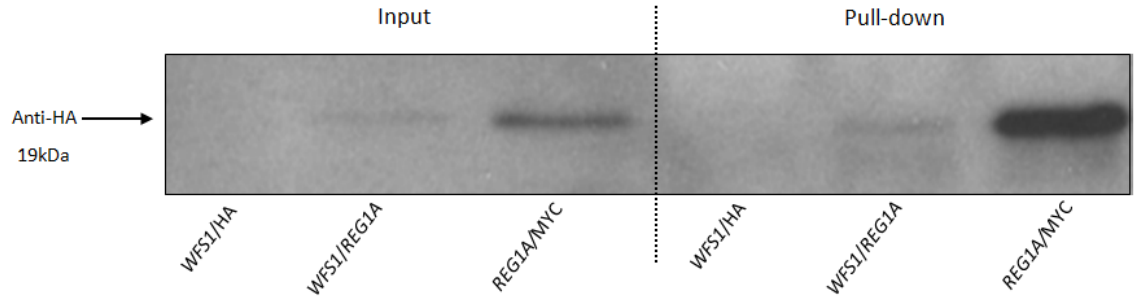


Figure 5.22 Co-immunoprecipitation of tagged wolframin (*WFS1*-Myc) and pancreatic stone protein (*REG1A*-HA). Over-expressed proteins harvested from HEK293 cells in RIPA buffer **(A)** supernatant and **(B)** insoluble material. Pull-down of protein complexes was performed with anti-Myc Ig, specific to *WFS1*.

5.7.4 Co-IP of *WFS1-CPB1*

The only *CPB1* protein detected so far was observed in resuspended insoluble material from RIPA buffer lysates. We attempted to Co-IP *CPB1*-HA and *WFS1*-Myc with anti-Myc Ig despite limitations imposed by partial insolubility. Very faint signals following immunoblot were captured both in sample inputs and pull-downs (fig 5.23), although appropriately sized bands were visible at 46kDa when probed with anti-HA, specific for *CPB1*-HA. Again signals were also detected in negative control samples.

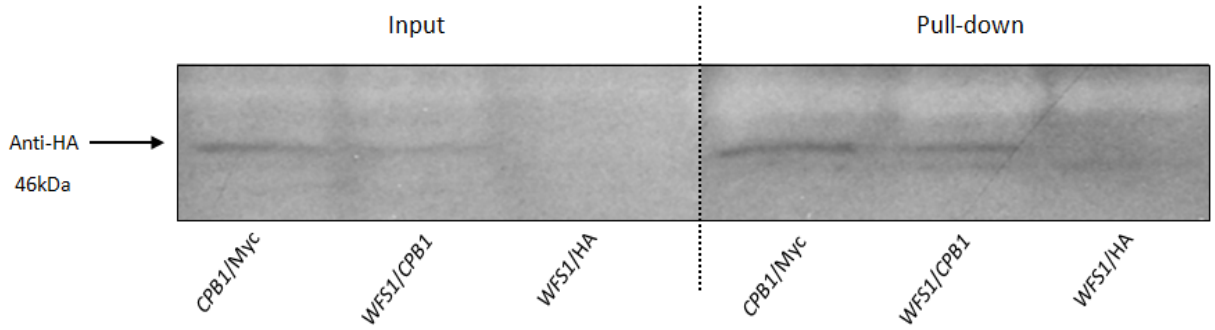


Figure 5.23 Co-immunoprecipitation of tagged wolframin (*WFS1*-Myc) and carboxypeptidase B1 (*CPB1*-HA). Over-expressed proteins harvested from HEK293 cells in RIPA buffer (pellets). Pull-down of protein complexes was performed with anti-Myc Ig, specific to *WFS1*.

Section summary

Our earliest attempts to confirm proposed WF-partner interactions in mammalian cells have been carried out with a limited supply of material. To date we have been able to confirm the 'potential' for *CLPS*, *REG1A* and *CPB1* to form a complex with *WFS1* using Co-IP although experiments are still at the stage of method development.

5.8 Ongoing and further work

5.8.0 Protein solubilisation

Since proteins with different physio-chemical properties are not easily amenable to the same treatment, it is difficult to study these on a systems level in a lab controlled environment. Our first goal is to define appropriate conditions for WF-partner co-expression for analysis by SDS-PAGE. A number of options remain open for optimising solubilisation of these proteins together, including varied lysis buffers and pre-treatments.

Export of secretory proteins and cleavage of signal peptides is known to play an important role in proper protein folding, formation of disulphide bonds and tertiary structure (Vitale, 1999; Phillips, 1996). Taken together with successful isolation of lipase and carboxypeptidase proteins from cell culture medium as well as cell pellets in earlier studies (Lowe, 1992; Reverter, 1998; Ventura, 1999), we hypothesise that our solubility issues may arise from incorrect folding of our candidate proteins, which all naturally pass through the secretory pathway. We may turn our attention to collection and purification of over-expressed proteins from culture medium, although perhaps isolating WF-partner pairs together in this way may not be possible. Conceivably, a potential interaction with wolframin may be necessary for correct folding or export of selected secretory proteins.

An *in vitro* transcription-translation method is another approach we could consider for tackling issues of combined protein solubility. Cell-free systems of synthesising target

proteins from DNA/RNA templates, have been employed successfully for study of cytotoxic products or those tending to aggregate or form inclusion bodies. Extracts containing the basic translational machinery including ribosomes, initiation factors, tRNAs and ATP can be used to generate proteins for a limited period. A number of commercial kits are available with components derived from prokaryotic and eukaryotic organisms, including human, which deliver functional proteins.

5.8.1 Optimising Co-IP conditions

Following successful production of our desired proteins we can concentrate on improving the specificity of our Co-IP experiments. A number of repeat experiments would enable modifications to be made to the stringency of pre-clearing lysates and wash buffers. Wash buffers containing increasingly higher concentrations of denaturants should serve to break non-specific interactions until a point is reached where all proteins bound to beads/immobilised antibodies are removed. In order to verify a positive association, under these conditions we would expect to maintain WF/partner pull-downs but observe a disappearing signal from the negative control pull-down.

Another approach we could try is direct Co-IP, which has proven more efficient in certain biological systems. Antibodies are immobilised onto a solid phase support such as beads before incubating with lysates. Antigen-antibody complexes are allowed to form before purification of these partners from solution with agarose beads already conjugated to antibodies.

Assuming adequate solubilisation of both proteins, we could also attempt Co-IP with one or more endogenous proteins. Endogenous wolframin, known to be expressed in COS7 and HEK293 cells, could be used to pull-down HA-tagged 'partner-proteins' with anti-*WFS1* Ig (and vice versa). In light of our earlier data (section 5.4) we may ultimately perform Co-IP of purely endogenous proteins found in acinar cells if possible.

5.8.2 Alternative approaches for confirming protein interactions in mammalian systems

If suitable conditions for solubilising both wolframin and our proposed partners (*CLPS*, *PNLIP*, *REG1A* and *CPB1*) together are not ascertained, we may consider an alternative approach for confirming/eliminating our findings from yeast in a mammalian system.

An altogether different method of producing proteins and investigating their interactions *in vitro* was described by Suzuki *et al* (Suzuki, 2004). Instead of using tagged proteins for Co-IP experiments, the target proteins are synthesised and biotinylated or radiolabeled through *in vitro* transcription-translation. Although not without drawbacks, this system produces proteins more likely to be soluble and eliminates the need for traditional fusion-tags.

In vivo protein-interaction assays such as FRET (fluorescence resonance energy transfer) and more recently BiFC (biomolecular fluorescence complementation) allow analysis of protein associations in a more physiologically relevant context. FRET technology is based upon constructing fusion-proteins with coupled fluorophores, which transfer energy when in close proximity. When energy is transferred from excited donor to acceptor, a change in fluorescence will take place that can be visualised with an inverted fluorescent microscope.

BiFC (Morell, 2008) involves the construction of fusion-proteins where proteins of interest are linked to complementary unfolded fragments of a fluorescent reporter. If the proteins interact, the reporter is able to reform its three-dimensional structure and emit a fluorescent signal, detectable under an inverted fluorescent microscope. FRET and BiFC are useful techniques that not only allow detection of interaction sites within live cells, but are also highly sensitive, enabling capture of short-lived or weak associations.

5.9 Summary and discussion

A number of potential wolframin molecular partners, proposed to interact with the cytoplasmic N-terminal domain of wolframin, have been detected from a yeast two-hybrid assay. *WFS*1-N was used as bait to screen a pancreatic cDNA library (Dr C. Ricketts) and from those positive yeast clones collected we have isolated and identified: *AMY2 A/B*, *CPB1*, *CLPS*, *PNLIP*, *PRSS2*, and *REG1A*, all of which encode secretory proteins synthesised in the exocrine pancreas. These interactions were all confirmed by 'one-to-one' reintroduction to yeast.

Before verifying these findings in a mammalian system, we firstly sought to provide evidence that these interactions could be biologically relevant, where both wolframin and proposed partner-protein may be present in the same locality. Since each of the candidate 'partners' are produced and secreted by exocrine cells, we considered this a reasonable place to look for the presence of wolframin, which has never previously been shown in this tissue type. A rat exocrine acinar cell line (AR42J) was used for this purpose, and we were

able to demonstrate the expression of *WFS1* on both RNA and protein levels. Immunohistochemical analyses of whole rat pancreas sections also provided inconclusive support for these observations but requires *WFS1* negative tissue as a control.

Four candidates were selected: *CLPS*, *PNLIP*, *REG1A* and *CPB1,* and corresponding cDNA sequences were cloned into HA-tagged expression vectors. Together with *WFS1*-Myc, each of the candidate partners was transfected and over-expressed in mammalian cells. Finding conditions for abundant, reproducible protein expression was problematic. Following a series of optimisation experiments, we found the most obvious obstacle was protein solubility, in particular preparing both wolframin and a designated 'partner' for immunoblot analysis in conditions suitable for both proteins.

A series of preliminary Co-IP experiments were carried out from the limited material we were able to generate. We demonstrated pull-down of *CLPS*, *REG1A* and *CPB1* using *WFS1*-myc Ig although adequate controls are needed to substantiate these early findings. Our initial results do not discredit the idea that wolframin has the 'potential' to bind each of the selected partners (*CLPS*, *PNLIP*, *REG1A*, *CPB1*) however, more work is needed to prove that the interaction is specific.

Limitations

Conditions *in vivo* may well allow interactions to take place that are technically complex or impractical to reproduce in the laboratory. The most challenging problem we have encountered is concerned with creating compatible conditions for solubilising WF-partner pairs for further analysis. Secondly, an issue that may in fact be related to solubility is the

need for further optimisation of Co-IP protocols. It has yet to be determined whether the observed non-specific pull-down during Co-IP results from improper folding of secretory proteins.

Alternative *in vitro* methods provide options for dealing with problems of protein solubility and purification although techniques such as FRET analysis may allow us to tackle both problems together, whilst providing a credible and biologically relevant means of confirming our proposed interactions *in vivo*.

Conclusions

Our findings suggest that wolframin may function in some capacity in the production or activity of one or more digestive enzymes and exocrine secretory proteins. The preliminary data show that the N-terminal domain of wolframin, sitting on the cytoplasmic side of the endoplasmic reticulum (ER), may interact with a number of previously unidentified pancreatic protein-partners.

It may be that possible wolframin interaction with our identified proteins could take place in regions other than the exocrine pancreas, perhaps in the gut or in the case of *REG1A*, in regenerating β-cells, which opens an interesting link with WS pathology. Indeed if interactions do occur endogenously in human systems, the regions of these proteins associated with wolframin are also important to understand. For example, pancreatic carboxypeptidases (CBP) do share 15-20% sequence homology with other CBP family members (Skidgel, 1988), which could hypothetically share a region of binding affinity to wolframin. This possibility could open a whole array of novel functions for wolframin.

In the case of *CLPS*, we would be interested to learn whether wolframin is able to interact with amino acids 18-22, which constitute the enterostatin pentapeptide, found in the brain as well as the digestive system. Of particular interest is the role of enterostatin in regulation of insulin secretion and whether a potential interaction of this peptide or its precursor *CLPS,* with wolframin has any significance.

Although wolframin expression is shown to be primarily localised to the ER membrane (Osman, 2003; Philbrook, 2005), a number of studies provide evidence that it may also engage in activity with plasma membrane proteins (Zatyka, 2008a) and is expressed in β-cell secretory granules (Hatanaka, 2011; Ishihara, 2004). Defining the sub-cellular location and understanding the context of these interactions is therefore difficult and the requirement of other associate-proteins or bridging molecules for complex formations has not been ruled out.

The likely nature of WF-interaction with our proposed candidates is therefore open to interpretation and further investigation, although we speculate that wolframin could be involved in the folding and processing of these proteins, since earlier evidence proposes wolframin as a molecular chaperone (Fonseca, 2005).

Our primary objective is to validate our preliminary findings from the yeast screen to demonstrate one or more of our candidates as an authentic wolframin molecular partner. We intend to incorporate any such novel findings in further functional study of *WFS1* variants (see chapter 6).

CHAPTER 6

GENERAL DISCUSSION, CONCLUSIONS AND FURTHER WORK

6.0 Thesis summary

The work presented in this thesis further highlights the genetic and phenotypic heterogeneity of WS, by initially expanding the mutational spectrum of WS patients. A number of individuals in our cohort were found to carry previously undescribed *WFS1* mutations, and together with the listed patients registered on our West Midlands database, comprise the largest single UK cohort to date. The first WS patient in the UK populace to harbour mutation of *WFS2* is also reported. Our data suggest this mutation (*WFS2*-E37Q) originates from a founder of Jordanian-Bedouin background, as all individuals found to carry this mutation, including our own, appear to share a common ancestor.

Possible causative mutations were not found in a number of WS patients included in our study despite full direct sequencing of all coding regions of *WFS1* and *WFS2.* This may be a consequence of undetected chromosomal abnormalities, large deletions, or important non-coding *WFS1* variants, or may alternatively suggest the existence of additional WS disease-loci. The phenotypic variability associated with WS complicates interpretation, as a number of other disorders share clinical characteristics with WS. This is underlined in our study by the discovery of a homozygous *SLC19A2* mutation (R162X) in one of our patients, who only presented with anaemia following a diagnosis of WS. We were therefore able to confirm a

genetic diagnosis of TRMA. Clinical profiles of remaining patients should perhaps be re-evaluated.

One or more *WFS1* mutations have been found in 67 of the listed patients on the Birmingham WS patient database. Of 121 identified mutations, approximately equal proportions (31-33%) of missense, nonsense and deletion mutation-types were found in *WFS1*. The remainder was accounted for by insertion, splice and synonymous changes.

All available clinical information compiled from a total of 42 patients revealed median ages of diabetes mellitus and optic atrophy onset of 6 years and 10 years respectively. Further genotype-phenotype correlation performed on patients grouped according to mutation-types, showed that individuals carrying inactivating *WFS1* mutations present with diabetes mellitus at an earlier median age (5 years) than those with at least one missense mutation (median: 10.5 years). A greater proportion of patients with inactivating mutations also appear to develop a more complete spectrum of WS-associated complications. Our findings are consistent with those of earlier studies (Smith, 2004; Cano, 2007; D'Annunzio, 2008; Zmyslowska, 2011) and suggest that patients carrying inactivating mutations are predisposed to develop a relatively more severe phenotype than those with missense mutations. In support of these observations are the reports of 'mild' cases of WS where patients develop a markedly later onset of clinical features (Van Den Ouweland, 2003; Cano, 2007; Gasparin, 2009). These patients have always harboured *WFS1*-missense mutations.

Single nucleotide variants in *WFS1* range from those considered as common benign polymorphisms to those associated with disease, including WS of variable severity, type 2 diabetes (Voight, 2010), non-syndromic deafness (Young, 2001; Bespalova, 2001), and psychiatric illness (Furlong, 1999; Torres, 2001; Swift, 2005; Zalsman, 2009). Perhaps

previously uncharacterised variants of *WFS1* are responsible, at least in part, for a number of other disorders of the visual, auditory or nervous systems and the clinical manifestations associated with this gene have a wider spectrum than previously thought? As more information comes to light, we may find a clearer relationship between genotype and phenotype and an understanding of wolframin function in general terms or specific to various biological systems.

We hoped to contribute to this end by functional analyses of selected wolframin variants, and intended to focus our efforts from the perspective of diabetogenesis, both an integral part of the WS phenotype and a prevalent, complex disease of an indeterminate molecular basis.

A selection of representative wolframin variants were characterised in mammalian over-expression systems and we provided the first evidence that correctly localised, stably expressed wolframin protein may be produced by *WFS1* harbouring some WS-associated mutations. The first molecular characterisation of the type 2 diabetes associated allele *WFS1*-H611 is also presented, laying the ground-work for further functional analyses. A new perspective on the effects of *WFS1* mutation was taken by comparing missense variants relating to phenotypes of variable severity. Our data, together with knowledge of variability in WS clinical presentation, support the hypothesis that some *WFS1* mutations may allow production of partially functional wolframin protein. We intend to test this idea further by assessing the relative ability of wild-type and mutant wolframin to function in a series of assays.

To broaden functional understanding of wolframin and enhance the scope of our prospective plans we have identified several new potential molecular partners of wolframin

in the pancreas, which have yet to be verified. A number of interesting candidates are described, all of which are components of exocrine pancreatic secretions, mostly associated with digestion, although some have other additional systemic functions perhaps relating to wolframin pathways. Our evidence that *WFS1* is expressed in exocrine cells raises the possibility for novel wolframin interactions to take place in acinar cells, potentially mediating synthesis or secretion of digestive enzymes.

The wide spectrum of reported *WFS1* mutations associated with WS by definition all result in presentation of diabetes mellitus and optic atrophy. What differs between affected individuals is the relative age of disease onset and the number of additional clinical features that present. We do not yet understand whether variability in disease severity and characteristics is solely dependent on background genetics, or whether the type of mutations involved play a large part in prognosis.

WS-associated diabetes is thought to be caused by ER stress-mediated apoptosis of pancreatic β-cells and mounting evidence suggests other WS clinical features may arise due to similar mechanisms (Yoshida, 2007). Unsurprisingly wolframin seems to have an important role in regulation of the UPR, and when this function is impaired an uncontrolled stress response results, ultimately culminating in execution of apoptosis.

Many specific links have been shown between wolframin and various components of stress response pathways (Yamada, 2005; Ueda, 2005; Fonseca, 2010), although it is unclear precisely how *WFS1* mutation disrupts the normal UPR. Complete absence of wolframin would clearly abolish all related molecular interactions, but the question is whether all or only some of these interactions are critical for cell survival and adaptation to stress conditions? If some missense mutations confer at least a degree of function, it is possible

that depending on genotype, only specific molecular interactions and downstream cascades are negatively affected. Could this partly explain the variation in WS phenotype, where not all wolframin pathways are disrupted allowing cells to cope with stress conditions for longer periods before the system breaks-down?

A lack of wolframin protein is the likely cause of ER stress in patients carrying nonsense or frameshift mutations as these types of gene aberrations are known to result in unstable products. Previous studies suggest that the effects of missense mutations are less clear. Absence or depleted expression of wolframin has been documented in patient fibroblasts carrying selected *WFS1* mutations (Philbrook, 2005). *In vitro* experiments have also shown that some *WFS1* mutations may cause reduced steady-state protein expression as a result of instability (Hofmann, 2006), and provide evidence that some mutants may be susceptible to aggregation (Fonseca, 2005). If the effects of different mutations are variable on a molecular level an important question is raised: do *WFS1* mutations directly contribute to ER stress, reduce the capability of a cell to effectively deal with stress conditions, or both? These outcomes may be dependent on the mutation-type.

The majority of research efforts are focused on *WFS1* knock-out or knock-down models, which draw parallels to situations of partial or complete deficiency in wolframin. This approach is fundamental for understanding what wolframin actually does by allowing a view of the system as a whole where wolframin function is totally abolished. Since *WFS1* missense mutations present not only in 31% of Wolfram syndrome patients (see section 3.7.0) but in several other disease states, it might also be prudent to view the disruption of wolframin function from a genotypic perspective, which may shed light on the phenotypic heterogeneity of WS and the role of *WFS1* in specific physiological systems, such as the

auditory or neuroendocrine systems. Although some functional studies have included a select few wolframin variants (Osman, 2003; Fonseca, 2005; Hofmann, 2006), they do not cover an extensive range of mutations, particularly missense changes, and also fail to correlate the functional data with associated physio-chemical characteristics and phenotypic effects of the mutation in question.

Genetic (locus/allelic) heterogeneity in WS suggests that one of many perturbations in the complex system of cellular stress management results in the same eventual outcome, an inability to adapt to stress conditions resulting in cell-death. Although uncontrolled apoptotic cell-death is the common molecular feature of WS, the patho-mechanisms leading to this break-down and presentation of parallel clinical features are clearly different between *WFS1* and *WFS2* associated WS. It seems then very reasonable to suggest that different mutations in *WFS1* could disrupt the function of wolframin in a number of ways with the same eventual outcome. If selected mutations do allow partial protein function, rather than abolishing protein function entirely, only specific molecular interactions may be interrupted, broadening the scope for therapeutic intervention in these cases.

6.1.0 Future directions

For comparative wild-type and mutant wolframin functional analyses an appropriate model system must firstly be developed. We intend to focus on cell types primarily affected in WS such as β-cells, neuronal cells, or retinal cells and compare wild-type with those carrying nonsense or missense *WFS1* mutations. Producing cell lines with suitable genotypes is the important first stage of this process, and can be accomplished in a number of ways. The

overall aim is to replace wild-type *WFS1* with variants of interest and is routinely accomplished by either temporarily or permanently 'turning-off' wild-type *WFS1*, then replacing this gene with an alternative copy harbouring a mutation.

6.1.1 Wolframin-depletion

Transient or stable *WFS1* knock-down can be achieved *in vitro* with the application of siRNA technology. Alternatively, whole animal or organ specific *WFS1* knock-out cells can be derived from selectively bred and genetically modified animals. Mutant copies of *WFS1* can then be reintroduced by means of transfection or infection with lentiviral or adenoviral delivery vectors.

RNA interference (RNAi) is a natural mechanism of gene silencing conserved in plant and mammalian cells, thought to play an important role in battling viral infection, fine tuning normal cellular gene expression and suppressing the movement of mobile genetic elements within the genome (McManus, 2002). The protein machinery comprising the RNAi pathway recognises double-stranded RNA and degrades it into smaller segments using an enzyme called *dicer*. These segments, known as short interfering RNAs (siRNA) are typically 21-23 nucleotides in length and are subsequently used in an enzyme complex (RISC) as templates to prevent expression of complementary sequences of single-stranded RNA (Ketting, 2001).

