Investigations of γ-secretase substrates Amyloid precursor protein and p75 neurotrophin receptor in zebrafish

Tanya Jayne

School of Biological Sciences

Discipline of Genetics and Evolution

A thesis submitted for the degree of Doctor of Philosophy

The University of Adelaide

November 2019

Table of Contents

Abstract

γ-secretase is an important protease complex responsible for the cleavage of over 100 substrates within their transmembrane domains. Among these substrates is the AMYLOID BETA (A4) PRECURSOR PROTEIN (APP), which is most well-known for its involvement in Alzheimer's disease (AD) pathology and the p75 neurotrophin receptor (p75^{NTR}).

This thesis aimed to investigate γ-secretase and its substrates, p75^{NTR} and APP, in a number of ways; Aim 1) to dissect the structural determinants contributing to selection of substrates by γ-secretase using an assay-based system, Aim 2) to investigate the molecular effects of fAD-like and null mutations in zebrafish *appa* and *appb*, by generating mutations in endogenous genes with genome editing technologies, and Aim 3) to investigate the aggregation propensity of Aβ-42-like peptides of Appa and Appb using predictive software.

How γ-secretase cleavage site specificity is determined is still unclear. A previous study investigating the proteolytic processing of p75^{NTR} and its homolog NRH1 (neurotrophin receptor homolog 1) found that transmembrane cleavage of NRH1 was not sensitive to the γ-secretase inhibitor DAPT, suggesting that it is not processed by γ-secretase. We have previously identified zebrafish orthologues of the *p75^{NTR}* and *Nrh1* genes and have developed *in-vivo* assays to assess cleavage of the p75^{NTR} and Nrh1 proteins.

To address Aim 1, we first improved upon our previous assay system by switching out the internal reference standard mCherry for a second GFP sequence. A chimeric construct was then designed, in which the Nrh1 transmembrane domain was replaced with the transmembrane domain of p75^{NTR}, to allow us to determine whether this domain can confer γ-secretase cleavage susceptibility to Nrh1. Our results indicate that the p75^{NTR} transmembrane domain alone is not sufficient to confer γ-secretase susceptibility to Nrh1.

Missense mutations in the APP gene cause approximately 15% of dominantly inherited familial Alzheimer's Disease (fAD). Currently 59 mutations are known and are listed on Alzforum (https://www.alzforum.org/mutations/app). The human APP gene has two co-orthologues in zebrafish, *appa* and *appb*. To address Aim 2, we attempted to utilise genome editing technology to generate fAD-like and null mutations in these zebrafish genes. Despite the various challenges associated with this project, a putative null mutation of the zebrafish *appb* gene was ultimately generated.

Hypoxia is thought to be a risk factor for AD. A previous study in our laboratory measured the hypoxic response of 6-month-old zebrafish carrying fAD-like mutations in *psen1* and observed an increased hypoxic response in these mutants under normoxia and a further increase under hypoxia. We measured the hypoxic response in our 6-month-old heterozygous *appb* putative null mutant zebrafish. However, there was no observable difference in the hypoxic response of our mutants compared to 6-month-old wildtypes. In future, our *appb* null mutation should serve to elucidate the role of this gene in neural function and its interactions with the other fAD genes.

Aβ is proposed to be a key pathogenic molecule in AD. It is unknown whether APP genes in zebrafish form aggregation prone Aβ42-like peptides as occurs in the aging human brain. To address Aim 3 a bioinformatics approach was employed. Analyses using multiple software revealed that the predicted Aβ-42-like peptide of Appa zebrafish had similar aggregation propensity potential to that of human Aβ-42. The peptides we analysed consistently showed two domains of high aggregation propensity, one at the C-terminus and one in the middle of the peptide. Interestingly, the Aβ-42-like peptide of zebrafish Appb had comparable aggregation