These can be engineered in the laboratory to complement specific gene sequences, and used to transiently induce gene knock-down following transfection into living cells (Fire, 1998). They are particularly useful as a research tool because they are highly specific, versatile and efficient. To increase the short half-life of synthetic RNAi, eukaryotic polymerase III promoters can be employed to drive expression of siRNA from expression

cassettes, to produce small hairpin RNA (shRNA) structures (Paddison, 2002). This method allows more 'stable' gene silencing as the vectors are usually passed on to daughter cells.

The greatest challenge for RNAi-based technology, is delivering these molecules to the cytoplasm of target cells as unmodified RNA cannot penetrate lipid plasma membranes. A number of techniques, each with their distinct advantages and limitations, are used to overcome this problem including: viral and bacterial delivery, liposomal and nanoparticle packaging, high-pressure injection and chemical modification.

Stable knock-down of *WFS1* in both mouse pancreatic (MIN6) and human neuroblastoma (SKNAS) cells are currently under development in the Birmingham laboratory. A selective murine breeding programme is also in place to create lines with conditional knock-out of *WFS1* in pancreatic β-cells using a *Cre-Lox* recombination system. *LoxP* sequences contain specific binding sites for *Cre*, a DNA recombinase enzyme, which serves to splice specific target genes *in vivo*. A recombination event can take place in the cells of transgenic animals containing *LoxP* sites, when they also express *Cre*. Inversion, deletion or translocation of DNA segments may occur depending on the orientation of the *LoxP* sites. Two mouse lines are used to achieve targeted deletion of *WFS1*, one that is genetically modified to contain 'floxed *WFS1*', where two *LoxP* sites flank a section of the gene, and another expressing *Cre* recombinase. Hybrid descendants containing both *LoxP* sites and the *Cre* recombinase consequently have gene inactivation.

For future use in these murine model systems, a selection of mutations corresponding to those we have characterised in human cells have been introduced in the mouse *wfs1* gene using site-directed mutagenesis (appendix 6).

6.1.2 Testing the model

Once a suitable level of knock-down can be confirmed, the effects of such a reduction in gene expression must be demonstrable on the phenotypic level to provide a negative control against which to measure 'normal' unaltered systems and experimental samples. Furthermore the successful delivery of replacement gene copies must also be tested and verified with wild-type genes before considering the consequences of introducing 'mutant' copies.

Depletion of *WFS1* is associated with increased ER stress and apoptosis (Fonseca, 2005; Riggs, 2005; Yamada, 2006), perturbation of calcium homeostasis (Takei, 2006; Zatyka, 2008b) and decreases in glucose stimulated insulin secretion (Zatyka, 2008b). Confirming the suitability of knock-down models can be tested by comparatively measuring these pathways in wild-type and *WFS1*-depleted cells. Once the effects of depletion have been verified, 'rescue' experiments can be performed to reintroduce wild-type *WFS1* into depleted cells as a positive control to confirm successful delivery of the gene and restoration of a normal phenotype. An established knock-down model with all the necessary controls completed will then be taken forward to the experimental stage where the mutated versions of *WFS1* can be re-introduced into depleted cells and the phenotype can be observed.

6.1.3 Wolframin functional assays

A top down approach could be taken for examining the overall functional effects selected *WFS1* missense variants have on ER stress and apoptosis, before considering more specific pathways in detail.

ER stress and apoptosis

Comparing efficiency of the UPR between wild-type cells and those harbouring mutations can be achieved by treating cells with chemical stress inducers, then probing for biomarkers of ER stress such as upregulation of UPR target genes *XBP1*, *BiP*, and *CHOP* (Samali, 2010). A number of reagents can be used to disrupt parts of the protein metabolism pathway causing accumulation of unfolded proteins; A23187 and thapsigargin disturb calcium homeostasis, tunicamycin suppresses glycosylation, and brefeldin A inhibits transport of proteins to the golgi complex (Ding, 2007). The relative survival rates of normal and mutant cells under conditions of cellular stress would also reveal information about the ability of these cells to cope with, or resolve such stimulus before making a 'switch' to apoptotic pathways.

Following up on previous evidence we could also examine how *WFS1* mutations affect distinct pathways known to be perturbed by loss of wolframin function such as insulin secretion and calcium signalling/homeostasis (Ishihara, 2004; Zatyka, 2008b). We anticipate some possible limitations in *WFS1*-H611 function in relation to insulin secretion pathways and hope to shed some light on the role of this allele in type 2 diabetes susceptibility.

Chaperone activity

Several lines of evidence suggest wolframin may act as a chemical chaperone (Hofmann, 2003; Zatyka, 2008b). This function can be tested using unique biosensors that enable measurement of chaperone activity, such as cdYFP (chaperone-dependent yellow fluorescent protein). CdYFP is a chromophore with low fluorescence intensity as a result of a single point-mutation that prevents the correct folding of the protein (Liman, 2005). This construct has been developed in such a way that in the presence of chaperone activity, correction to the folding defects enables an increase in fluorescent signal intensity. Applying this technology to the study of wolframin may provide some solid evidence about the ability of this protein to assist in protein folding, and could provide some interesting clues about the role of wolframin in the UPR. If chaperone activity can be demonstrated, we may take this line of study further by examining the relative ability of our WF-mutants to perform this function.

The effects of *WFS1* mutation on protein interactions

Testing the ability of selected wolframin variants to interact with known molecular partners may enable us to pinpoint a more clearly defined area hampered by loss of function.

ATF6α

A recent report detailed the negative regulation of ATF6α by wolframin (Fonseca, 2010). Wolframin was shown to stabilise HRD1, an E3 ligase for ATF6α and enhance the ubiquitination and proteasomal degradation of ATF6α. In stress conditions ATF6α is thought to be released from wolframin to initiate the UPR. Some *WFS1* mutations (G695V and

P742L) previously shown to induce insoluble protein aggregates in the ER, result in loss of interaction with HRD1. Consequently ATF6α is hyper-activated leading to an uncontrolled UPR.

Our data have shown that not all WS-associated missense mutations result in mislocalisation and aggregation of wolframin, even those corresponding to a severe phenotype (*WFS1*-G736A). We therefore propose that the loss of HRD1 binding and consequent effects on the UPR may also differ, dependent on mutation-type. We would like to investigate this possibility by looking more closely at ATF6α regulatory pathways in conjunction with our panel of selected variants to further define the importance of this link in WS pathogenesis. To this end, production of ATF6α expression constructs are currently in progress.

The ability of wolframin variants to bind with other known partner-proteins can also be explored. We would like to concentrate our efforts primarily on interaction with the sodium-potassium ATPase β1 subunit, which was first demonstrated in the Birmingham laboratory (Zatyka, 2008a). Following confirmation of our proposed molecular partners (chapter 5) we may also have one or more novel interactors to include in co-immunoprecipitation experiments with *WFS1* mutants.

6.2.0 Conclusions

The heterogenic nature of WS can make accurate diagnosis difficult. This is exemplified in our study by the discovery of TRMA-associated mutations (in the *SLC19A2* gene) in one of our probands who was initially thought to have Wolfram syndrome.

Mutational analysis of WS patients and compilation of clinical data continues to widen the spectrum of disease-associated *WFS1* mutations and phenotypic variability. We have contributed to this end by expanding the mutational spectrum of WS with a number of novel *WFS1* mutations and reporting the first case of *WFS2*-associated WS in a UK cohort. Larger data sets enable us to perform more detailed analysis so that we can uncover relationships between genetic anomalies and clinical manifestation. Our data concur with previous reported correlations between WS genotype and phenotype (Smith, 2004; Cano, 2007; D'Annunzio, 2008) as we found that patients harbouring inactivating *WFS1* mutations appeared susceptible to a relatively more severe phenotype than those with at least one missense mutation. In our cohort, patients with inactivating mutations presented with diabetes mellitus at an earlier median age and a higher proportion of these individuals also developed additional clinical features compared to patients carrying missense mutations.

Several previous studies have reported 'mild' cases of WS where individuals presented with a relatively late onset of disease (Van Den Ouweland, 2003; Cano, 2007; Gasparin, 2009). In each of these cases, the affected individuals carried at least one missense mutation, which suggests that perhaps some *WFS1* mutations allow partial protein function (wolframin). Our data support this hypothesis by providing evidence that some *WFS1* missense mutations allow expression of *WFS1* RNA and stable wolframin protein, which is localised correctly to the endoplasmic reticulum membrane. This provides the basis and rationale for further investigation of the ability of these 'mutated' proteins to function in various capacities known to be disrupted by depletion of *WFS1* such as glucose-stimulated insulin secretion, calcium homeostasis and mediation of the UPR (Ishihara, 2004; Zatyka, 2008b; Fonseca,

2005). Selected variants with residual function may potentially be enhanced or restored by chemical chaperones, offering new perspectives for pharmacologic intervention.

The inclusion and molecular characterisation of the *WFS1*-H611 allele, associated with type 2 diabetes mellitus, also provides a foundation for studying the role of this variant in a more prevalent disease. To our knowledge, no functional analyses have yet been conducted on this diabetes 'risk-gene', therefore we intend to investigate the impact this *WFS1* allele has on glucose-stimulated insulin secretion and the ability of pancreatic β-cells to effectively cope with ER-stress conditions. We hope to shed some light on this aspect of type 2 diabetes pathogenesis and β-cell function.

Any verifiable wolframin partner-proteins from our panel of candidates will also be included in our future work. Our evidence that *WFS1* is expressed in the exocrine pancreas, together with identification of a number of pancreatic digestive enzymes as potential novel wolframin molecular-partners gives us grounds to examine the function of wolframin from a new perspective. Given the localisation to insulin secretory granules (Hatanaka, 2011), we speculate that wolframin may also have a role in the exocrine pancreas, possibly in the exocytotic secretory pathway. Possible association between wolframin and lipase/colipase perhaps suggests that wolframin could be an important link between lipid toxicity (hyperlipidemia) and type 2 diabetes mellitus?

As a clearer picture of WS natural history emerges we may be able to relate genetic mutations more effectively to phenotypic outcomes enabling more accurate prognoses to be made, perhaps opening up the catchment for identifying affected individuals who may not present with the classical complement of WS features. The increasingly diverse range of clinical manifestations associated with *WFS1* mutations raises an important question about

the stringency of current diagnostic criteria, principally juvenile-onset diabetes mellitus and optic atrophy.

Our study of specific missense mutation of *WFS1* relating to phenotype provides a complementary and novel approach to the classical study of wolframin mutation, which has previously been thought to cause loss of function. We contend that WS may be a disease characterised by loss of protein function to different degrees, which could in part explain the extreme phenotypic variability associated with mutation of this gene.

REFERENCES

ADA (American Diabetes Association), (2010). Diagnosis and classification of Diabetes Mellitus, *Diabetes Care*, **33** (1); S62-S69.

ADA (American Diabetes Association), (2004). Diagnosis and classification of diabetes mellitus, Diabetes Care, **27**: S5-S10.

Alstrom, C.H., *et al* (1959). A specific syndrome (not hitherto described) distinct from Laurence-Moon-Beidl syndrome. A clinical endocrinological and genetic examination based on a large pedigree, *Acta Psychiactica et Neurologica Scandanavica*, **43**(suppl 129): 1-35.

Amr S., *et al* (2007). A homozygous mutation in a novel zinc-finger protein, ERIS, is responsible for Wolfram syndrome 2, *The American Journal of Human Genetics*, **81**: 673-83.

Amr, S. (2010). 'Identification and characterisation of a second Wolfram syndrome gene, PhD thesis, Virginia Coomonwealth University.

Aneesh, T.P., *et al* (2009). Pharmacogenomics: The right drug to the right person, *Journal of Clinical Medicine Research*, **1**(4): 191-94.

Ashcroft, F.J., *et al* (2004). Control of expression of the lectin-like protein Reg-1 by gastrin: role of the Rho family GTPase RhoA and C-rich promoter element, *The Journal of Biochemistry*, **381**: 397-403.

Awata, T., et al (2000). Missense variations of the gene responsible for Wolfram syndrome (WFS1/wolframin) in Japanese: possible contribution of the Arg456His mutation to type 1 diabetes as a nonautoimmune genetic basis, *Biochemical and Biophysical Research Communications*, **268**(2): 612-6.

Barrett T.G., *et al* (1995). Neurodegeneration and diabetes: UK nationwide study of Wolfram (DIDMOAD) syndrome, *The Lancet*, **346**: 1458-63.

Barrientos A., *et al* (1996a). Wolfram syndrome associated with an 8.5-kb mtDNA single deletion, *American Journal of Human Genetics*, **58**: 963-70.

Barrientos, A., *et al* (1996b). A nuclear defect in the 4p16 region predisposes to multiple mitochondrial DNA deletions in families with Wolfram syndrome, *Journal of Clinical Investigation*, **97**(7): 1570-76.

Beck, I.T., (1973). The role of pancreatic enzymes in digestion, *American Journal of Clinical Nutrition*, **26**: 311-25.

Bespalova, I.N., *et al* (2001). Mutations in the Wolfram syndrome 1 gene (*WFS1*) are a common cause of low frequency sensorineural hearing loss, *Human Molecular Genetics*, **10** (22): 2501-08.

Birnboim, H.C., and Doly, J., 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA, *Nucleic Acids Research*, **7**(6): 1513-23.

Blackinton, J., et al (2005). Effects of DJ-1 mutations and polymorphisms on protein stability and subcellular localisation, Molecular Brain Research, **134**(1): 76-83.

Bram, R.J., *et al* (1986). A GAL family of upstream activating sequences in yeast: roles in both induction and repression of transcription, *EMBO*, **5**(3): 603-8.

Butler, A.E., *et al* (2003). β-cell deficit and increased β-cell apoptosis in humans with type 2 diabetes, *Diabetes*, **52**: 102-10.

Cano, A., *et al* (2007). Identifications of novel mutations in WFS1 and genotype-phenotype correlation in Wolfram syndrome, *American Journal of Medical Genetics*, Part A **143A**: 1605-12.

Cardozo, A.K., *et al* (2001). Identification of novel cytokine-induced genes in pancreatic β-cells by high-density oligonucleotide arrays, *Diabetes*, **50**: 909-20.

Cardozo, A.K., *et al* (2005). Cytokines downregulate the sarcoendoplasmic reticulum pump Ca^{2+} ATPase 2b and deplete endoplasmic reticulum Ca^{2+}, leading to induction of endoplasmic reticulum stress in pancreatic β-cells, *Diabetes*, **54**: 452-61.

Cavard, C., *et al* (2006). Overexpression of regenerating islet-derived 1 alpha and 3 alpha genes in human primary liver tumors with B-catenin mutations, *Oncogene*, **25**: 599-608.

CDC (Centers for Disease Control and Prevention) (2010). Diabetes public health resource, [online], available at: <www.cdc.gov/diabetes> [Accessed: Aug, 2012].

Chaussenot, A., *et al* (2011). Neurologic features and genotype-phenotype correlation in Wolfram syndrome, *Annals of Neurology*, **69**(3): 501-8.

Chen, Y., *et al* (2009). Cisd2 deficiency drives premature aging and causes mitochondria-mediated defects in mice, *Genes and Development*, **23**: 1183-94.

Chen, X.I., et al (2002). The luminal domain of ATF6 senses endoplasmic reticulum stress and causes translocation of ATF6 from the ER to the golgi, *Journal of Biological Chemistry*, **277**(15): 13045-52.

Cnop, M., *et al* (2007). Selective inhibition of eukaryotic translation initiation factor 2α dephosphorylation potentiates fatty acid-induced endoplasmic reticulum stress and causes pancreatic b-cell dysfunction and apoptosis, *Journal of Biological Chemistry*, **282**: 3989-97.

Colca, J.R., *et al* (2004). Identification of a novel mitochondrial protein ('mitoNEET') cross-linked specifically by a thiazolidinedione photoprobe, *American Journal of Physiology, Endocrinology and Metabolism*, **286**: E252-60.

Collier, D.A., *et al* (1996). Linkage of Wolfram syndrome to chromosome 4p16.1 and evidence for heterogeneity, *American Journal of Human Genetics*, **59**: 855-63.

Colosimo, A., *et al* (2003). Molecular detection of novel WFS1 mutations in patients with Wolfram syndrome by a DHPLC-based assay, *Human Mutation*, **21**: 622-9.

Crenon, I., *et al* (1994). Molecular cloning and expression of two horse pancreatic cDNA encoding colipase A and B, *Biochemica et Biophysica Acta*, **1213**: 357-60.

D'Annunzio, G., *et al* (2008). Wolfram syndrome (diabetes insipidus, diabetes mellitus, optic atrophy and deafness) clinical and genetic study, *Diabetes Care*, **31** (9): 1743-5.

Davidson, M.W., (2010). Molecular Expressions, Cell biology and microscopy, Structure and function of cells and viruses, Florida state university, [online], available at:

<http://micro.magnet.fsu.edu/cells/endoplasmicreticulum/endoplasmicreticulum.html> [Accessed September 2011]

De Caro A., *et al* (1979). Characterization of a protein isolated from men suffering from chronic calcifying pancreatitis, *Biochemical and Biophysical Research Communications*, **87**(4): 1176-82.

Di Lorenzo, C., *et al* (2007). The wolframin His611Arg polymorphism influences medication overuse headache, *Neuroscience Letters*, **424**: 179-84.

Diabetes UK (2010). [online], available at: <www.diabetes.org.uk> [Accessed March 2011].

Dieckgraefe, B.K., *et al* (2000). Analysis of mucosal gene expression in inflammatory bowel disease by parallel oligonucleotide arrays, *Physiological Genomics*, **4**: 1-11.

Ding, W, *et al* (2007). Differential effects of endoplasmic reticulum stress-induced autophagy on cell survival, *The Journal of Biological Chemistry*, **282**: 4702-10.

Domenech, E., et al (2004). Study of the *WFS1* gene and mitochondrial DNA in Spanish Wolfram syndrome families, *Clinical Genetics*, **65**: 463-69.

El-Shanti, H., *et al* (2000). Homozygosity mapping identifies an additional locus for Wolfram syndrome on chromosome 4q, *American Journal of Human Genetics*, **66**: 1229-36.

Eizirik, D.L., *et al* (1992). Interleukin-1β induces the expression of an isoform of nitric oxide synthase in insulin-producing cells, which is similar to that observed in activated macrophages, *FEBS Lett*, **308**: 248-52.

El-Shanti H., *et al* (2000). Homozygosity mapping identifies an additional locus for Wolfram syndrome on chromosome 4q, *American Journal of Human Genetics*, **66**: 1229-36.

Emi, M., *et al* (1986). Cloning, characterization and nucleotide sequences of two cDNAs encoding human pancreatic trypsinogens, *Gene*, **41**(2-3): 305-10.

Ensembl Genome Browser (2012). [online], available at: <www.ensemble.org> [Accessed December 2011].

Erlanson-Albertson, C., *et al* (1997). Pancreatic procolipase propeptide, enterostatin, specifically inhibits fat intake, *Physiology and Behaviour*, **49**: 1191-4.

Espino-Paisan, L., et al (2011). Early and Late Onset Type 1 Diabetes: One and the Same or Two Distinct Genetic Entities?, Type 1 diabetes complications:29-54.

Fajans, S.S., *et al* (2010). MODY, History, genetics, pathophysiology, and clinical decision making, *Diabetes Care*, **34**(8): 1878-84.

Fawcett, K.A., *et al* (2010). Detailed investigation of the role of common low-frequency *WFS1* variants in type 2 diabetes risk, *Diabetes*, **59** (3): 741-6.

Fire, A., *et al* (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, *Nature*, **391**: 806-11.

Florez, J.C., *et al* (2008). Testing of diabetes-associated *WFS1* polymorphisms in the Diabetes Prevention Program, *Diabetologia*, **51**(3): 451-57.

Fonseca, S. *et al.* (2005). WFS1 is a novel component of the unfolded protein response and maintains homeostasis of the endoplasmic reticulum in pancreatic β-cells, *The Journal of Biological Chemistry*, **280** (47): 39609-15.

Fonseca, S., *et al* (2010). Wolfram syndrome 1 gene negatively regulates ER stress signalling in rodent and human cells, *The Journal of Clinical Investigation*, **120** (3): 744-55.

Fowler, M.J., (2008). Microvascular and macrovascular complications of diabetes, *Clinical Diabetes*, **26**(2): 77-82.

Franks, P.W., *et al* (2008). Replication of the association between variants in *WFS1* and risk of type 2 diabetes in European populations, *Diabetologia*, **51** (3): 458-63.

Fricker, L.D., (1988). Carboxypeptidase E, *Annual Review of Physiology*, **50**: 309-21.

Fujikawa, T., *et al* (2010). Additional heterozygous 2507A>C mutation of *WFS1* in progressive hearing loss at lower frequencies, *Laryngoscope*, **120**: 166-71.

Furlong, R.A., *et al* (1999). A rare coding variant within the *wolframin* gene in bipolar and unipolar affective disorder cases, *Neuroscience Letters*, **277**: 123-26.

Gasparin, M.R., *et al* (2009). Identification of novel mutations of the *WFS1* gene in Brazilian patients with Wolfram syndrome, *European Journal of Endocrinology*, **160**: 309-16.

Gomez-Zaera, M., *et al* (2001). Presence of major *WFS1* mutation in Spanish Wolfram syndrome pedigrees, *Molecular Genetics and Metabolism*, **72**: 72-81.

Guo, X., *et al* (2011). The E3 ligase Smurf1 regulates Wolfram syndrome protein stability at the endoplasmic reticulum, **286**(20): 18037-47.

Harding H.P., *et al* (2000). Regulated translation initiation controls stress-induced gene expression in mammalian cells, *Mol Cell*, **6**: 1099-1108.

Hardy, C., *et al* (1999). Clinical and molecular genetic analysis of 19 Wolfram syndrome kindreds demonstrating a wide spectrum of mutations in WFS1, *American Journal of Human Genetics*, **65** (5): 1279-90.

Hasan, M.A., *et al* (2000). Wolfram's (DIDMOAD) syndrome and chronic renal failure, *Saudi Journal of Kidney Disease and Transplantation*, **11** (1): 53-58.

Hatanaka, M., *et al* (2011). Wolfram syndrome 1 gene (WFS1) product localizes to secretory granules and determines granule acidification in pancreatic beta-cells, *Human Molecular Genetics*, **20**(7): 1274-84.

Hildebrand, M.S., *et al* (2008). Autoimmune disease in a DFNA6/14/38 family carrying a novel missense mutation in WFS1, *American Journal of Medical Genetics*, **146A**(17): 2258-65.

Heiman, M., (1997). Webcutter 2.0 sequence analysis server. [online] Available at: *bio.lundberg.gu.se/cutter2.*

Hofmann, S., et al (2003). Wolfram syndrome: structural and functional analyses of mutant and wild-type wolframin, the *WFS1* gene product, *Human Molecular Genetics*, **12** (16): 2003-12.

Hofmann, S., and Bauer, M.F. (2006). Wolfram syndrome-associated mutations lead to instability and proteasomal degradation of wolframin, *FEBS Letters*, **580**: 4000-04.

Hollien, J., *et al* (2009). Regulated Ire1-dependent decay of messenger RNAs in mammalian cells, *Journal of Biological Chemistry*, **186**(3): 323-31.

Hosszufalusi, N., *et al* (2003). Similar genetic features and different islet cell autoantibody pattern of latent autoimmune diabetes in adults (LADA) compared with adult-onset type 1 diabetes with rapid progression, *Diabetes Care*, **26**: 452-57.

Inoue, H., *et al* (1998). A gene encoding a transmembrane protein is mutated in patients with diabetes mellitus and optic atrophy (Wolfram syndrome), Nature Genetics, **20**, 143 – 48.

Ishihara, H., *et al* (2004). Disruption of the *WFS1* gene in mice causes progressive B-cell loss and impaired stimulus-secretion coupling in insulin secretion, *Human Molecular Genetics*, **13**(11): 1159-70.

Jackson, M.J., *et al* (1994). Biochemical and molecular studies of mitochondrial function in diabetes insipidus, diabetes mellitus, optic atrophy and deafness, *Diabetes Care*, **17**(7): 728-33.

Jeffrey, K.D., *et al* (2008). Carboxypeptidase E mediates palmitate-induced beta-cell ER stress and apoptosis, *Proceedings of the National Academy of Sciences in the United States of America*, **105**: 8457-7.

Jörns, A., *et al* (2005). Immune cell infiltration, cytokine expression, and β-cell apoptosis during the development of type 1 diabetes in the spontaneously diabetic LEW.1AR1/Ztm-iddm rat, *Diabetes*, **54**: 2041-52.

Kakiuchi, C., *et al* (2007). Association of HSP90B1 with bipolar disorder, *Journal of Human Genetics*, **52**(10): 794-803.

Kakiuchi, C., et al (2009). Valproate, a mood stabilizer, induces WFS1 expression and modulates its interaction with ER stress protein GRP94, *PLoS ONE*, **4**(1): e4134.

Kang, S.W., *et al* (2006). Substrate-specific translocation attenuation during ER stress defines a pre-emptive quality control pathway, *Cell*, **127**: 999-1013.

Karaskov, E., *et al* (2006). Chronic palmitate but not oleate exposure induces endoplasmic reticulum stress, which may contribute to INS-1 pancreatic β-cell apoptosis, *Endocrinology*, **147**: 3398-3407.

Kawano, J., *et al* (2005). Wolfram syndrome 1(*Wfs1*) gene expression in the normal mouse visual system, *The Journal of Comparative Neurology*, **510** (1): 1-23.

Ketting, R.F., *et al* (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*, *Genes and Development*, **15**: 2654-59.

Khanim, F., *et al* (2001). WFS1/wolframin mutations, Wolfram syndrome, and associated diseases, *Human Mutation*, **17**:357–67.

Kharroubi, I., *et al* (2004). Free fatty acids and cytokines induce pancreatic beta-cell apoptosis by different mechanisms: role of nuclear factor-kappaB and endoplasmic reticulum stress, *Endocrinology*, **145**: 5087-96.

King, D., 2009, Histology at SIU .com southern Illinois University School of Medicine, [online], available at: <www.siumed.edu/~dking2/index.htm> [Accessed October 2011].

King, M.P., *et al* (1992). Defects in mitochondrial protein synthesis and respiratory chain activity segregate with the tRNA (Leu (UUR)) mutation associated with mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes, *Molecular and Cellular Biology*, **12**: 480-90.

Koido, K., *et al* (2005). Polymorphisms in wolframin gene (*WFS1*) are possibly related to increased risk for mood disorders, *International Journal of Neuropsychopharmacology*, **8**: 235-44.

Krawczak, M., *et al* (2007). Single base-pair substitutions in exon-intron junctions of human genes: Nature, distribution and consequences for mRNA splicing, *Human Mutation*, **28** (2): 150-58.

Kudo, T., *et al* (2008). A molecular chaperone inducer protects neurones from ER stress, *Cell Death and Differentiation*, **15**(2): 364-75.

Kutlu, B., *et al* (2003). Discovery of gene networks regulating cytokine-induced dysfunction and apoptosis in insulin-producing INS-1 cells, *Diabetes*, **52**: 2701-19.

Labay, V., *et al* (1999). Mutations in SLC19A2 cause thiamine-responsive megaloblastic anaemia associated with diabetes mellitus and deafness, *Nature Genetics*, **22** (3): 300-4.

Lawrance, I.C., *et al* (2001). Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate gene, *Human Molecular Genetics*, **10**: 445-56.

Laybutt, D.R., *et al* (2007). Endoplasmic reticulum stress contributes to β-cell apoptosis in type 2 diabetes, *Diabetologia*, **50**: 752-63.

Lees, G.J., *et al* (2001). Inhibition of sodium-potassium-ATPase: a potentially ubiquitous mechanism contributing to central nervous system neuropathology, *Brain Research Reviews*, **16**(3): 283-300.

Levitzki, A., *et al* (1973). The allosteric activation of mammalian a-amylase by chloride, *European Journal of Biochemistry*, **41**: 171-80.

Liman, J., *et al* (2005). Interaction of BAG1 and Hsp70 mediates neuroprotectivity and increases chaperone activity, *Molecular Cell Biology*, **25**(6): 3715-25.

Lin, L., *et al* (2001). 5-HT 1b antagonist in the PVN blocks the feeding response to enterostatin in the amygdale, *Obesity Research*, **9**: S56.

Lin, L., *et al* (2002). Procolipase gene and enterostatin expression in the rat brain *(Abstract)*, *The Journal of the Federation of American Societies for Experimental Biology*, A783.

Lindholm, D., *et al* (2006). ER stress and neurodegenerative diseases, *Cell Death and Differentiation*, **13**: 385-92.

Livesey, F.J., *et al* (1997). A Schwann cell mitogen accompanying regeneration of motor neurons, *Nature*, **390**: 614-18.

Lowe, M.E., *et al* (1989). Cloning and characterization of human pancreatic lipase cDNA, *The Journal of Biological Chemistry*, **264**(33): 20042-48.

Lowe, M.E., *et al* (1990). Cloning and characterization of the human colipase cDNA, *Biochemistry*, **29**(3): 823-28.

Lowe, M.E., *et al* (1992). The catalytic site residues and interfacial binding of human pancreatic lipase, *The Journal of Biological Chemistry*, **267**(24): 17069-73.

McManus, M.T., *et al* (2002). Gene silencing in mammals by small interfering RNAs, *Nature Reviews: Genetics*, **3**: 737-47.

McBain, S.C., *et al* (2003). Functional effects of expression of wolframin-antisense transcripts in BRIN-BD11 beta-cells, *Biochemical and Biophysical Research Communications*, **307(**3): 684-8.

Marchetti, P., *et al* (2007). The endoplasmic reticulum in pancreatic β-cells of type 2 diabetes patients, *Diabetologia*, **50**: 2486-94.

Martindale, J.J., *et al* (2006). Endoplasmic reticulum stress gene induction and protection from ischemia/reperfusion injury in the hearts of transgenic mice with a tamoxifen-regulated form of ATF6, *Circulation Research*, **98**(9): 1186-93.

Mathews, L., *et al* (2009). Thiamine-responsive megaloblastic anemia, *Indian Pediatrics*, **46**: 172-74.

Meex, S.J., *et al* (2007). Activating transcription factor 6 polymorphisms and haplotypes are associated with impaired glucose homeostasis and type 2 diabetes in Dutch Caucasians, *The Journal of Clinical Endocrinology and Metabolism*, **92**(7): 2720-25.

Mersch, B., *et al* (2008). Automatic detection of exonic splicing enhancers (ESEs) using SVMs, *BMC Bioinformatics*, **9**: 369.

Mezghani, N., *et al* (2011). The mitochondrial ND1 m.3337G>A mutation associated to multiple mitochondrial DNA deletions in a patient with Wolfram syndrome and cardiomyopathy, Biochemical and Biophysical Research Communications, [online], available at: <http://www.checkorphan.org/research/view/wolfram-syndrome/page/6> [Accessed September 2011].

Minton, J.A.L., *et al* (2002). Association studies of genetic variation in the *WFS1* gene and type 2 diabetes in U.K. populations, *Diabetes*, **51**: 1287-90.

Miyashita, H., *et al* (1995). Human *REG* family genes are tandemly ordered in a 95-kilobase region of chromosome 2p12, *FEBS Letters*, **377**: 429-33.

Morell, M., *et al* (2008). Study and selection of *in vivo* protein interactions by coupling bimolecular fluorescence complementation and flow cytometry, *Nature Protocols*, **3**: 22-33.

Neufeld, E.J., *et al* (1997). Localization of the gene for thiamine-responsive megaloblastic anemia syndrome on the long arm of chromosome 1 by homozygosity mapping, *American Journal of Human Genetics*, **61**: 1335-41.

Neuwirth, A.K., *et al* (2009). Thiamine responsive megaloblastic anemia: Identification of novel compound heterozygotes and mutation update, *Journal of Paediatrics*, **155** (6): 888-92.

O'Brien, B.A., *et al* (1997). Apoptosis is the mode of β-cell death responsible for the development of IDDM in the nonobese diabetic (NOD) mouse, *Diabetes*, **46**: 750-57.

Obrig, T.G., *et al* (1971). The mechanism by which cyclohexamide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes, *Journal of Biological Chemistry*, **246**: 174-81.

Ohtsuki, T., *et al* (2000). *WFS1* gene mutation search in depressive patients: detection of five missense polymorphisms but no association with depression or bipolar affective disorder, *Journal of Affective Disorders*, **58**: 11-17.

Oikawa, D., *et al* (2007). Self-association and BiP dissociation are not sufficient for activation of the ER stress sensor Ire1, *Journal of Cell Science*, **120**: 1681-8.

Okada, S., *et al* (1993). Procolipase mRNA: tissue localization and effects of diet and adrenalectomy, *Biochemical Journal*, **292**: 787-9.

Olivecrona, T., *et al* (1993). Lipoprotein lipase and hepatic lipase, *Current Opinion in Lipidology*, **4**: 187-96.

Osman, A.A., *et al* (2003). Wolframin expression induces novel ion channel activity in endoplasmic reticulum membranes and increases intracellular calcium, *Journal of Biological Chemistry*, **278**: 52755-62.

Oyadomari, S., *et al* (2001). Nitric oxide-induced apoptosis in pancreatic β-cells is mediated by the endoplasmic reticulum stress pathway, *Proc Natl Acad Sci USA*, **98**: 10845-50.

Oyadomai, s., et al (2002). Targeted disruption of the *Chop* gene delays endoplasmic reticulum stress–mediated diabetes, *Journal of Clinical Investigation*, **109**(4):525–32.

Ozturk, M., *et al* (1989). Elevated levels of an exocrine pancreatic secretory protein in Alzheimer's disease brain, *Proceedings of the National Academy of Sciences in the United States of America*, **86**: 419-23.

Palmer, J.P., *et al* (2005). Is latent autoimmune diabetes in adults distinct from type 1 diabetes or just type 1 diabetes at an older age? *Diabetes*, **54**(2): S62-67.

Paddison, P.J., *et al* (2002). Short hairpin RNas (shRNAs) induce sequence-specific silencing in mammalian cells, *Genes and Development*, **16**: 948-58.

Park, M.E., *et al* (2009). Enterostatin alters protein trafficking to inhibit insulin secretion in Beta-TC6 cells, *Peptides*, **30**(10): 1866-73.

Peltonen, L., *et al* (2006). Lessons from studying monogenic disease for common disease, *Human Molecular Genetics*, **15** (Review issue 1): R67-R74.

Philbrook, C., *et al* (2005). Expressional and functional studies of Wolframin, the gene function deficient in Wolfram syndrome, in mice and patient cells, *Experimental Gerontology*, **40**(8-9):671-78.

Phillips, M., *et al* (1996). Role of the prodomain in folding and secretion of rat pancreatic carboxypeptidase A1, *Biochemistry*, **35**(21): 6771-76.

Pirot, P., *et al* (2007). Transcriptional regulation of the endoplasmic reticulum stress gene chop in pancreatic insulin-producing cells, *Diabetes*, **56**: 1069-77.

Polymeropoulos, M.H., *et al* (1994). Linkage of the gene for Wolfram syndrome to markers on the short arm of chromosome 4, *Nature Genetics*, **8**: 95-97.

Puca, A.A., *et al* (2001). A genome-wide scan for linkage to human exceptional longevity identifies a locus on chromosome 4q, *Proceedings of the National Academy of Sciences in the United States of America*, **98**: 10505-08.

Raz, T., *et al* (2000). The spectrum of mutations, including four novel ones, in the thiamine-responsive megaloblastic anemia gene, *SLC192A* of eight families, *Human Mutation*, **16** (1): 37-42.

Ren, J., *et al* (2007). Glucose stimulated insulin secretion, *Journal of Translational Medicine*, **5**: 1.

Reverter, D., *et al* (1998). Overexpression of human procarboxypeptidase A2 in *Pichia pastoris* and detailed characterisation of its activation pathway, *The Journal of Biological chemistry*, **273**: 3535-41.

Reznik, S.E., *et al* (2001). Carboxypeptidases from A-Z: implications in embryonic development and Wnt binding, *Cellular and Molecular Life Sciences*, **58**: 1790-1804.

Riggs, A.C., *et al* (2005). Mice conditionally lacking the Wolfram gene in pancreatic islet beta cells exhibit diabetes as a result of enhanced endoplasmic reticulum stress and apoptosis, *Diabetologia*, **48**(11): 2313-21.

Rogers, L.E., *et al* (1969). Thiamine responsive megaloblastic anemia, *Journal of Paediatrics*, **74**: 494-504.

Rohayem, J., *et al* (2011). Diabetes and neurodegeneration in Wolfram syndrome: a multicenter study of phenotype and genotype, *Diabetes Care*, **34**(7): 1503-10.

Rötig, A., *et al* (1993). Deletion of mitochondrial DNA in a case of early-onset diabetes mellitus, optic atrophy and deafness (Wolfram syndrome, MIM 222300), *Journal of Clinical Investigation*, **91**: 1095-98.

Rozen, S., and Skaletsky, H.J. (2000). Primer3. [online] (updated 27/2/2009) Available at: <http://frodo.wi.mit.edu/primer3> [Accessed December 2010].

Ruffner, H., *et al* (2007). Human protein-protein interaction networks and the value for drug discovery, *Drug Discovery Today*, **12**(17/18): 709-16.

Rutkowski, D.T., and Kaufman, R.J., (2007). That which does not kill me makes me stronger: adapting to chronic ER stress, *Trends in Biochemical Sciences*, **32**(10): 469-76.

Saleh, M., *et al* (2006). Enhanced bacterial clearance and sepsis resistance in caspase-12 deficient mice, *Nature*, **440**: 1064-68.

Sandu, M.S., *et al* (2007). Common variants in *WFS1* confer risk of type 2 diabetes, *Nature Genetics*, **39**(8): 951-953.

Schröder, M., *et al* (2005). The mammalian unfolded protein response, *Annual Reviews in Biochemistry*, **74**: 739-89.

Sequeira, A., *et al* (2003). Wolfram syndrome and suicide: Evidence for a role of WFS1 in suicidal and impulsive behaviour, American Journal of Medical Genetics; B Neuropsychiatric Genetics, **119B**(1): 108-13.

Silvestre, R.A., *et al* (1996). Effect of enterostatin on insulin, glucagon, and somatostatin secretion in the perfused rat pancreas, *Diabetes*, **45**(9): 1157-60.

Skidgel, R.A., *et al* (1988). Basic carboxypeptidases: regulators of peptide hormone activity, *Trends in Pharmacological Sciences*, **9**(8): 299-304.

Smith, C.J.A., *et al* (2004). Genotype-phenotype correlations in a series of Wolfram syndrome families, *Diabetes Care*, **27** (8): 2003-09.

Song, J.J., *et al* (2011). Plasma carboxypeptidase B downregulates inflammatory responses in autoimmune arthritis, *Journal of Clinical Investigation*, **121**(9): 3517-27.

Southern, C., *et al* (1990). Inhibition of insulin secretion by interleukin-1β and tumour necrosis factor-α via an L-arginine-dependent nitric oxide generating mechanism, *FEBS Lett*, **276**: 42-44.

Steck, A.K., *et al* (2011). Genetics of Type 1 Diabetes, *Clinical Chemistry*, **57**(2): 176-85.

Steer, M., *et al* (1974). The role of sulfhydryl groups in the action and structure of mammalian alpha-amylase, *Biochemica et Biophysica Acta*, **334**: 389.

Stewart, T.A., *et al* (1989). The human reg gene encodes pancreatic stone protein, *Biochemical Journal*, **260**: 622-23.

Strom, T.M., *et al* (1998). Diabetes insipidus, diabetes mellitius, optic atrophy and deafness (DIDMOAD) caused by mutations in a novel gene (*Wolframin*) coding for a predicted transmembrane protein, *Human Molecular Genetics*, **7**(13): 2021-28.

Suzuki, H., *et al* (2004). In vitro pull-down assay without expression constructs, *BioTechniques*, *BENCHMARKS*, **37**(6): 918-19.

Swift, M., and Swift, R.G., (2005). Wolframin mutations and hospitalization for psychiatric illness, *Molecular Psychiatry*, **10**: 799-803.

Szegezdi, E., *et al* (2003). Caspase-12 and ER-stress-mediated apoptosis: the story so far. *Ann NY Acad Sci*, **1010**: 186-94.

Takeda, K., *et al* (2001). *WFS1* (Wolfram syndrome 1) gene product: Predominant sub-cellular localisation to endoplasmic reticulum in cultured cells and neuronal expression in rat brain, *Human Molecular Genetics*, **10**(5): 477-84.

Takei, D., *et al* (2006). WFS1 protein modulates the free $Ca2^{+}$ concentration in the endoplasmic reticulum, *FEBS Letters*, **580**: 5635-40.

Terazono, K., *et al* (1988). A novel gene activated in regenerating islets, *Journal of Biological Chemistry*, **263**: 2111-14.

Thorel, F., *et al* (2010). Conversion of adult pancreatic α-cells to β-cells after extreme β-cell loss, *Nature*, **464**: 1149-54.

Torres, L.E.R., *et a*l (2001). Mutation screening of the Wolfram syndrome gene in psychiatric patients, *Molecular Psychiatry*, **6**(1): 39-43.

Traven, A., *et al* (2006). Yeast Gal4: a transcriptional paradigm revisited, *EMBO reports*, **7**: 496-99.

Ueda, K., *et al* (2005). Endoplasmic reticulum stress induces *Wfs1* gene expression in pancreatic beta-cells via transcriptional activation, *European Journal of Endocrinology*, **153**(1): 167-76.

Uno, S., *et al* (2007). Macrophages and dendritic cell infiltrating islets with or without β-cells produce tumour necrosis factor-α in patients with recent-onset type 1 diabetes, *Diabetologia*, **50**: 596-601.

Urano, F., *et al* (2000). Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1, *Science*, **28** (5453): 664-66.

Van Den Ouweland, J.M.W., *et al* (2003). Molecular characterisation of *WFS1* in patients with Wolfram syndrome, *Journal of Molecular Diagnostics*, **5** (2): 88-95.

Ventura, S., *et al* (1999). Mapping the pro-region of Carboxypeptidase B by protein engineering, cloning, overexpression, and mutagenesis of the porcine proenzyme, *The Journal of Biological Chemistry*, **274**: 19925-33.

Vitale, A., *et al* (1999). The endoplasmic reticulum- gateway of the secretory pathway, Plant *Cell*, **11**: 615-28.

Voight, B.F., *et al* (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis, *Nature Genetics*, **42**(7): 579-90.

Wang, X., *et al* (2007). Thiamine deficiency induces endoplasmic reticulum stress in neurons, *Neuroscience*, **144**(3): 1045-56.

Watanabe, T., *et al* (1994). Pancreatic β-cell replication and amelioration of surgical diabetes by Reg protein, *Proceedings of the National Academy of Sciences in the United States of America*, **91**: 3589-92.

World Health Organisation, (2012). [online], available at: <www.thewho.org> [Accessed May 2011].

Wild, S., *et al* (2004). Global prevalence of diabetes; Estimates for the year 2000 and projections for 2030, *Diabetes Care*, **27**:1047-53.

Wiley, S., *et al* (2007). MitoNEET is an iron-containing outer mitochondrial membrane protein that regulates oxidative capacity, *Proceedings of the National Academy of Sciences*, **104**(13): 5318-23.

Whitcomb (2009). Pancreas - Cybergene diagnostics [online], available at: <www.pancreas.org> [Accessed November 2011].

Wolfram D.J., and Wagner H.P., (1938). Diabetes mellitus and simple optic atrophy among siblings: Report of four cases, *Mayo. Clin. Proc*., **1**: 715-18.

Wu, J., *et al* (2007). ATF6alpha optimizes long-term endoplasmic reticulum function to protect cells from chronic stress, *Developmental Cell*, **13**(3): 351-64.

Wong, H., *et al* (1992). The lipase gene family, *Journal of Lipid Research*, **43**: 993-99.

Xie, H-G., *et al* (2005). Pharmacogenomics steps toward personalized medicine, *Personalized Medicine*, **2**(4): 325-37.

Xu, R., *et al* (2009). Expression and localization of *Wolfram syndrome 1* gene in developing rat pancreas, *World Journal of Gastroenterology*, **15**(43): 5425-31

Yamada, T., *et al* (2006). *WFS1*-deficiency increases endoplasmic reticulum stress, impairs cell cycle progression and triggers the apoptotic pathway specifically in pancreatic β-cells, *Human Molecular Genetics*, **15**(10): 1600-09.

Yamaguchi, S., *et al* (2004). Endoplasmic reticulum stress and N-glycosylation modulate expression of WFS1 protein, *Biochemical and Biophysical Research Communications*, **325**: 250-56.

Yan, W., *et al* (2002). Control of PERK eIF2α activity by the endoplasmic reticulum stress-induced molecular chaperone $P58^{IPK}$, *Proceedings of the National Academy of Sciences in the United States of America*, **99**(25): 15920-5.

Ye, R., *et al* (2010). Grp78 heterozygosity promotes adaptive unfolded protein response and attenuates diet-induced obesity and insulin resistance, *Diabetes*, **59**(1): 6-16.

Young, T.L., *et al* (2001). Non-syndromic progressive hearing loss *DFNA38* is caused by heterozygous missense mutation in the Wolfram syndrome gene *WFS1*, *Human Molecular Genetics*, **10** (22): 2509-14.

Yoshida, H., (2007). ER stress and diseases, *FEBSJ*, **274**(3): 630-58.

Yurimoto, S., *et al* (2009). Identification and characterisation of wolframin, the product of the Wolfram syndrome gene (*WFS1*), as a novel Calmodulin-binding protein, *Biochemistry*, **48**: 3946-55.

Zalsman, G., *et al* (2009). Wolfram gene H611R polymorphism: No direct association with suicidal behaviour but possible link to mood disorders, *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, **33**: 707-10.

Zatyka, M., *et al* (2008a). Sodium-potassium ATPase β1 subunit is a molecular partner of wolframin, an endoplasmic reticulum protein involved in ER stress, *Human Molecular Genetics*, **17**(2): 190-200.

Zatyka, M., *et al* (2008b). Wolfram depletion in MIN6 cells leads to defective glucose signalling and insulin secretion via altered ATP generation and endoplasmic reticulum calcium homeostasis, (unpublished data).

Zhou, Q., *et al* (2008). *In vivo* reprogramming of adult pancreatic exocrine cells to β-cells, *Nature*, **455**: 627-33.

Zimmet, P., *et a*l (2001). Global and societal implications of the diabetes epidemic, *Nature*, **6865**: 782-7.

Zymslowska, A., *et al* (2011). Wolfram syndrome in the Polish population: novel mutations and genotype-phenotype correlation, *Clinical Endocrinology*, **75**(5): 636-41.

Supplier	Location/Distribution	Web
Abcam	330 Cambridge Science Park, Cambridge, CB4 0FL, UK	www.abcam.com
ABI/Applied Biosystems	Kelvin Close, Birchwood Science Park North, Warrington, Cheshire, WA7 7PV, UK	www.appliedbiosystems.com
AltaBioscience	AltaBioscience, Building Y10, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK	www.altabioscience.bham.ac.uk
Appleton Woods Ltd	Linden House, Heeley Road, Selly Oak, Birmingham, UK	www.corning.com
Bioline	Bioline Ltd, Unit 16 The Edge Business Centre, Humber Road, London, NW2 6EW, UK	www.bioline.com
Biorad	Bio-Rad Laboratories Ltd, Bio-Rad House, Maxted Road, Hemel Hempstead, Hertfordshire, HP2 7DX, UK	www.bio-rad.com
Bright Instruments Co Ltd	Bright Instrument Co Ltd, St Margarets Way, Huntingdon, Cambridgeshire, PE29 6EU, UK	www.brightinstruments.co.uk
Carl Zeiss Ltd	15 – 20 Woodfield Road, Welwyn Garden City, Hertfordshire, AL7 1JQ	www.zeiss.co.uk
Clontech	Unit 2, Intec 2, Wade Road, Basingstoke, Hampshire, RG24 8NE, UK.	www.clontech.co.uk
European collection of cell cultures (ECACC)	Distributed by Sigma-Aldrich, Fancy road, Poole, Dorset, BH12 4QH, UK.	www.hpacultures.org.uk
Fermentas	Sheriff House, Sheriff Hutton Industrial Park, York, YO60 RZ, UK	www.fermentas.com
Fisher Scientific	Fisher Scientific UK Ltd, Bishop Meadow Road, Loughborough, LE11 5RG, UK	www.fisherscientific.co.uk
Geneservice (Source Bioscience)	Units 24 and 25 William James House, Cowley Road, Cambridge, CB4 0WU, UK.	www.geneservice.co.uk
GE Healthcare	Amersham Place, Little Chalfont, Buckinghamshire, HP7 9NA, UK.	www.amershambiosciences.com
Invitrogen	Inchinnan Business Park, 3 Fountain Drive, Paisley, PA4 9RF, UK.	www.invitrogen.com

Appendix 1A. Supplier details for laboratory equipment and reagents.

Supplier	Location/Distribution	Web
MERCK	Bedfont Cross, Stanwell Road, Feltham, Middlesex, TW14 8NX, UK	www.merckgroup.com
New England Biolabs	New England Biolabs (UK) Ltd., 75/77 Knowl Piece, Wilbury Way, Hitchin, Herts. SG4 0TY, UK	www.NEB.com
Promega	Delta House, Southampton Science Park, Southampton, Hampshire, SO16 7NS, UK.	www.promega.com
Proteintech	Proteintech Group, Inc., 2201 Campbell Park Drive, Chicago, IL 60612, USA	www.ptglab.com
Qiagen Ltd.	Qiagen House, Flemming Way, Crawley, RH10 9NQ, UK.	www.qiagen.com
Roche Diagnostics Ltd.	Charles Avenue, Burgess Hill, West Sussex, RH15 9RY.	www.roche-diagnostics.co.uk
Santa-Cruz Biotechnology, Inc.	Bergheimer Str. 89-269115 Heidelberg, Germany	www.scbt.com
Sigma	Sigma-Aldrich, Fancy Road, Poole, Dorset, BH12 4QH, UK.	www.sigmaaldrich.com
Stratagene	1 Mostyn Street, Leicester, Leicestershire, LE3 6DT, UK.	www.stratagene.com
Stressgen	Enzo Life Sciences (UK) Ltd., Palatine House, Matford Court, Exeter, EX2 8NL, UK.	www.enzolifesciences.com
Surgipath Europe Ltd	Unit 1/Venture Park, Stirling Way, Peterborough, Cambridgeshire, PE3 8YD	www.icambridgeshire.co.uk
Syngene	Beacon House, Nuffield Road, Cambridge, CB4 1TF, UK	www.syngene.com
Thermo Scientific	ABgene House, Blenheim Road, Epsom, Surrey, KT19 9AP, UK	www.thermoscientific.com
Vector Laboratories Ltd	3 Accent Park, Bakewell Road, Orton Southgate, Peterborough, PE2 6XS, UK	www.vectorlabs.com

Appendix 1B. Supplier details for laboratory equipment and reagents.

Human *WFS1* genomic DNA	Sequence
Exon 2F	GGA TGT GCC TGA CCT TGA CT (20mer)
Exon 2R	CAA TGC TGA ACT GCA GAG GA (20mer)
Exon 3F	CCC TCC ATC CTG ACA AGT GA (20mer)
Exon 3R	ACT TCT CTG TGG GCT GTG GA (20mer)
Exon 4F	GAG AGG GTC GGA GAA TCT GG (20mer)
Exon 4R	GCA GCA TTA CAA GCT GCT CA (20mer)
Exon 5F	AAG CCT AGG CAG GGC ACA (18mer)
Exon 5R	CCC CTA TGG GAA GGT CCT G (19mer)
Exon 6F	AGG GCA CGA GGA GAT AGT CA (20mer)
Exon 6R	AGA TAG GGC AGG CCA GGA T (19mer)
Exon 7F	CCC ATT GCT CTG TGT GAG G (19mer)
Exon 7R	GAA GGT GCC CTG CCT GAG (18mer)
Exon 8aF	AGA GGC AGG GTG GTC AGA G (19mer)
Exon 8aR	CAC GGT GAT GAA GGT CTG G (19mer)
Exon 8bF	GAC CAG CTA CCT GAG CCT GA (20mer)
Exon 8bR	GCA CAG CGC ACC ATA CTG (18mer)
Exon 8cF	ATC GTG CTG TTC TGC TGG TT (20mer)
Exon 8cR	CAG TCG TGC TCG ATC TTC AC (20mer)
Exon 8dF	TCA TCG AGT TCA GCA CCA TC (20mer)
Exon 8dR	AAG GGG AAG AGC TGC TAA GG (20mer)
Exon 8eF	AAG CAC CCT GTT CCC TCT TT (20mer)
Exon 8eR	AGG CTC TGC TGA TTC TCA CG (20mer)

Appendix 2.1. Primer sequences

Human *WFS2* genomic DNA	Sequence
Exon 1F	GCT CGG GAG AGG AGT TGA C (19mer)
Exon 1R	GGA TTT TTA CGC CTC CTT CC (20mer)
Exon 2F	AGC ACT GCA GAT TCT GAC ACA (21mer)
Exon 2R	GGG GAT TTA AGA GGC GAA AC (20mer)
Exon 3F	GCT TTC TTT CTG AGA GCA TTT C (22mer)
Exon 3R	CCA GTA GTA ATA ATT AAG ACC ACC ATT (27mer)

Appendix 2.2. Primer sequences

Human *ATP1-B1* genomic DNA	Sequence
H - ATPB1 1F	AGA GCC AGG CCG GAG AAG (18mer)
H - ATPB1 1R	CAT TCC GCA CAG GAC TGC (18mer)
H - ATPB1 2F	TGT TCA TCC GTG GGA AGA TT (20mer)
H - ATPB1 2R	CCT TTT CAT GGT TGC TTT AAG A (22mer)
H - ATPB1 3F	CCT TTC TTT TTA GCC AAG TGG A (22mer)
H - ATPB1 3R	TGG CCA GAG TTC AAT CTT TCA (21mer)
H - ATPB1 4F	CAA GGA AGA CAG TTT CTT TTT CTT A (25mer)
H - ATPB1 4R	GCT ACG TTT CAA AGT CAG CAC T (22mer)
H - ATPB1 5F	TGT TTG GAC TAC GGA GTA GTT TT (23mer)
H - ATPB1 5R	CCT GTC ACT TAA ACA GTC TTC TTC A (25mer)
H - ATPB1 6F	TCT CAA GGC TGC AAT GGA AT (20mer)
H - ATPB1 6R	TGA CAG TTT GTT CAA GAC TAG TAG GTT (27mer)
H - ATPB1 7F	CCC ATC GAT GAG CAT TTT TA (20mer)
H - ATPB1 7R	GGA GGA CAA GGC AAG AGA TG (20mer)

Appendix 2.3. Primer sequences

Human *SLC19A2* genomic DNA	Sequence
Exon 1F	TAC AGG GAG AAG GCG TCA CT (20mer)
Exon 1R	CGC TTT TCT CGG TCC TCT CT (20mer)
Exon 2aF	TTC CAG GTC CTT TCA TCA CT (20mer)
Exon 2aR	CTG AGA CAA GGA TTT GCC CTA (21mer)
Exon 2bF	GGC ATG TAC CAG AAA GTC ACA A (22mer)
Exon 2bR	CCA GCC CCC ATA GTA GCA AT (20mer)
Exon 3F	GAG CAT TGT TCC CTT TTG CT (20mer)
Exon 3R	CCC ACC ACG ACC CTC TAT TA (20mer)
Exon 4F	GCA ACA GCA TTT GTG TAG CAA (21mer)
Exon 4R	TGT TAC AAT TTT TCC TAA GGC TTC (24mer)
Exon 5F	TGA CAG TAG AAA GAA GGA CTC TTC A (25mer)
Exon 5R	TCA CAC ATA CTA CCT TCG ATG C (22mer)
Exon 6F	GTA ATC CTC TTG TCA TTG ATA ATT CA (26mer)
Exon 6R	CAA ACA CAT CCA GGC AGT TG (20mer)

Appendix 2.4. Primer sequences

Human *WFS1* cDNA	Sequence
CR - WFS1 1R	GGC TCC TGG TAT GCT GGG CC (20mer)
CR - WFS1 2F	CCT GGT GTT AGA GAC GCA GC (20mer)
CR - WFS1 3F	CTC AAC AGC TGC ACC GCT GT (20mer)
CR - WFS1 3R	CCA TCG TGC TCG TTG ACC TG (20mer)
CR - WFS1 4F	GGA AGC TCA ACC CCA AGA AG (20mer)
CR - WFS1 4R	CCA GCT CGT CAT CAT CTT CG (20mer)
CR - WFS1 5F	CCA AGG GCG TCA TCC CCA GC (20mer)
CR - WFS1 5R	GGA CAG CCA GTG CAT GCC TG (20mer)
CR - WFS1 6F	GGA GAT CAA GGA GTA CCT GA (20mer)
MZ - WFS1 6F	GGC CTG CTA TCG CTG CTG CCC (21mer)
MZ - WFS1 7F	CCT GGT GGC CGG CCT GGC CCT (21mer)
MZ - WFS1 8F	CAG ATC CTC TGC AGC CAC CTG (21mer)
MZ - WFS1 8F	CGT CAC CAA GGA CAT CGT GCT (21mer)

Appendix 2.5. Primer sequences

Mouse *WFS1* cDNA	Sequence
MZ - WFS1 1F	CCT AGC TGA CCG GAA AGG CAT (21mer)
MZ - WFS1 2F	CAG AAG GTG GTG AAG TAC CCT (21mer)
MZ - WFS1 3F	GCT GGC CAG CAA GGA CTG CAT (21mer)
MZ - WFS1 4F	CCC CTG CCA CAT CAA GAA GTT (21mer)
MZ - WFS1 5F	GAA GAG CTC CAT GGT GAA GCT (21mer)

Appendix 2.6. Primer sequences

pGBKT7/pACT2 vector sequencing primers	Sequence
T7	TAA TAC GAC TCA CTA TAG GG
3'-DNA-AD	AGA TGG TGC ACG ATG CAC AG

Appendix 2.7. Primer sequences

Site-Directed Mutagenesis	Sequence
Human WFS1 - E158K-S	GAG GCA TCA CGT CCA AGA ACG AAC GGG AG (29mer)
Human WFS1 - E158K-A	CTC CCG TTC GTT CTT GGA CGT GAT GCC TC (29mer)
Human WFS1 - L511P-S	CCT GCT CTA TGT CTA CCC GCT CTA TCT CTT CTT CC (35mer)
Human WFS1 - L511P-A	GGA AGA AGA GAT AGA GCG GGT AGA CAT AGA GCA GG (35mer)
Human WFS1 - R558H-S	GGG GCT GCT CCA GCG CTC CAT CG (23mer)
Human WFS1 - R558H-A	CGA TGG AGG CGT GGA GCA GCC CC (23mer)
Human WFS1 - R611H-S	CCC CTG CTG TTG CAC TGG TGG ACC AAG (27mer)
Human WFS1 - R611H-A	CTT GGT CCA CCA GTG CAA CAG CAG GGG (27mer)
Human WFS1 - G736A-S	GCT GCC TCT ACG CCG AGG CCT ACC C (25mer)
Human WFS1 - G736A-A	GGG TAG GCC TCG GCG TAG AGG CAG C (25mer)
Human WFS1 - E737K-S	CTG CCT CTA CGG CAA GGC CTA CCC TGC (27mer)
Human WFS1 - E737K-A	GCA GGG TAG GCC TTG CCG TAG AGG CAG (27mer)
Human WFS1 - R818C-S	CTG CTC AGC CTG TGC CAG GGC AGC C (25mer)
Human WFS1 - R818C-A	GGC TGC CCT GGC ACA GGC TGA GCA G (25mer)

Appendix 2.8. Primer sequences

Site-Directed Mutagenesis	Sequence
Mouse WFS1 - Clonal change S	CCA CCT TCT TCA CGG TGA CCA GCT ACA TGA G (31mer)
Mouse WFS1 - Clonal change A	CTC ATG TAG CTG GTC ACC GTG AAG AAG GTG G (31mer)
Mouse WFS1 - E158K-S	CGG AAA GGC ATC ACT TCT AAG AAC GAG GCT GAG (33mer)
Mouse WFS1 - E158K-A	CTC AGC CTC GTT CTT AGA AGT GAT GCC TTT CCG (33mer)
Mouse WFS1 - L511P-S	GCC TGC TCT ATG TCT ATC CCT TTT ACC TCT TCT TCC G (37mer)
Mouse WFS1 - L511P-A	CGG AAG AAG AGG TAA AAG GGA TAG ACA TAG AGC AGG C (37mer)
Mouse WFS1 - R558H-1S	TGG GCT TGG TCC AGG CCT CCA TCG G (25mer)
Mouse WFS1 - R558H-1A	CCG ATG GAG GCC TGG ACC AAG CCC A (25mer)
Mouse WFS1 - R558H-2S	CCT GGG CTT GGT CCA CGC CTC CAT CG (26mer)
Mouse WFS1 - R558H-2A	CGA TGG AGG CGT GGA CCA AGC CCA GG (26mer)
Mouse WFS1 - R611H-S	TAC CCC TGC TTT TCC ATT GGT GGA CCA AGG C (31mer)
Mouse WFS1 - R611H-A	GCC TTG GTC CAC CAA TGG AAA AGC AGG GGT A (31mer)
Mouse WFS1 - G736A-S	GCG CTG CCT GTA TGC CGA GGC CTA CC (26mer)
Mouse WFS1 - G736A-A	GGT AGG CCT CGG CAT ACA GGC AGC GC (26mer)
Mouse WFS1 - E737K-S	GCT GCC TGT ATG GCA AGG CCT ACC CAT CT (29mer)
Mouse WFS1 - E737K-A	AGA TGG GTA GGC CTT GCC ATA CAG GCA GC (29mer)
Mouse WFS1 - R818C-S	CTG CTG AAC CTG TGC CAG GGG AGC C (25mer)
Mouse WFS1 - R818C-A	GGC TCC CCT GGC ACA GGT TCA GCA G (25mer)

Appendix 2.9. Primer sequences

Microsatellite marker analysis	Sequence
D4S3256 (F)	TGG AAC ACC ATT TGT TTC TG
D4S3256 (R)	GGC TAA GTT TCC AGA GAA TCA
D4S1531 (F)	TGG GTG ACA GAG CTA GTC C
D4S1531 (R)	GAA CCA GCC TCG CAT ACC
D4S1564 (F)	GAA TGC TCT CCA GTT ACT GAT AAA G
D4S1564 (R)	AAA CCG TGG CAA AGG AAA TAA AAA C

Appendix 2.10. Primer sequences

PCR primers for adding restriction sites	Sequence
Human CLPS restriction site primer *EcoR1* (F)	GGG AAT TCC CAT GGA GAA GAT CCT GAT C (28mer)
Human CLPS restriction site primer *Xho1* (R)	GGC TCG AGT ACA GCA TTC TGG GCT AGG T (28mer)
Human *REG1A* restriction site primer *Sal 1* (F)	GGG TCG ACC ATG GCT CAG ACC AGC TCA TA (29mer)
Human *REG1A* restriction site primer *Xho 1*(R)	GGC TCG AGC TAG TTT TTG AAC TTG CAG A (28mer)
Human CPB1 restriction site primer *EcoR1* (F)	GGG AAT TCC CAT GTT GGC ACT CTT GGT TCT (30mer)
Human CPB1 restriction site primer *Xho1* (R)	GGC TCG AGC TAG TAC AGG TGT TCC AGG A (28mer)
Human *ATF6* restriction site primer *Sal1* (F)	GGG TCG ACC ATG GGG GAG CCG GCT GGG GTT (30mer)
Human *ATF6* restriction site primer *Nhe1* (F)	GGG CTA GCA TGG GGG AGC CGG CTG GGG TT (29mer)
Human *ATF6* restriction site primer *Not1* (R)	CCG CGG CCG CCT ATT GTA ATG ACT CAG GGA TGG TGC TG (38mer)

Appendix 2.11. Primer sequences

Human *PNLIP* cDNA	**Sequence**
1R	AGT AAC CGA ACG CCG ACT GA (20mer)
1F	TCA GTC GGC GTT CGG TTA CT (20mer)
2F	GTG CCT CTT ACA ACG TCT TCA (21mer)
2R	TGA AGA CGT TGT AAG AGG CAC (21mer)
Human *CPB1* cDNA	**Sequence**
H - CPB1 1F	TGA AGG TTG GCA AAG CTG GAC (21mer)
H - CPB1 1R	TCT CTT ACA AAC CAC TGG CAG A (22mer)
H - CPB1 2F	CCA CTC GTA CTC CCA AAT GA (20mer)
***ATF6* cDNA**	**Sequence**
ATF6-1R	GAA AGT GGC TGA GGT TCT GC (20mer)
ATF6-1F	GCA GAA CCT CAG CCA CTT TC (20mer)
ATF6-2F	CCA AGC CTT TAT TGC TTC CA (20mer)
ATF6-3F	TGC TAG GGT TAG AGG CGA GA (20mer)
ATF6-4F	TGC TAA AGA GGC ACA GGA CA (20mer)
ATF6-5F	CAG GAA CTC AGG GAG TGA GC (20mer)

Appendix 2.12. Primer sequences

Rat *WFS1/β-actin* reverse-transcription PCR primers	**Sequence**
R-WFS1-RT-1F	CAT GCT GGA GCG TCT AGT GA (20mer)
R-WFS1-RT-1R	GGC CAC GTC AAT CAG GTA CT (20mer)
R-BACT-1F	GAT ATC GCT GCG CTC GTC (18mer)
R-BACT-1R	GGG GTG TTG AAG GTC TCA AA (20mer)

Appendix 2.13. Primer sequences

Patient number	DM	OA	Deafness	DI	Other	AGE	SEX
D94.1270	3	5	9	3	Neuro (9)	11	M
D94.1271	4	4	6	4	none	9	F
D94.1547	10	10	32	32	Renal (16), Neuro (11)	18 RIP	F
D94.1588	10	15	no	40	none	42	F
D94.5534	12	15	15	12	Renal (33), Neuro (36), Gonadal	39	M
D94.5533	7	10	30	12	Renal (34), Neuro (36), Gonadal	40	M
D94.2781	12	12	16	16	Neuro (35)	38	F
D94.3304	6	6	6	11	Renal (20), Neuro (30)	35	M
D94.2016	8	8	no	12	Neuro (28)	29	M
D94.2017	11	9	no	11	none	25	F
D94.6032	5	12	no	2.5	Renal (27), Neuro (29)	31	F
D94.6031	4	12	no	14	Renal (20), Neuro (34)	48	F
D94.1709	5	9	12	no	none	18	M
D95.3328	4	11	no	11	none	23	M
D94.5461	11	7	32	23	Renal (30), Neuro (32)	37 RIP	F
D94.5514	3	11	16	24	Renal (34), Neuro (36)	39	M
D94.4815	2	7	5	7	none	9	M
D94.1945	5	9	no	no	Renal (10), Neuro (26), Respiratory	29	M
D92.5373	4	10	26	10	Renal (10), Neuro (33), Psychiatric (37)	40	F
D94.5216	7	11	no	13	none	26	F
D94.8277	7	11	no	13	Neuro (22)	24	M
D94.5284	2	8	9	15	Renal (18), Neuro (32)	38	M
D94.6190	12	15	no	no	none	34	F
D94.7451	6	12	27	27	Renal (31), Neuro (31)	40	F
D94.7944	16	16	16	no	none	30	M

Appendix 3. West Midlands Regional Wolfram syndrome patient database clinical information. Reported age of onset of all primary Wolfram syndrome features.

KEY: DM – diabetes mellitus, OA – optic atrophy, DI – diabetes insipidus.

Patient number	DM	OA	Deafness	DI	Other	AGE	SEX
D94.7509	5	11	yes	yes	Renal/Neuro/Psychiatric	34	M
D94.7508	4	17	18	17	Renal (23), Neuro (38), Psychiatric	43	M
D94.1558	4	12	11	11	Renal (11), Neuro (30), Psychiatric	34	M
D953028	10	18	no	no	Sleep apnoea (30), Neurogenic bladder (13), Brain stem atrophy (30), Balance problems (30)	34	F
D95.3651	5	12	no	no	none	18	F
D95.7087	4	7	no	no	none	34	F
D95.7668	6	8	7	18	Renal (28)	31	M
D95.7669	6	8	8	24	Neuro (30)	32	F
D95.7672							
D95.8485	12	9	no	yes	Renal (27), Neuro (28)	41	M
D95.8941						37	F
D96.0073							
D96.0074							
D96.0075							
D05.25212	6	8	10	no	none	11	F
D02.09864						16	F
D01.05858	7		birth	no	Congenital blindness and deafness, not thought to be optic atrophy	10	F
D05.15697	4	no	5	no		6	F
D04.01235	30	12			Neuro (16), Neuropathic bladder (19), Mood swings (35)	43	F
D01.06664	14	13	no	no	none	21	F
D01.11436						45	F
D03.20431	5	4		5	Neuropathic bladder (16)	17	F
D00.8316						8	F
D05.20565	no	1	birth		Gross motor delay, Cerebellar vermis hypoplasia, Generalised cerebral atrophy		F
D02.03148							

Appendix 3. West Midlands Regional Wolfram syndrome patient database clinical information. Reported age of onset of all primary Wolfram syndrome features.

KEY: DM – diabetes mellitus, OA – optic atrophy, DI – diabetes insipidus.

Patient number	DM	OA	Deafness	DI	Other	AGE	SEX
D02.03147							
D01.16238							
D01.14183							
D01.12590	no	10	birth	no		16	M
D01.11844	11 (post renal transplant)	4	4	11	End stage renal disease, dysplastic kidneys, Bilat vesicoureteric reflux, Psychiatric (emotional instability)	11	M
D01.11341	6	6	39	40	Renal (40), Neuro - central respiratory failure (40)	45	F
D01.09361							
D01.08816							
D00.21491	3	11	no	11	Cardiac: ventricular septal defect and infundibular stenosis	12	M
D00.20766	3	12	no	no	none	26	M
D00.14257							
D00.12582	4	10	1	no	Congenital heart defect (Fallot's tetralogy)	13	M
D00.11041	10	8	no	no	Depression (psychiatric)	21	M
D00.7870						29	F
D00.6616							
D98.15053	4	10	2	17	none	24	M
D05.25758	1.5	no	3 months	no	Eye disease: Cone rod dystrophy (2) (not optic atrophy); Sideroblastic anaemia (9), Diagnosed thiamine responsive megaloblastic anaemia, diabetes and deafness (Rogers syndrome)	11	M
D02.11822							
D04.15546	8	8	7	No	Bladder instability. Pain in his limbs and muscles	8	M
D03.09228							
D03.09215							
D00.18057							
D00.18855							
D01.06359							
D00.15331	18	12	no	18	Neuropathic Bladder	41	M

Appendix 3. West Midlands Regional Wolfram syndrome patient database clinical information. Reported age of onset of all primary Wolfram syndrome features.

KEY: DM – diabetes mellitus, OA – optic atrophy, DI – diabetes insipidus.

Patient number	DM	OA	Deafness	DI	Other	AGE	SEX
D99.11552							
D99.6014							
D00.5338	5	8	no	no	Pain in limbs and muscles	11	F
D00.5339	4	7	no	6	Bilateral hydronephrosis; bladder instability, Pain in limbs and muscles	10	M
D01.04544							
D00.5345						30	F
D00.10649	14	15	no	16	none	14	F
D00.5923	5	7	no	11	Mood swings (psychiatry)	13	M
D00.5342						33	M
D00.12593						14	M
D00.8316						30	F
D00.6610						12	F
D99.11552						16	M
D094.1519	5	13	22	11	Renal (24), Neuro (30)	33	F
D094.2054	1.5	8	9	18	Renal (20), Neuro (29)	34 RIP	M
unknown							
unknown							
unknown							

Appendix 3. West Midlands Regional Wolfram syndrome patient database clinical information. Reported age of onset of all primary Wolfram syndrome features.

KEY: DM – diabetes mellitus, OA – optic atrophy, DI – diabetes insipidus.

Patient number	Protein (p.)	Mutation type	Zygosity
D94.1270	W648X	Nonsense	Comp. het
	Del517fs/ter521	Del	
D94.1271	W648X	Nonsense	Comp. het
	Del517fs/ter521	Del	
D94.1547	G437R	Missense	Comp. het
	Del567-568L	Del	
D94.1588	G437R	Missense	Comp. het
	Del567-568L	Del	
D94.5534	W478X	Nonsense	Comp. Het
	Del 538-542fster537	Del	
D94.5533	W478X	Nonsense	Comp. het
	Del 538-542fster537	Del	
D94.2781	W478X	Nonsense	Comp. het
	Q668X	Nonsense	
D94.3304	W478X	Nonsense	Comp. het
	Q668X	Nonsense	
D94.2016	Y302X	Nonsense	Comp. het
	G736X	Nonsense	
D94.2017	Y302X	Nonsense	Comp. het
	G736X	Nonsense	
D94.6032	Del517fs/ter521	Del	Comp. het
	Del888fs/ter949	Del	
D94.6031	Del517fs/ter521	Del	Comp. het
	Del888fs/ter949	Del	
D94.1709	P885L	Missense	Het
	Unknown		
D95.3328	Del883fs/ter949	Del	Comp. het
	E752X ; W700C	Nonsense/missense	
D94.5461	E169K	Missense	Comp. het
	C690R ; I296S	Missense	
D94.5514	Del354F	Del	Comp. het
	W478X	Nonsense	
D94.4815	E237X	Nonsense	Comp. het
	502Sins/8aa	Insertion	
D94.1945	Del883fs/ter949	Deletion	Het
	Unknown		
D92.5373	E452fsX471	?	Comp. het
	V871M	Missense	
D94.5216	Q667X	Nonsense	Comp. het
	Del883fs/ter949	Deletion	

Appendix 4. Genetic data (where available) for individuals listed on the West Midlands Regional Genetics WS database.

Patient number	Protein (p.)	Mutation type	Zygosity
D94.8277	Q667X	Nonsense	Comp. het
	Del883fs/ter949	Deletion	
D94.5284	Del812fs	Deletion	Hom
	Del812fs	Deletion	
D94.6190	R558H	Missense	Het
	Unknown		
D94.7451	E752X	Nonsense	Het
	Unknown		
D94.7944	P292S	Missense	Comp. het
	E273X	Nonsense	
D94.7509	Q136X	Nonsense	Hom
	Q136X	Nonsense	
D94.7508	Q136X	Nonsense	Hom
	Q136X	Nonsense	
D94.1558	Q136X	Nonsense	Hom
	Q136X	Nonsense	
D953028	P885L	Missense	?
	Silent?		
D95.3651	E273X	Nonsense	Comp. het
	G736S	Missense	
D95.7087	Del415V	Deletion	Hom
	Del415V	Deletion	
D95.7668	Del415V	Deletion	Hom
	Del415V	Deletion	
D95.7669	Del415V	Deletion	Hom
	Del415V	Deletion	
D95.7672	L804P	Missense	Comp. het
	C690R	Missense	
D95.8485	R558C	Missense	Comp. het
	Q366X	Nonsense	
D95.8941	C537X	Nonsense	Hom
	C537X	Nonsense	
D96.0073	W700X	Nonsense	Het
	Unknown		
D96.0074	W540X	Nonsense	Het
	Unknown		
D96.0075	W540X	Nonsense	Hom
	W540X	Nonsense	

Appendix 4. Genetic data (where available) for individuals listed on the West Midlands Regional Genetics WS database.

Patient number	Protein (p.)	Mutation type	Zygosity
D05.25212	Unknown		
	Unknown		
D02.09864	V412fsX440	Deletion	Hom
	V412fsX440	Deletion	
D01.05858	Unknown		
	Unknown		
D05.15697	Unknown		
	Unknown		
D04.01235	E169K	Missense	Comp. het
	A326V	Missense	
D01.06664	E202G	Missense	Comp. het
	E273X	Nonsense	
D01.11436	Unknown		
	Unknown		
D03.20431	P885L	Missense	Het
	Unknown		
D00.8316	Unknown		
	Unknown		
D05.20565	Unknown		
	Unknown		
D02.03148	Y528X	Nonsense	Hom
	Y528X	Nonsense	
D02.03147	Unknown		
	Unknown		
D01.16238	DNA unavailable		
D01.14183	R457S	Missense	Het
	Unknown		
D01.12590	E864K	Missense	Het
D01.11844	WFS2		
	WFS2		
D01.11341	Unknown		
	Unknown		
D01.09361	P533S	Missense	?
	silent?		

Appendix 4. Genetic data (where available) for individuals listed on the West Midlands Regional Genetics WS database.

Patient number	Protein (p.)	Mutation type	Zygosity
D01.08816	L432V	Missense	Comp. het
	G702C	Missense	
D00.21491	DNA unavailable		
D00.20766	DNA unavailable		
D00.14257	Silent?		
	Unknown		
D00.12582	V142fsX251	Deletion	Comp. het
	V871M	Missense	
D00.11041	R629W	Missense	Comp. het
	A126T	Missense	
D00.7870	W540C	Missense	Comp. het
	G820V	Missense	
D00.6616	Y669S	Missense	Hom
	Y669S	Missense	
D98.15053	Q112X	Nonsense	Hom
	Q112X	Nonsense	
D05.25758	Unknown/TRMA?		
D02.11822	DNA unavailable		
D04.15546	Del468L	Deletion	Hom
	Del468L	Deletion	
D03.09228	V412fsX440	Deletion	Het
	Unknown		
D03.09215	F247S	Missense	Het
	Unknown		
D00.18057	Unknown		
	Unknown		
D00.18855	Unknown		
	Unknown		
D01.06359	1228delCTCT	Deletion	Hom
	1228delCTCT	Deletion	
D00.15331	1689delCTTCCT	Deletion	Hom
	1689delCTTCCT	Deletion	
D99.11552	S430X	Nonsense	Hom
	S430X	Nonsense	

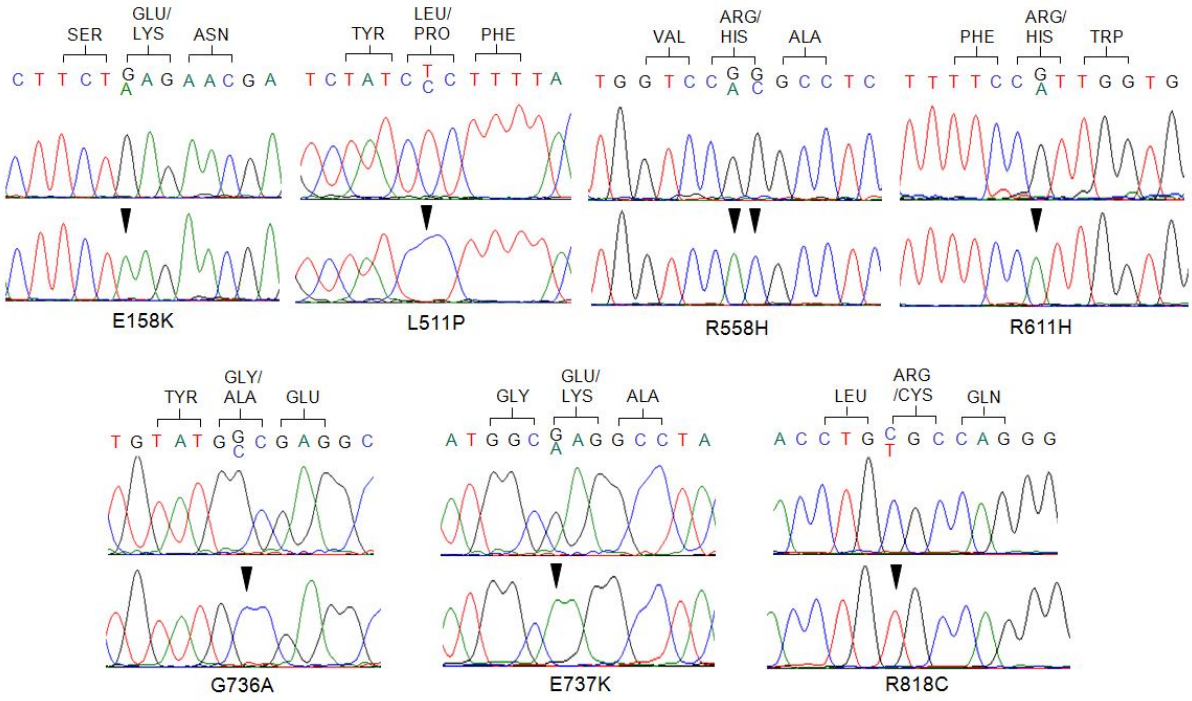
Appendix 4. Genetic data (where available) for individuals listed on the West Midlands Regional Genetics WS database.

Patient number	Protein (p.)	Mutation type	Zygosity
D99.6014	C647X	Nonsense	Hom
	C647X	Nonsense	
D00.5338	1399delCTG	Deletion	Hom
	1399delCTG	Deletion	
D00.5339	1399delCTG	Deletion	Hom
	1399delCTG	Deletion	
D01.04544	1237delTTC	Deletion	Hom
	1237delTTC	Deletion	
D00.5345	E752X		Het
	ND		
D00.10649	1240delTTC	Deletion	Comp. het
	1689delCTTCTT	Deletion	
D00.5923	W478X	Nonsense	Comp. het
	2648delCTTC	Deletion	
D00.5342	2648delCTTC	Deletion	Het
	ND		
D00.12593	1440del7nt	Del	Hom
	1440del7nt	Del	
D00.8316	A559T	Missense	Comp. het
	2642delTC	Del	
D00.6610	Q667X	Nonsense	Comp. het
	P742S	Missense	
D99.11552	S430X	Nonsense	Hom
	S430X	Nonsense	
D094.1519	E752X	Nonsense	Het
	ND		
D094.2054	P885L	Missense	Comp. het
	INV7-1G>A	Splice	
unknown	L511P	Missense	Hom
	L511P	Missense	
unknown	ND		
unknown	ND		

Appendix 4. Genetic data (where available) for individuals listed on the West Midlands Regional Genetics WS database.

	Marker		
	D4S3256	D4S1531	D4S1564
Patient D01.11844	227-227	323-323	129-129
WFS2 positive control	227-227	323-323	129-129
Normal control DNA	223-227	319-341	129-129

Appendix 5. **Comparative haplotype analysis of individuals with *WFS2*-associated Wolfram syndrome.** Microsatellite markers across a critical region of chromosome 4q22-24 analysed on an automated ABI3730 sequencer using GeneScan Analysis software. Data contains the estimated size (base pairs) of detected fragments, confirming a shared haplotype between patient D01.11844 in this study and an affected individual from a previously described cohort of Jordanian-Bedouin descent (Amr, 2000).



Appendix 6. Site-directed mutagenesis of mouse *wfs1*.

T1	*WFS1*	*WFS1* background	*WFS1* minus background	β-actin	β-actin background	β-actin minus background	*WFS1* divided by β-actin
Wt	1229330	1005740	223590	654948	605230	49718	4.497164
E158K	1228910	1005740	223170	662184	605230	56954	3.918425
L511P	1029160	1005740	23420	662698	605230	57468	0.407531
R558H	1074740	1005740	69000	669043	605230	63813	1.081284
R611H	1206880	1005740	201140	670017	605230	64787	3.104635
G736A	1180400	1005740	174660	666628	605230	61398	2.844718
E737K	1169840	1005740	164100	662768	605230	57538	2.852028
R818C	1133820	1005740	128080	655931	605230	50701	2.526183
T2							
Wt	1249950	968379	281571	801559	730758	70801	3.976935
E158K	1250310	968379	281931	799317	730758	68559	4.112239
L511P	1051060	968379	82681	804859	730758	74101	1.115788
R558H	1112200	968379	143821	798730	730758	67972	2.115886
R611H	1222230	968379	253851	801175	730758	70418	3.604916
G736A	1198910	968379	230531	805002	730758	74244	3.105046
E737K	1212250	968379	243871	793807	730758	63049	3.86796
R818C	1197180	968379	228801	781713	730758	50955	4.490256
T3							
Wt	1448210	1212280	235930	738227	647177	91050	2.591214
E158K	1474140	1212280	261860	734095	647177	86918	3.012725
L511P	1360570	1212280	148290	726559	647177	79382	1.868056
R558H	1440880	1212280	228600	722542	647177	75365	3.033238
R611H	1518360	1212280	306080	719425	647177	72248	4.236519
G736A	1506840	1212280	294560	720815	647177	73639	4.000054
E737K	1413300	1212280	201020	720545	647177	73368	2.739887
R818C	1460610	1212280	248330	701205	647177	54028	4.59632
T4							
Wt	1345770	1083650	262120	632134	579345	52789	4.965428
E158K	1350850	1083650	267200	616647	579345	37302	7.163155
L511P	1234910	1083650	151260	619655	579345	40310	3.752419
R558H	1283460	1083650	199810	619799	579345	40454	4.93919
R611H	1368510	1083650	284860	616988	579345	37643	7.56741
G736A	1326830	1083650	243180	629538	579345	50193	4.844899
E737K	1293640	1083650	209990	636591	579345	57246	3.668204
R818C	1321440	1083650	237790	639565	579345	60220	3.948688

Appendix 7. Clone set 1 harvested in RIPA buffer. Raw data figures for expression of *WFS1* variants in HEK293 cells normalised to β-actin. Measurements are calculated in arbitrary units of absorbance (Gel.Doc). T= transfection replicate.

WFS1 variant	Relative expression (AU)				Mean relative expression (AU) (n=4)	STD error	T test (p.value)
	T1	T2	T3	T4			
Wt	1	1	1	1	1	0	-
E158K	0.87131	1.034022	1.162669	1.442606	1.127652	0.139403	0.367957
L511P	0.09062	0.280565	0.720919	0.755709	0.461953	0.189779	0.046646
R558H	0.240437	0.532039	1.170586	0.994716	0.734445	0.245622	0.300435
R611H	0.690354	0.906456	1.634955	1.524019	1.188946	0.266582	0.473075
G736A	0.632558	0.780763	1.543699	0.975726	0.983187	0.230494	0.93818
E737K	0.634184	0.972598	1.057376	0.738749	0.850727	0.114011	0.227761
R818C	0.561728	1.129074	1.77381	0.795236	1.064962	0.304151	0.82111

Appendix 8 Clone set 1 harvested in RIPA buffer. Steady-state expression in arbitrary units (AU) of *WFS1* variants relative to wild-type over four repeat experiments (T=transfection replicate), and mean values for each variant. P.values are generated from two-tailed students T-test comparing relative expression of each variant to wild-type.

T1	*WFS1*	*WFS1* background	*WFS1* minus background	β-actin	β-actin background	β-actin minus background	*WFS1* divided by β-actin
Wt	604070	367082	236988	366093	255167	110926	2.136451
E158K	548562	367082	181480	379739	255167	124572	1.456828
L511P	403511	367082	36429	387692	255167	132525	0.274884
R558H	524363	367082	157281	353449	255167	98282	1.600303
R611H	553754	367082	186672	327849	255167	72682	2.568339
G736A	474772	367082	107690	313150	255167	57983	1.857269
E737K	470033	367082	102951	306191	255167	51024	2.017698
R818C	433957	367082	66875	305189	255167	50022	1.336912
T2							
Wt	593704	440194	153510	331782	265909	65873	2.330393
E158K	580342	440194	140148	315919	265909	50010	2.8024
L511P	440194	440194	0.001	307293	265909	41384	2.42E-08
R558H	501824	440194	61630	313762	265909	47853	1.287903
R611H	510082	440194	69888	316799	265909	50890	1.373315
G736A	475748	440194	35554	318851	265909	52942	0.671565
E737K	479387	440194	39193	314706	265909	48797	0.803185
R818C	501170	440194	60976	322396	265909	56487	1.07947
T3							
Wt	150304	81344	68960	132265	76902	55363	1.245597
E158K	155795	81344	74451	113908	76902	37006	2.011863
L511P	93966	81344	12622	127387	76902	50485	0.250015
R558H	123920	81344	42576	110541	76902	33639	1.265674
R611H	185821	81344	104477	108586	76902	31684	3.297469
G736A	118216	81344	36872	113450	76902	36548	1.008865
E737K	151460	81344	70116	131765	76902	54863	1.27802
R818C	120958	81344	39614	123795	76902	46893	0.844774

Appendix 9 Clone set 2 harvested in RIPA buffer. Raw data figures for expression of *WFS1* variants in HEK293 cells normalised to β-actin. Measurements are calculated in arbitrary units of absorbance (Gel.Doc). T= transfection replicate.

WFS1 variant	Relative expression (AU) 1	2	3	Mean relative expression (AU) (n=3)	STD error	T test (p.value)
Wt	1	1	1	1	0	-
E158K	0.681892	1.202544	1.615179	1.166538	0.330703	0.600243
L511P	0.128664	0.001	0.200719	0.110128	0.071518	0.004278
R558H	0.749047	0.552655	1.016118	0.772607	0.164493	0.232519
R611H	1.202152	0.589306	2.647299	1.479586	0.747182	0.514149
G736A	0.869324	0.288177	0.809945	0.655815	0.226108	0.203288
E737K	0.944415	0.344656	1.02603	0.7717	0.263097	0.399241
R818C	0.625763	0.463213	0.678208	0.589061	0.079265	0.023918

Appendix 10 Clone set 2 harvested in RIPA buffer. Steady-state expression in arbitrary units (AU) of *WFS1* variants relative to wild-type over three repeat experiments (T=transfection replicate), and mean values for each variant. P.values are generated from two-tailed students T-test comparing relative expression of each variant to wild-type.

Set 1	*WFS1*	*WFS1* Background	WFS1 – background	β-actin	β-actin background	β-actin – background	*WFS1*/β-actin	Relative
Wt	1448210	1212280	235930	738227	647177	91050	2.591214	1
E158K	1474140	1212280	261860	734095	647177	86918	3.012725	1.162669
L511P	1360570	1212280	148290	726559	647177	79382	1.868056	0.720919
R558H	1440880	1212280	228600	722542	647177	75365	3.033238	1.170586
R611H	1518360	1212280	306080	719425	647177	72248	4.236519	1.634955
G736A	1506840	1212280	294560	720816	647177	73639	4.000054	1.543699
E737K	1413300	1212280	201020	720545	647177	73368	2.739887	1.057376
R818C	1460610	1212280	248330	701205	647177	54028	4.59632	1.77381
Set 2								
Wt	1345770	1083650	262120	632134	579345	52789	4.965428	1
E158K	1350850	1083650	267200	616647	579345	37302	7.163155	1.442606
L511P	1234910	1083650	151260	619655	579345	40310	3.752419	0.755709
R558H	1283460	1083650	199810	619799	579345	40454	4.93919	0.994716
R611H	1368510	1083650	284860	616988	579345	37643	7.56741	1.524019
G736A	1326830	1083650	243180	629538	579345	50193	4.844899	0.975726
E737K	1293640	1083650	209990	636591	579345	57246	3.668204	0.738749
R818C	1321440	1083650	237790	639565	579345	60220	3.948688	0.795236
Set 3								
Wt	150304	81344	68960	132265	76902	55363	1.245597	1
E158K	155795	81344	74451	113908	76902	37006	2.011863	1.615179
L511P	93966	81344	12622	127387	76902	50485	0.250015	0.200719
R558H	123920	81344	42576	110541	76902	33639	1.265674	1.016118
R611H	185821	81344	104477	108586	76902	31684	3.297469	2.647299
G736A	118216	81344	36872	113450	76902	36548	1.008865	0.809945
E737K	151460	81344	70116	131765	76902	54863	1.27802	1.02603
R818C	120958	81344	39614	123795	76902	46893	0.844774	0.678208

Appendix 11 *WFS1* measurements in supernatant harvested in RIPA buffer. Raw data figures for expression of *WFS1* variants in HEK293 cells normalised to β-actin. Measurements are calculated in arbitrary units of absorbance (Gel.Doc). T= transfection replicate.

Set 1	*WFS1*	*WFS1* Background	*WFS1* - background	β-actin	β-actin background	β-actin - background	*WFS1*/β-actin	Relative
Wt	651378	457469	193909	345031	291094	53937	3.595101693	1
E158K	585995	457469	128526	346368	291094	55274	2.325252379	0.64678348
L511P	561045	457469	103576	350393	291094	59299	1.746673637	0.48584819
R558H	533146	457469	75677	341926	291094	50832	1.488766918	0.41410982
R611H	682857	457469	225388	338025	291094	46931	4.802539899	1.33585648
G736A	670729	457469	213260	337613	291094	46519	4.584363378	1.27516932
E737K	633980	457469	176511	338932	291094	47838	3.689765458	1.02633132
R818C	699437	457469	241968	345738	291094	54644	4.428079936	1.2316981
Set 2								
Wt	503459	382124	121335	355194	320605	34589	3.507907138	1
E158K	503984	382124	121860	355166	320605	34561	3.525939643	1.00514053
L511P	426239	382124	44115	352538	320605	31933	1.381486237	0.39382064
R558H	428588	382124	46464	347701	320605	27096	1.714791851	0.48883616
R611H	548148	382124	166024	343490	320605	22885	7.254708324	2.06810159
G736A	546048	382124	163924	354402	320605	33797	4.850252981	1.38266288
E737K	523769	382124	141645	358266	320605	37661	3.761052548	1.07216423
R818C	499194	382124	117070	360190	320605	39585	2.957433371	0.8430763
Set 3								
Wt	299581	186026	113555	107734	76647	31087	3.652813073	1
E158K	265623	186026	79597	105116	76647	28469	2.795918367	0.76541512
L511P	276014	186026	89988	94371	76647	17724	5.07718348	1.38993794
R558H	216895	186026	30869	91931	76647	15284	2.019693797	0.55291463
R611H	310763	186026	124737	103148	76647	26501	4.706878986	1.28856278
G736A	261738	186026	75712	129554	76647	52907	1.431039371	0.39176365
E737K	292105	186026	106079	145911	76647	69264	1.531517094	0.41927059
R818C	294446	186026	108420	132596	76647	55949	1.937836244	0.53050518

Appendix 12 *WFS1* measurements in pellets harvested in RIPA buffer. Raw data figures for expression of *WFS1* variants in HEK293 cells normalised to β-actin. Measurements are calculated in arbitrary units of absorbance (Gel.Doc). T= transfection replicate.

Supernatant	1	2	3	Average	STD error
Wt	1	1	1	1	0
E158K	1.162669343	1.442605594	1.615179351	1.406818096	0.161480472
L511P	0.720919231	0.755708964	0.200718858	0.559115685	0.219816736
R558H	1.170585944	0.994715821	1.016117996	1.060473254	0.067853212
R611H	1.634955384	1.52401948	2.647299342	1.935424735	0.437693263
G736A	1.543699172	0.975726221	0.809944838	1.109790077	0.272101722
E737K	1.05737581	0.738748723	1.026029726	0.940718086	0.124176008
R818C	1.773809922	0.795236145	0.678208209	1.082418092	0.425406207

Pellets	1	2	3	Average	STD error
Wt	1	1	1	1	0
E158K	0.646783479	1.005140531	0.765415123	0.805779711	0.129087023
L511P	0.485848186	0.393820641	1.389937941	0.756535589	0.389240396
R558H	0.414109821	0.488836159	0.552914632	0.485286871	0.049123019
R611H	1.335856482	2.068101588	1.288562785	1.564173618	0.309044271
G736A	1.275169319	1.382662878	0.391763647	1.016531948	0.384473854
E737K	1.026331318	1.072164228	0.41927059	0.839255379	0.25769709
R818C	1.231698103	0.8430763	0.530505176	0.868426527	0.248394702

Super/pellets	t-tests
E158K	0.025505913
L511P	0.624419418
R558H	0.001629035
R611H	0.448862211
G736A	0.821572831
E737K	0.69445442
R818C	0.629227814

Appendix 13 Mean steady-state expression *WFS1* variants relative to wild-type in both soluble and insoluble fractions of RIPA buffer lysates (n=3).

T-Tests comparing differences between relative expression of *WFS1* variants to wild-type in the soluble and insoluble fractions of RIPA lysates.

	WFS1	*WFS1* background	*WFS1* – background	β-actin	β-Actin Background	B-actin – background	*WFS1/β*-actin	Relative
T1								
Endog	180343	179572	771	82016	59805	22211	0.03471253	0.006351
Wt	297931	179572	118359	81460	59805	21655	5.465666128	1
E158K	296416	179572	116844	80489	59805	20684	5.649004061	1.0335436
L511P	254998	179572	75426	80277	59805	20472	3.684349355	0.6740897
R558H	298207	179572	118635	79248	59805	19443	6.101681839	1.1163656
R611H	246324	179572	66752	78704	59805	18899	3.532038732	0.6462229
G736A	236555	179572	56983	80613	59805	20808	2.738514033	0.5010394
E737K	258451	179572	78879	82191	59805	22386	3.52358617	0.6446764
R818C	226190	179572	46618	82074	59805	22269	2.093403386	0.3830097
T2								
Endog	100014	99464	550	111343	95963	15380	0.035760728	0.0035795
Wt	325018	99464	225554	118540	95963	22577	9.990432741	1
E158K	282067	99464	182603	118878	95963	22915	7.968710452	0.7976342
L511P	107248	99464	7784	121817	95963	25854	0.301075269	0.0301364
R558H	186658	99464	87194	115775	95963	19812	4.401070059	0.4405285
R611H	183705	99464	84241	120219	95963	24256	3.472996372	0.3476322
G736A	325023	99464	225559	114430	95963	18467	12.21416581	1.2225863
E737K	315495	99464	216031	117286	95963	21323	10.1313605	1.0141063
R818C	356472	99464	257008	116753	95963	20790	12.36209716	1.2373936
T3								
Endog	113446	109814	3632	110179	67127	43052	0.084363096	0.0177376
Wt	329787	109814	219973	113377	67127	46250	4.756172973	1
E158K	248172	109814	138358	105003	67127	37876	3.652920055	0.7680377
L511P	143429	109814	33615	99117	67127	31990	1.050797124	0.2209333
R558H	304286	109814	194472	97434	67127	30307	6.416735408	1.3491384
R611H	255784	109814	145970	103401	67127	36274	4.024094393	0.8460782
G736A	193566	109814	83752	102719	67127	35592	2.353113059	0.4947493
E737K	238076	109814	128262	107383	67127	40256	3.186158585	0.6698996
R818C	298121	109814	188307	112301	67127	45174	4.16848187	0.8764361

Appendix 14 Raw data figures for expression of *WFS1* variants harvested from HEK293 cells in Laemmli buffer normalised to β-actin. Measurements are calculated in arbitrary units of absorbance (Gel.Doc). T= transfection replicate.

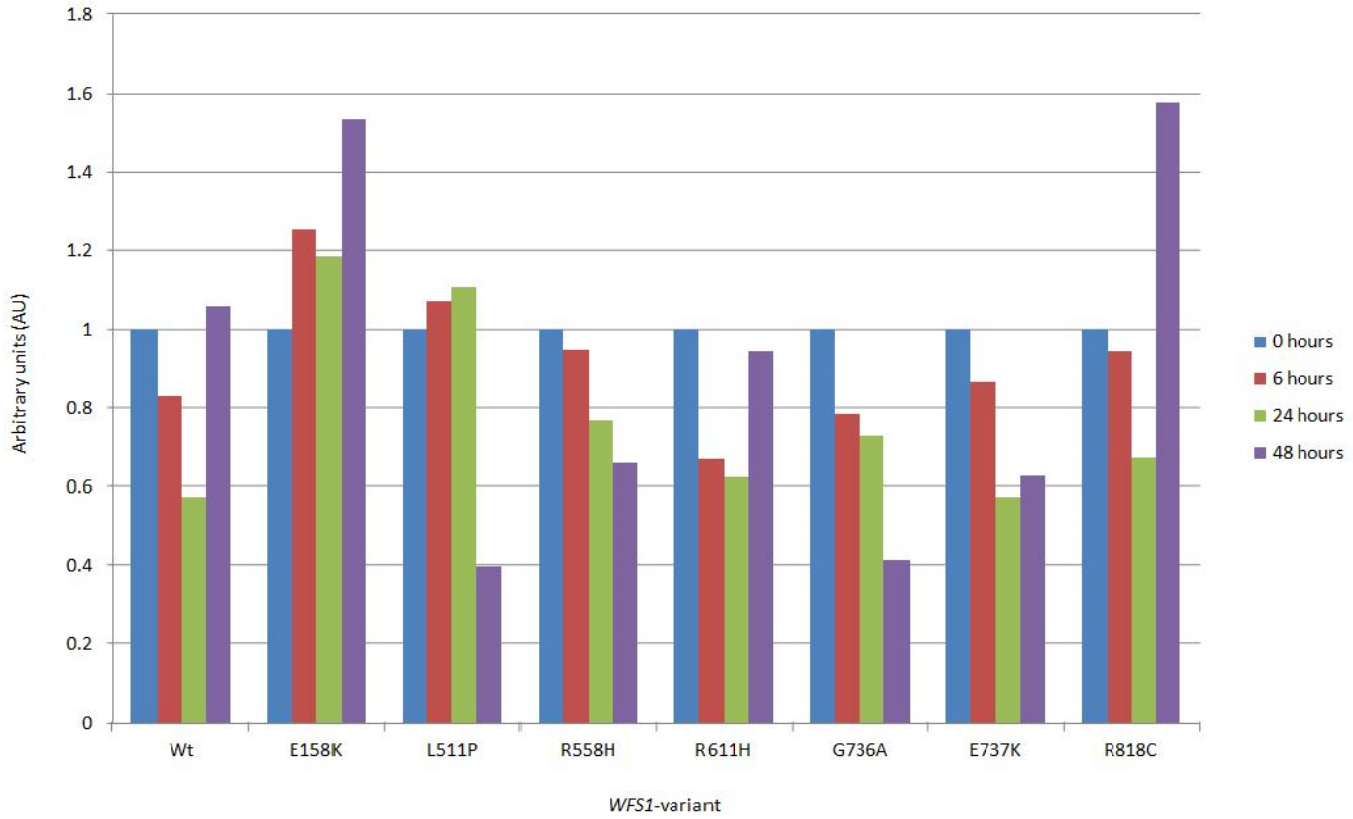
	Set 1	Set 2	Set 3	Average	STD error	t-test (P.value)	
Endog	0.006351	0.0035795	0.0177376	0.009223	0.005306	1.91164E-05	Wt/Endog
Wt	1	1	1	1	0	-	-
E158K	1.033544	0.79763416	0.76803768	0.866405	0.102884	0.252728918	Wt/E158K
L511P	0.67409	0.03013636	0.220933328	0.308386	0.233886	0.068499925	Wt/L511P
R558H	1.116366	0.44052847	1.349138361	0.968677	0.33373	0.918985704	Wt/R558H
R611H	0.646223	0.34763223	0.846078226	0.613311	0.177376	0.116305209	Wt/R611H
G736A	0.501039	1.22258626	0.49474926	0.739458	0.295863	0.393588527	Wt/G736A
E737K	0.644676	1.01410627	0.669899645	0.776227	0.145943	0.201185945	Wt/E737K
R818C	0.38301	1.23739356	0.876436138	0.83228	0.303278	0.568051145	Wt/R818C

Appendix 15 WFS1 variants harvested in Laemmli buffer. Steady-state expression in arbitrary units (AU) of *WFS1* variants relative to wild-type over three repeat experiments (set=transfection replicate), and mean values for each variant. P.values are generated from two-tailed students T-test comparing relative expression of each variant to wild-type.

***WFS1*-variant** **Hours post treatment**					
Wt	**Set 1**	**Set 2**	**Set 3**	**Average relative**	**STD error**
0	1	1	1	**1**	0
6	0.768847383	0.9859312	0.732699082	**0.829159222**	0.096849793
24	0.658128233	0.700956232	0.370449547	**0.576511337**	0.127091808
48	2.420536425	0.704978011	0.058126674	**1.061213703**	0.863256294
E158K					
0	1	1		**1**	0
6	1.301070297	1.208990922		**1.255030609**	0.06510995
24	1.472621649	0.899462588	-	**1.186042118**	0.405284659
48	2.116532954	0.953876319		**1.535204636**	0.822122391
L511P					
0	1	1	1	**1**	0
6	0.980609438	1.013540794	1.230173161	**1.074774465**	0.095871489
24	1.34304584	0.969642505	1.011423493	**1.108037279**	0.144668899
48	0.387825708	0.505766599	0.304675616	**0.399422641**	0.071450201
R558H					
0	1	1		**1**	0
6	0.745715371	1.152012452		**0.948863912**	0.287295421
24	0.643431822	0.892678365	-	**0.768055093**	0.17624392
48	0.20846351	1.110284512		**0.659374011**	0.637683746
R611H					
0	1	1		**1**	0
6	0.42797281	0.909887883		**0.668930346**	0.340765416
24	0.271100271	0.977749903	-	**0.624425087**	0.499676747
48	0.828572295	1.060518798		**0.944545547**	0.164010945
G736A					
0	1	1		**1**	0
6	0.667438556	0.899251381		**0.783344969**	0.16391642
24	0.395700715	1.061940093	-	**0.728820404**	0.471102382
48	0.370100941	0.459314708		**0.414707825**	0.06308366
E737K					
0	1	1	1	**1**	0
6	0.966247794	0.953689989	0.676983878	**0.865640554**	0.11561343
24	0.926410351	0.472780774	0.32337428	**0.574188468**	0.222065042
48	0.879797755	0.606654424	0.399020776	**0.628490985**	0.170505502
R818C					
0	1	1	1	**1**	0
6	0.627761841	0.929973215	1.274178915	**0.943971324**	0.22870365
24	0.638482281	0.445992492	0.93222282	**0.672232531**	0.173146315
48	0.451263452	3.668213582	0.619465699	**1.579647578**	1.280362051

Appendix 16A. Pilot data (n=2/3) showing relative stability of *WFS1* variants over 48 hours. HEK293 cells transfected with *WFS1* constructs were treated with 100µM cyclohexamide and harvested at time intervals of 0, 6, 24 and 48 hours. Following SDS-PAGE and western blot analysis the relative stability of each variant was calculated by comparing protein levels (measured in arbitrary units) at each time interval relative to 0 hours.

Marked inconsistencies (see standard error) in results suggested that extended chase-periods following cyclohexamide treatment was unsuitable for our system.



Appendix 16B. Graphical representation of pilot data (n=2/3) showing relative stability of *WFS1* variants over 48 hours (see appendix 15A for details).

WFS1 variant	1	2	3	Average	STD error
Wt	0.768847	0.985931	0.732699	0.829159	0.09685
E158K	1.208991	1.30107	-	1.255031	0.06511
L511P	0.980609	1.013541	1.230173	1.074774	0.095871
R558H	0.745715	1.152012	-	0.948864	0.287295
R611H	0.427973	0.909888	-	0.66893	0.340765
G736A	0.667439	0.899251	-	0.783345	0.163916
E737K	0.966248	0.95369	0.676984	0.865641	0.115613
R818C	0.627762	0.929973	1.274179	0.943971	0.228704

Appendix 17 Pilot data (n=2/3) showing relative stability measurements of *WFS1* variants 6 hours following 100µM cyclohexamide treatment.

WFS1 variant	set 1	set 2	set 3	set 4	Average	STD error
Endog	1.30889086	1.330133839	1.549321185	-	1.396115296	0.094119204
Wt-pcDNA	1.15149678	2.168606701	1.523611772	-	1.614571753	0.363891079
L511P-pcDNA	1.84683522	1.285692865	0.340690191	-	1.157739427	0.538236601
R611H-pcDNA	1.34663221	0.691437654	3.175253682	-	1.737774514	0.910241656
Wt-myc	1.95434088	0.903201771	0.970127634	1.076638979	1.226077315	0.283325193
N-term-myc	0.09638927	0.215099071	0.29131433	0.050078352	0.163220254	0.063567091

Appendix 18 Expression of *WFS1* variants 6 hours following cyclohexamide treatment. Measured in arbitrary units relative to expression at 0 hours.

t-test	Comparisons
0.542720412	Endog v Wt-pcDNA
0.366456755	Wt-pcDNA v Wt-myc
0.443946025	Wt-pcDNA v L511P
0.888728592	Wt-pcDNA v R611H
0.019984106	Wt-myc v N-term-myc

Appendix 19 Two-tailed students T-test results comparing the differences in relative expression of variants 6 hours after cyclohexamide treatment.

WFS1 variant	Hours	*WFS1*	*WFS1* background	*WFS1* – background	β-actin	β-actin background	β-actin background	*WFS1*/β-actin	Relative
Endog									
set 1	0	233497	108145	125352	105399	68673	36726	3.413167783	1
	3	244283	101660	142623	104692	68673	36019	3.959660179	1.16011296
	6	262553	97851	164702	105540	68673	36867	4.467464128	1.308890864
set 2	0	107313	84674	22639	158624	127697	30927	0.732014098	1
	3	134486	84674	49812	162496	127697	34799	1.431420443	1.955454748
	6	118889	84674	34215	162837	127697	35140	0.973676722	1.330133839
set 3	0	131672	88451	43221	160204	127130	33074	1.30679688	1
	3	146551	88451	58100	161949	127130	34819	1.668629197	1.276884895
	6	158929	88451	70478	161940	127130	34810	2.02464809	1.549321185
Wt-pcDNA									
set 1	0	225168	103214	121954	108015	63299	44716	2.72730119	1
	3	246271	107333	138938	113641	63299	50342	2.759882404	1.011946321
	6	271007	106157	164850	115791	63299	52492	3.140478549	1.151496784
set 2	0	268418	225444	42974	145083	112392	32691	1.314551406	1
	3	250544	215026	35518	140742	112392	28350	1.252839506	0.953054784
	6	275529	209138	66391	135681	112392	23289	2.850744987	2.168606701
set 3	0	124155	86174	37981	162606	116874	45732	0.830512551	1
	3	138378	86174	52204	160432	116874	43558	1.198493962	1.443077483
	6	138834	86174	52660	158490	116874	41616	1.2653787	1.523611772

Appendix 20 Raw data for *WFS1* expression following 100μM cyclohexamide treatment normalised with β-actin measurements. Figures are arbitrary units of absorbance (Gel.Doc).

WFS1 variant	Hours	*WFS1*	*WFS1* background	*WFS1* – background	β-actin	β-actin background	β-actin background	*WFS1*/β-actin	Relative
L511P-pcDNA									
set 1	0	212412	122873	89539	116122	94554	21568	4.151474407	1
	3	229170	114601	114569	112044	94554	17490	6.550543168	1.577883548
	6	245366	106239	139127	112700	94554	18146	7.667089166	1.846835224
set 2	0	160875	100956	59919	123142	81307	41835	1.432269631	1
	3	154721	102927	51794	127560	81307	46253	1.119797635	0.781834377
	6	191655	103195	88460	129345	81307	48038	1.841458845	1.285692865
set 3	0	198938	118997	79941	208617	151821	56796	1.407511092	1
	3	164422	118997	45425	221362	151821	69541	0.653211774	0.464089966
	6	153741	118997	34744	224276	151821	72455	0.479525223	0.340690191
R611H-pcDNA									
set 1	0	185678	134330	51348	166301	102532	63769	0.805218837	1
	3	162838	135211	27627	151579	102532	49047	0.563276041	0.699531625
	6	187664	142135	45529	144520	102532	41988	1.084333619	1.346632207
set 2	0	246451	152866	93585	129032	91577	37455	2.498598318	1
	3	224692	158057	66635	142751	91577	51174	1.30212608	0.521142622
	6	249272	168566	80706	138292	91577	46715	1.72762496	0.691437654
set 3	0	146609	125508	21101	161194	134337	26857	0.785679711	1
	3	156986	125509	31477	170687	134337	36350	0.865942228	1.10215679
	6	219991	125508	94483	172210	134337	37873	2.494732395	3.175253682

Appendix 21 Raw data for *WFS1* expression following 100μM cyclohexamide treatment normalised with β-actin measurements. Figures are arbitrary units of absorbance (Gel.Doc).

WFS1 variant	Hours	*WFS1*	*WFS1* background	*WFS1* – background	β-actin	β-actin background	β-actin background	*WFS1*/β-actin	Relative
Wt-myc									
set 1	0	188460	145945	42515	143995	111084	32911	1.291817326	1
	3	190314	133003	57311	141177	111084	30093	1.904462832	1.474250882
	6	195448	127189	68259	138121	111084	27037	2.524651404	1.954340876
set 2	0	189901	118438	71463	137818	107377	30441	2.347590421	1
	3	190581	117363	73218	146561	107377	39184	1.868568804	0.795951793
	6	206546	121202	85344	147627	107377	40250	2.120347826	0.903201771
set 3	0	153692	129729	23963	138370	113564	24806	0.966016286	1
	3	135786	129729	6057	145011	113564	31447	0.192609788	0.199385653
	6	158422	129729	28693	144181	113564	30617	0.937159095	0.970127634
set 4	0	190670	140326	50344	144583	112020	32563	1.546049197	1
	3	192223	140326	51897	148388	112020	36368	1.42699626	0.922995376
	6	201025	140326	60699	148486	112020	36466	1.664536829	1.076638979
N-term-myc									
set 1	0	584761	435665	149096	145471	131925	13546	11.00664403	1
	3	472449	435665	36784	141814	124517	17297	2.126611551	0.193211623
	6	461891	435665	26226	141952	117232	24720	1.06092233	0.096389265
set 2	0	848160	477955	370205	134844	116417	18427	20.09035654	1
	3	618563	477955	140608	140126	123603	16523	8.509834776	0.423578086
	6	544803	477955	66848	149333	133864	15469	4.321417028	0.215099071
set 3	0	189123	45322	143801	105131	76233	28898	4.97615752	1
	3	68165	45322	22843	104221	76233	27988	0.816171216	0.164016355
	6	82531	45322	37209	101901	76233	25668	1.449625993	0.29131433
set 4	0	218767	65822	152945	101625	75488	26137	5.85166622	1
	3	169744	65822	103922	99052	75488	23564	4.410202003	0.753666022
	6	72952	65822	7130	99819	75488	24331	0.293041799	0.050078352

Appendix 22 Raw data for *WFS1* expression following 100μM cyclohexamide treatment normalised with β-actin measurements. Figures are arbitrary units of absorbance (Gel.Doc).

WFS1 variant	Hours	set 1	set 2	set 3	set 4	average	STD error
Endog							
	0	1	1	1		1	0
	3	1.16011296	1.955454748	1.276884895	-	1.464150868	0.303680382
	6	1.308890864	1.330133839	1.549321185		1.396115296	0.094119204
Wt-pcDNA							
	0	1	1	1		1	0
	3	1.011946321	0.953054784	1.443077483	-	1.136026196	0.189179046
	6	1.151496784	2.168606701	1.523611772		1.614571753	0.363891079
L511p-pcDNA							
	0	1	1	1		1	0
	3	1.577883548	0.781834377	0.464089966	-	0.941269297	0.405708428
	6	1.846835224	1.285692865	0.340690191		1.157739427	0.538236601
R611H-pcDNA							
	0	1	1	1		1	0
	3	0.699531625	0.521142622	1.10215679	-	0.774277012	0.210457265
	6	1.346632207	0.691437654	3.175253682		1.737774514	0.910241656
Wt-myc							
	0	1	1	1	1	1	0
	3	1.474250882	0.795951793	0.199385653	0.922995376	0.848145926	0.302071363
	6	1.954340876	0.903201771	0.970127634	1.076638979	1.226077315	0.283325193
N-term-myc							
	0	1	1	1	1	1	0
	3	0.193211623	0.423578086	0.164016355	0.753666022	0.383618021	0.157413757
	6	0.096389265	0.215099071	0.29131433	0.050078352	0.163220254	0.063567091

Appendix 23 Average relative expression of *WFS1* variants following 100µM cyclohexamide treatment

Untransfected/ Wt WFS1	del ct Untransfected	del ct Wt WFS1	del-del Ct	Fold change	Average Fold change	Std error	p.value
1	6.6615	0.2055	6.456	87.79093047	155.1989716	20.88409703	1.77396E-13
2	7.0275	0.2335	6.794	110.9680081			
3	7.0815	-0.0085	7.09	136.2393834			
4	6.883	-0.2215	7.1045	137.6155798			
5	6.8555	-1.1885	8.044	263.9278896			
6	6.7705	-1.347	8.1175	277.7224513			
7	7.1905	1.027	6.1635	71.68006299			
8	7.285	1.0235	6.2615	76.71836132			
9	7.372	0.1895	7.1825	145.2606368			
10	7.2565	0.0565	7.2	147.0333894			
11	6.9375	-0.5875	7.525	184.1834982			
12	6.9745	-0.828	7.8025	223.2474676			

Untransfected/ WFS1-L511P	del ct Untransfected	del ct WFS1-L511P	del-del Ct	Fold change	Average Fold change	Std error	p.value
1	6.6615	1.518	5.1435	35.34661119	94.02548345	17.2746348	1.22345E-11
2	7.0275	1.5135	5.514	45.69612735			
3	7.0815	0.789	6.2925	78.38469047			
4	6.883	0.764	6.119	69.5028388			
5	6.8555	-0.1545	7.01	128.8903104			
6	6.7705	-0.027	6.7975	111.237545			
7	7.1905	1.024	6.1665	71.82927257			
8	7.285	0.777	6.508	91.01295428			
9	7.372	1.8475	5.5245	46.02991904			
10	7.2565	1.679	5.5775	47.75235569			
11	6.9375	-0.7055	7.643	199.8813421			
12	6.9745	-0.689	7.6635	202.7418345			

Appendix 24 Del-del-Ct and fold-change calculations from Taqman real-time PCR measurements of wild-type and mutant *WFS1*.

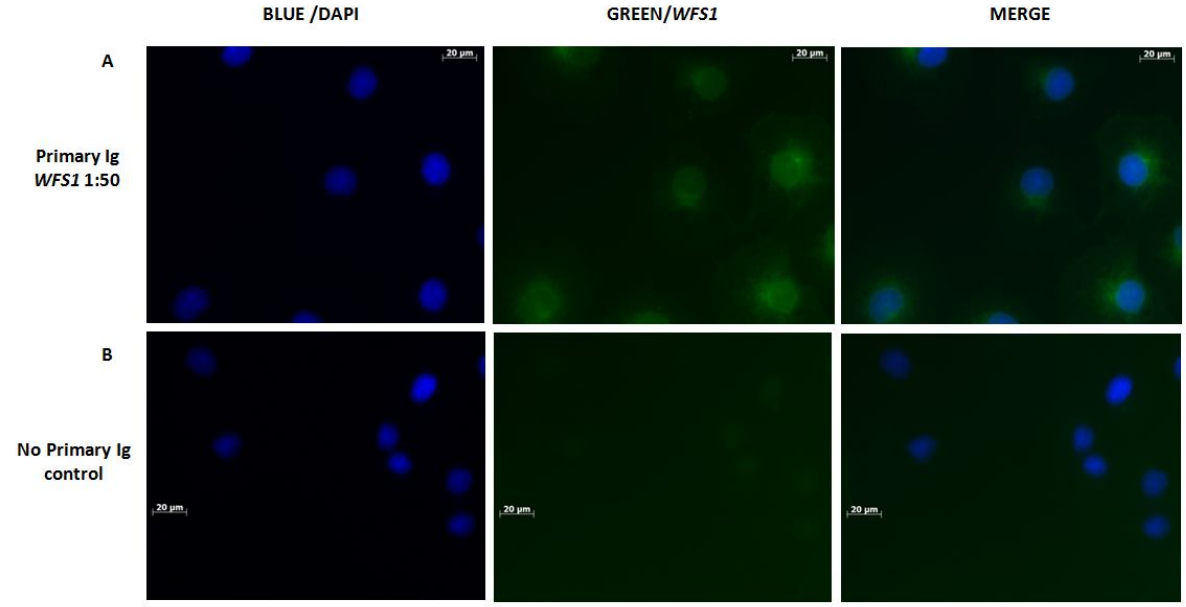
Empty pcDNA/ Wt-WFS1	del ct Empty	del ct Wt WFS1	del-del Ct	fold change	Average fold change	Std error	p.value
1	7.2395	0.2055	7.034	131.0524033	153.8135544	24.34088061	1.52752E-15
2	7.2815	0.2335	7.048	132.3303342			
3	6.5985	-0.0085	6.607	97.47767934			
4	6.2735	-0.2215	6.495	90.19652833			
5	7.103	-1.1885	8.2915	313.3215087			
6	6.711	-1.347	8.058	266.501529			
7	7.3555	1.027	6.3285	80.36525509			
8	7.577	1.0235	6.5535	93.92908173			
9	6.93	0.1895	6.7405	106.9283058			
10	6.5425	0.0565	6.486	89.63560454			
11	7.3275	-0.5875	7.915	241.3528412			
12	6.835	-0.828	7.663	202.6715817			

Empty pcDNA/ WFS1-L511P	del ct Empty	del ct WFS1-L511P	del-del Ct	Fold change	Average fold change	Std error	p.value
1	7.2395	1.518	5.7215	52.76465711	97.46553183	21.35194539	6.12564E-13
2	7.2815	1.5135	5.768	54.4930373			
3	6.5985	0.789	5.8095	56.0833258			
4	6.2735	0.764	5.5095	45.55381577			
5	7.103	-0.1545	7.2575	153.0118949			
6	6.711	-0.027	6.738	106.7431736			
7	7.3555	1.024	6.3315	80.53254381			
8	7.577	0.777	6.8	111.4304721			
9	6.93	1.8475	5.0825	33.88324164			
10	6.5425	1.679	4.8635	29.11115147			
11	7.3275	-0.7055	8.033	261.9231923			
12	6.835	-0.689	7.524	184.0558762			

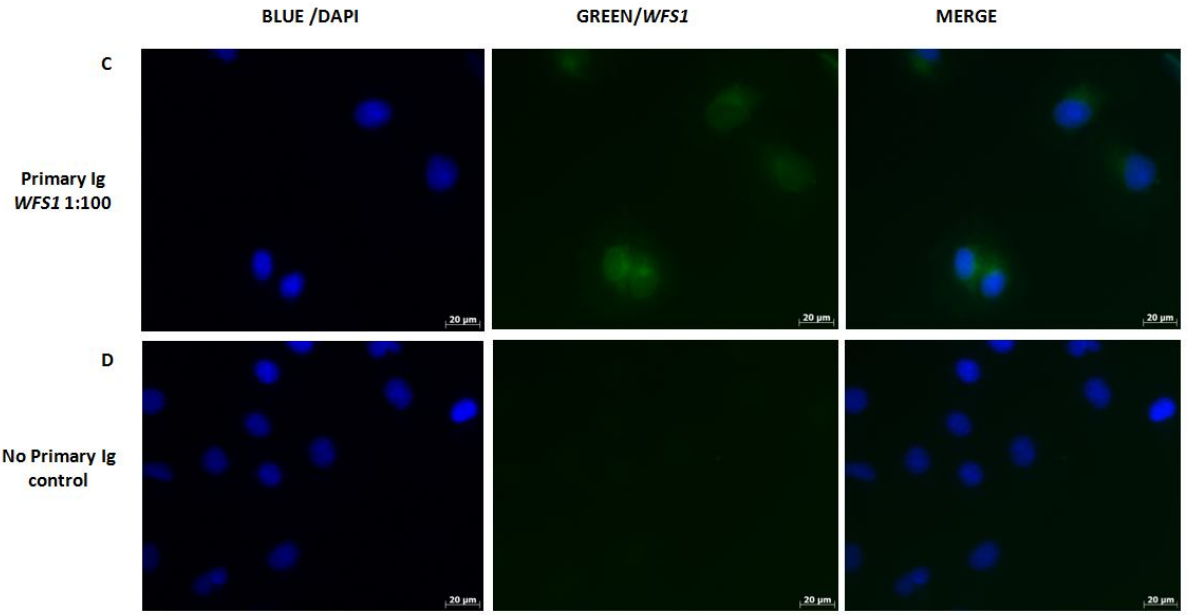
Appendix 25 Del-del-Ct and fold-change calculations from Taqman real-time PCR measurements of wild-type and mutant *WFS1*.

Untrans/Empty pcDNA3.1	del ct Untransfected	del ct Empty	del-del Ct	Fold change	Average fold change	Std error	p.value
1	6.6615	7.2395	-0.578	0.669891801	1.074039932	0.098778147	0.743596358
2	7.0275	7.2815	-0.254	0.838568184			
3	7.0815	6.5985	0.483	1.397646972			
4	6.883	6.2735	0.6095	1.525730339			
5	6.8555	7.103	-0.2475	0.842354841			
6	6.7705	6.711	0.0595	1.042104532			
7	7.1905	7.3555	-0.165	0.891928519			
8	7.285	7.577	-0.292	0.816768991			
9	7.372	6.93	0.442	1.358486285			
10	7.2565	6.5425	0.714	1.640345822			
11	6.9375	7.3275	-0.39	0.763129604			
12	6.9745	6.835	0.1395	1.101523291			
Wt-WFS1/ WFS1-L511P	**del ct Wt WFS1**	**del ct WFS1-L511P**	**del-del Ct**	**Fold change**	**Average Fold change**	**Std error**	**p.value**
1	0.2055	1.518	-1.3125	0.402622583	0.633928213	0.095959626	0.025712256
2	0.2335	1.5135	-1.28	0.411795509			
3	-0.0085	0.789	-0.7975	0.575345312			
4	-0.2215	0.764	-0.9855	0.505050656			
5	-1.1885	-0.1545	-1.034	0.488354264			
6	-1.347	-0.027	-1.32	0.400534939			
7	1.027	1.024	0.003	1.002081605			
8	1.0235	0.777	0.2465	1.186325577			
9	0.1895	1.8475	-1.658	0.316878131			
10	0.0565	1.679	-1.6225	0.324772189			
11	-0.5875	-0.7055	0.118	1.085229372			
12	-0.828	-0.689	-0.139	0.908148418			

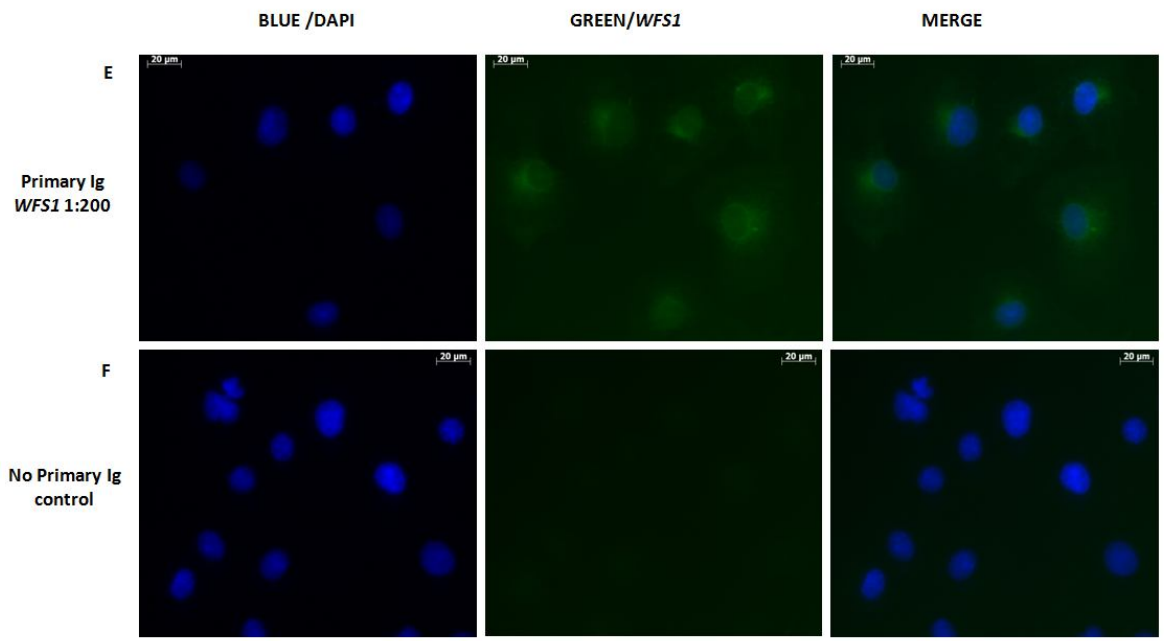
Appendix 26 Del-del-Ct and fold-change calculations from Taqman real-time PCR measurements of wild-type and mutant *WFS1*.



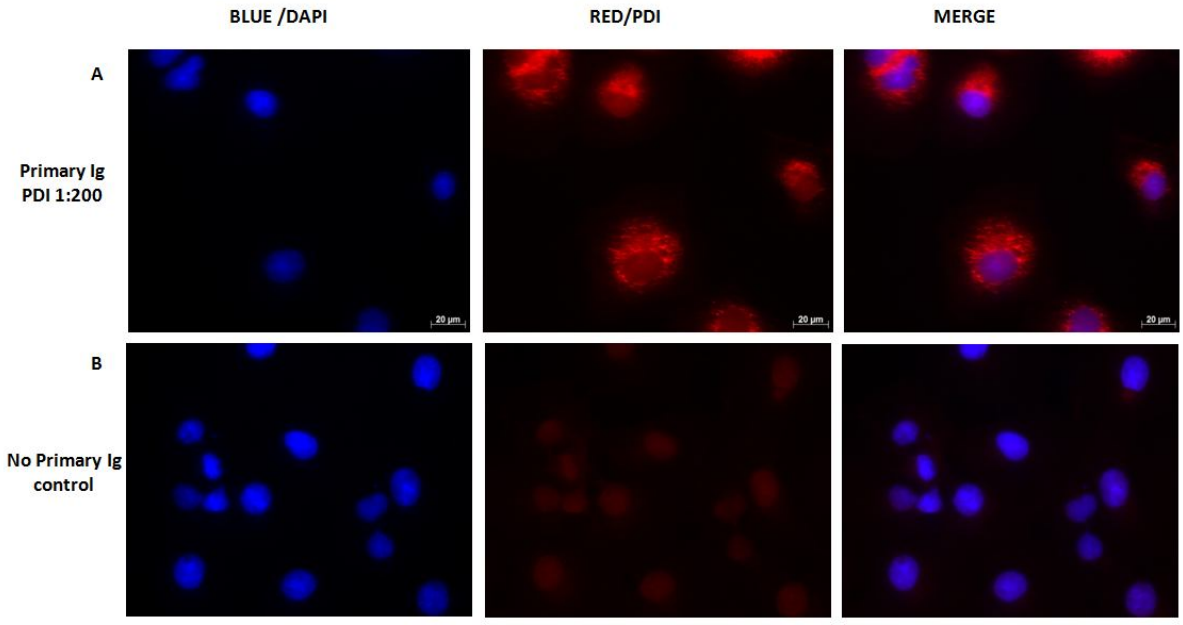
Appendix 27 Antibody optimisation. COS7 cells fixed and stained with a dilution series of anti-*WFS1* Ig (x40 mag). *WFS1* Ig 1:50



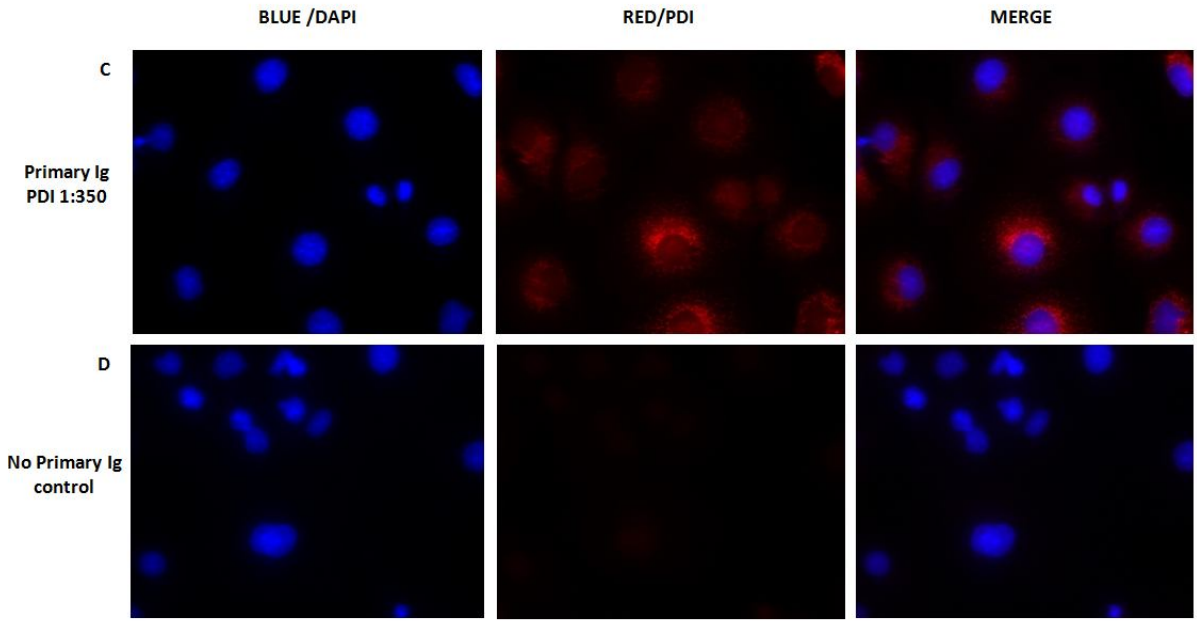
Appendix 27 Antibody optimisation. COS7 cells fixed and stained with a dilution series of anti-*WFS1* Ig (x40 mag). WFS1-Ig 1:100



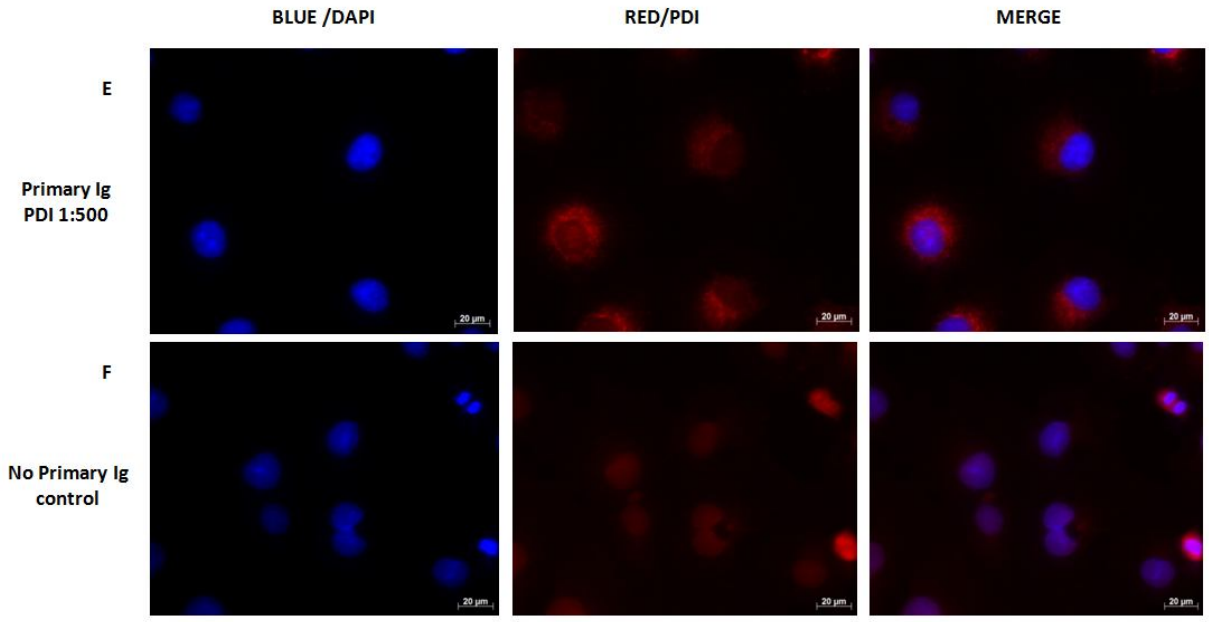
Appendix 27 Antibody optimisation. COS7 cells fixed and stained with a dilution series of anti-*WFS1* Ig (x40 mag). *WFS1* Ig 1:200



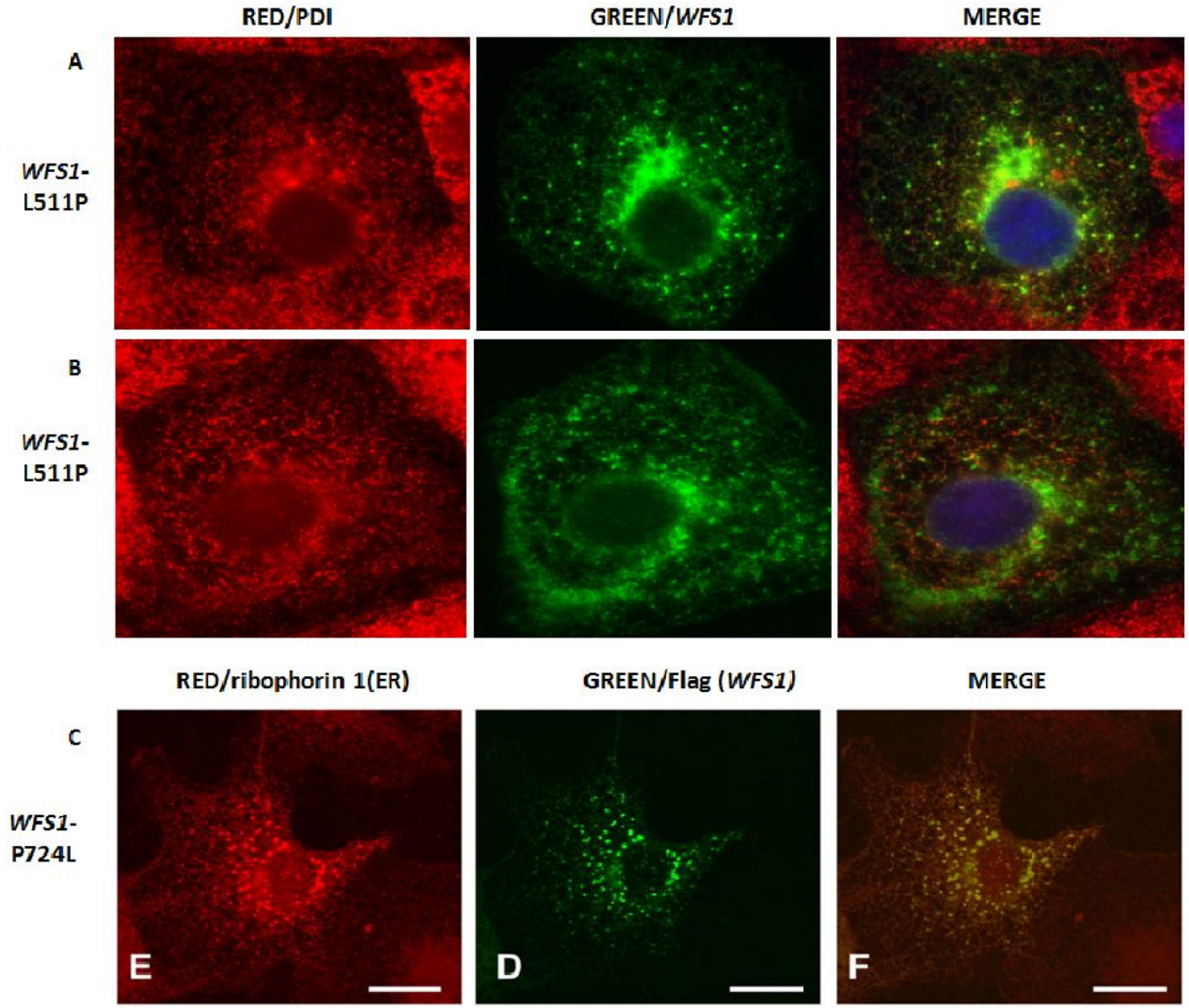
Appendix 28 Antibody optimisation. COS7 cells fixed and stained with a dilution series of anti-PDI Ig (x40 mag). PDI Ig 1:200



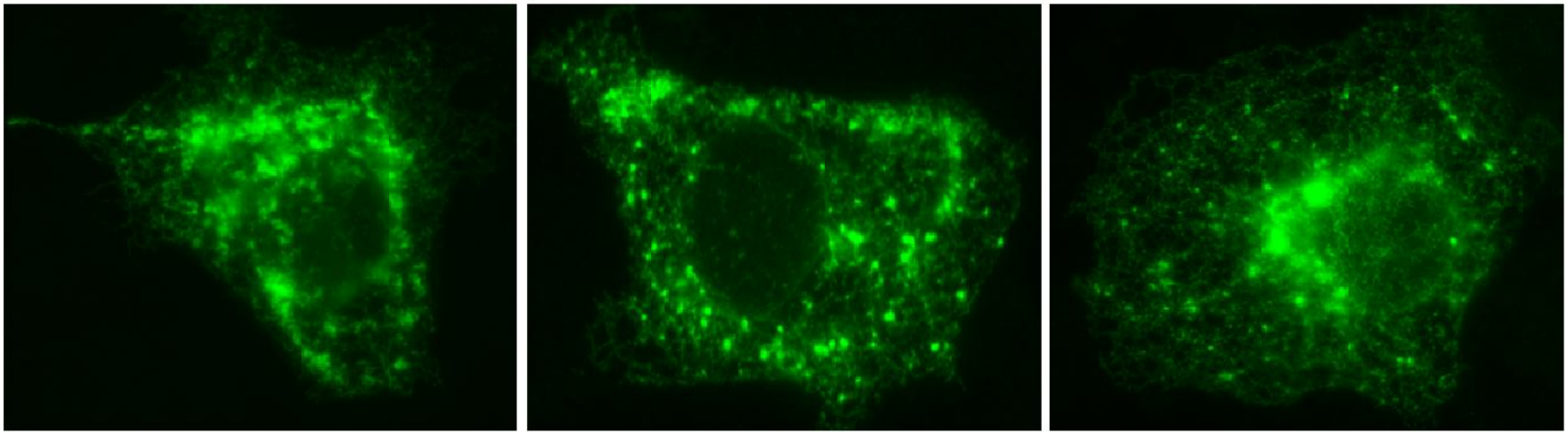
Appendix 28 Antibody optimisation. COS7 cells fixed and stained with a dilution series of anti-PDI Ig (x40 mag). PDI Ig 1:350



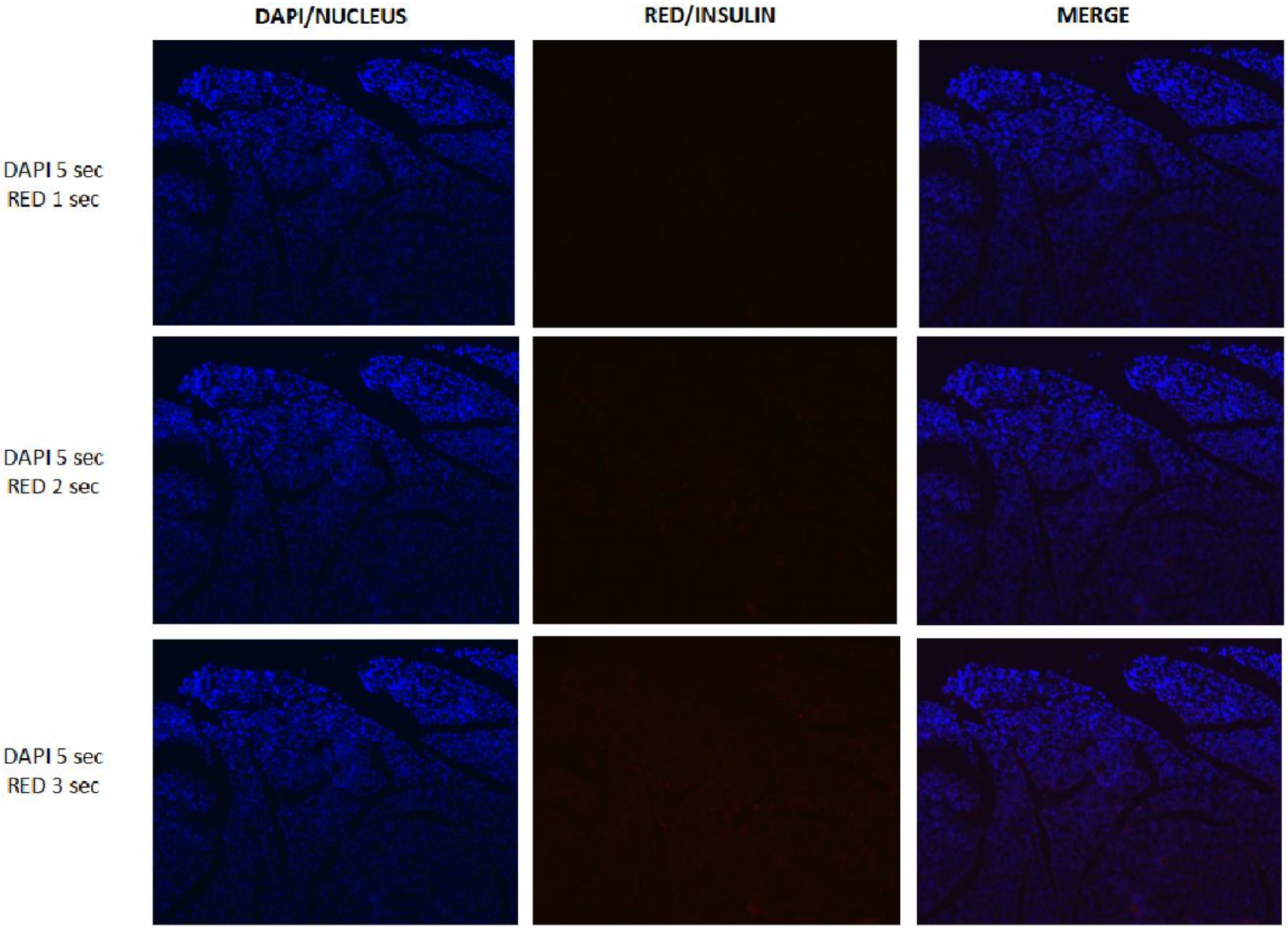
Appendix 28 Antibody optimisation. COS7 cells fixed and stained with a dilution series of anti-PDI Ig (x40 mag). PDI Ig 1:500



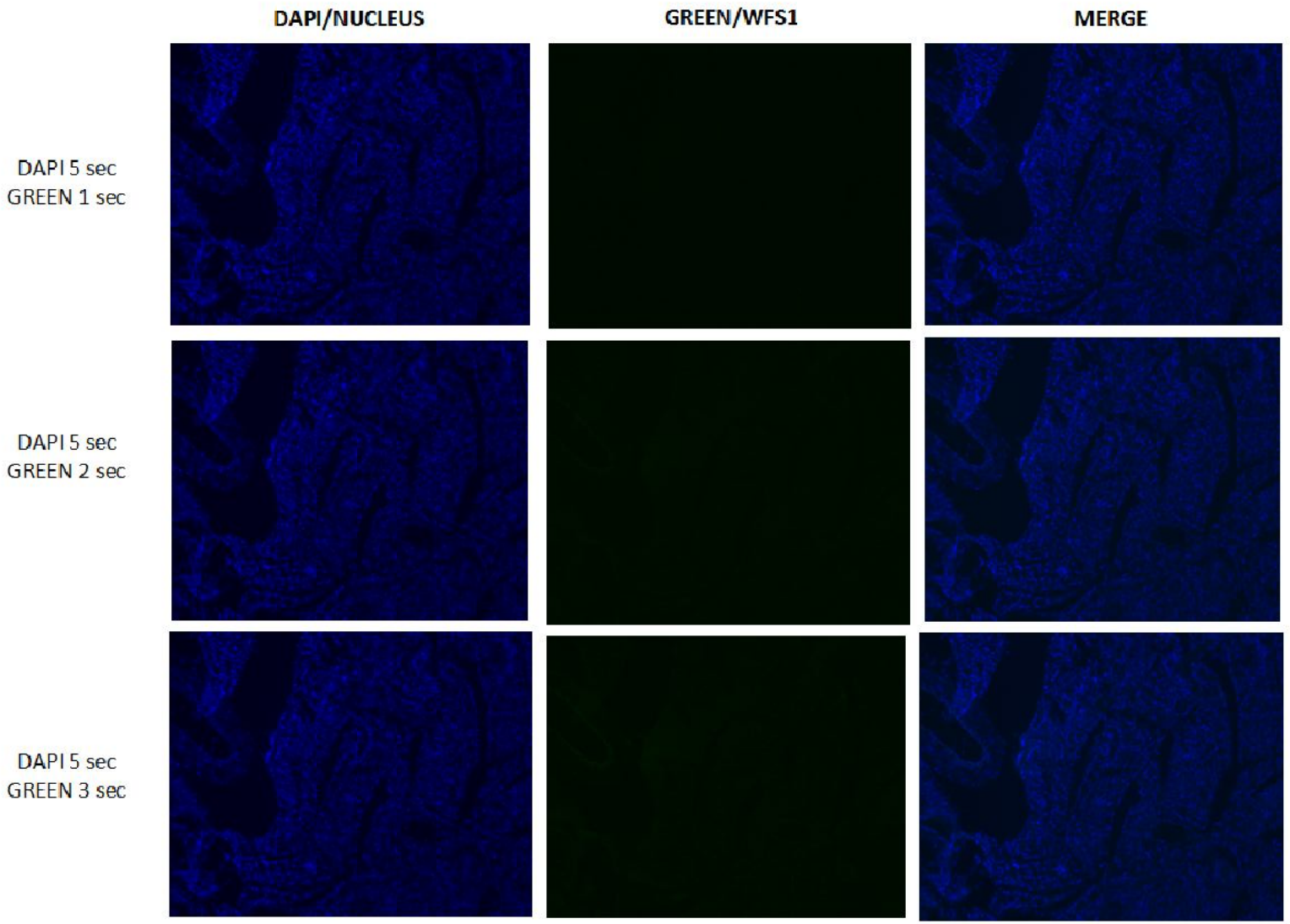
Appendix 29 COS7 cells fixed and stained for markers of the ER and presence of *WFS1* (x100), contrast adjusted. **(A** and **B)** Examples of over-expressed *WFS1*-L511P. **(C)** Similar patterns of *WFS1*-P742L expression reported in previous study (From Fonseca, 2005).



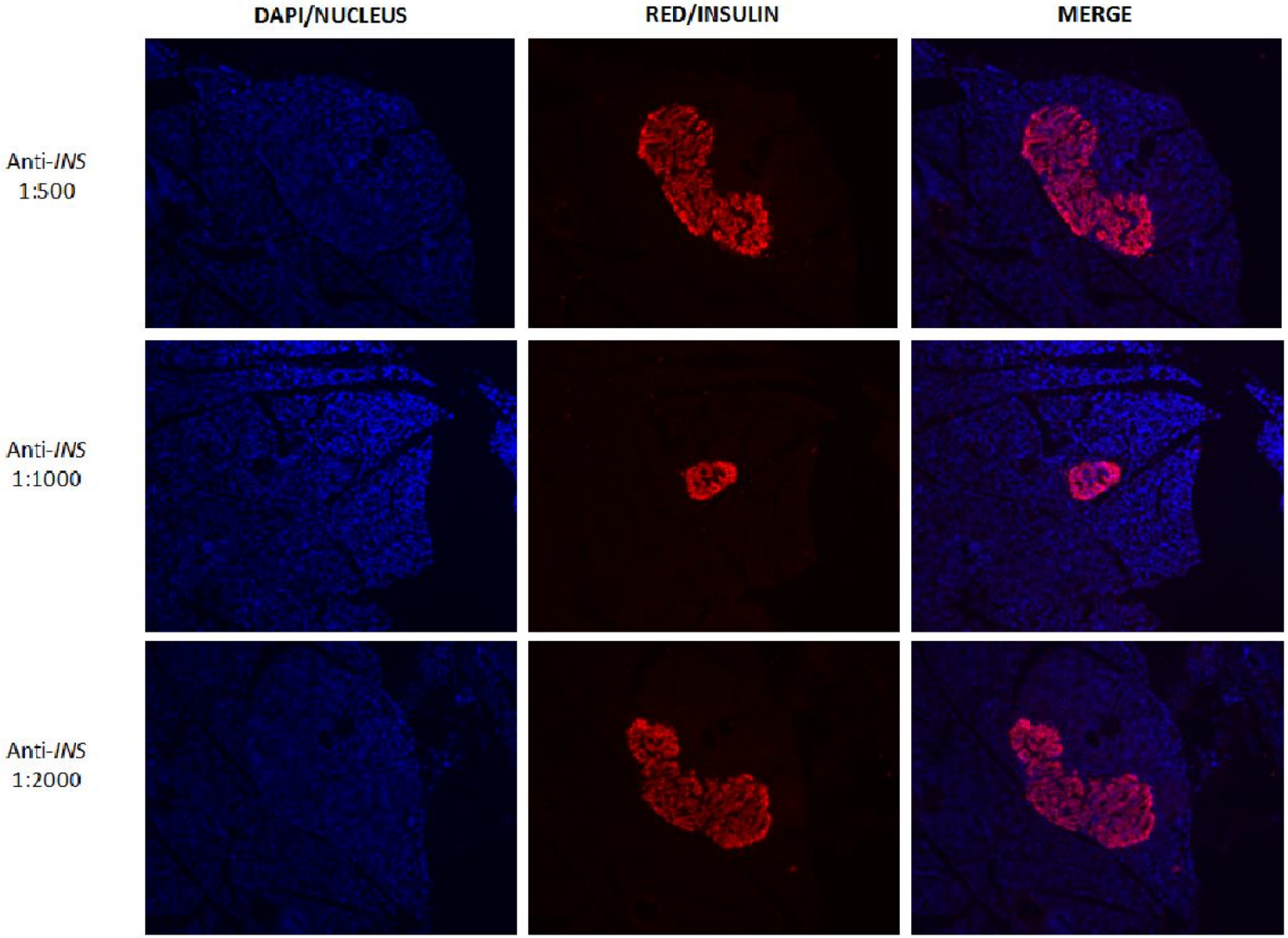
Appendix 30 COS7 cells fixed and stained with anti-*WFS1* Ig (x100), contrast adjusted. Further examples of over-expressed *WFS1*-L511P which produces a reticular, punctuate pattern of staining.



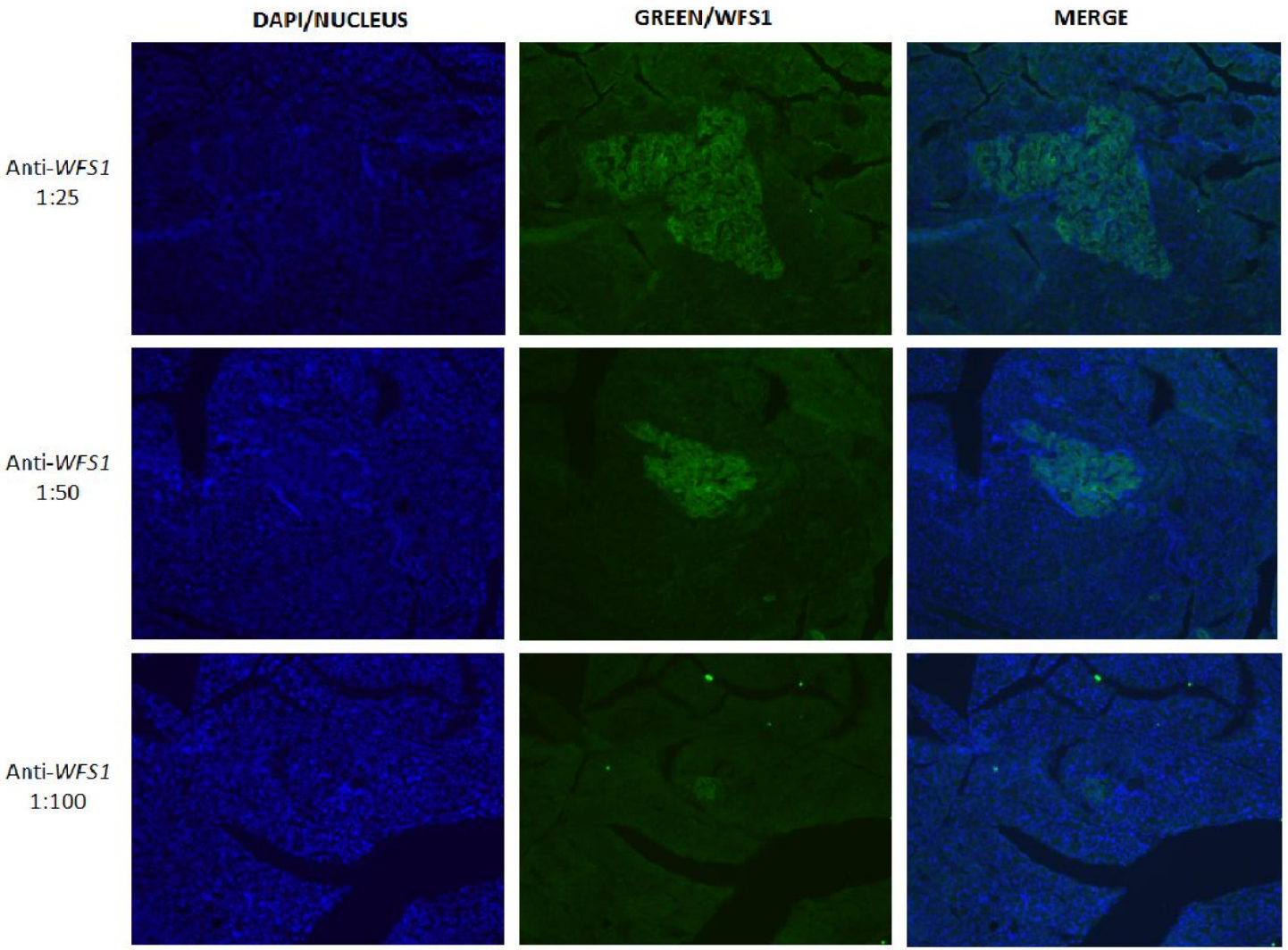
Appendix 31 Antibody optimisation for immunohistochemistry of rat pancreas (x100mag). No primary controls with increasing RED channel exposure settings.



Appendix 32 Antibody optimisation for immunohistochemistry of rat pancreas (x100mag). No primary controls with increasing GREEN channel exposure settings.



Appendix 33 Antibody optimisation for immunohistochemistry of rat pancreas (x100mag). Anti-insulin antibody dilution series, exposure red 1 sec.



Appendix 34 Antibody optimisation for immunohistochemistry of rat pancreas (x100mag). Anti-*WFS1* antibody dilution series, exposure green 3 sec.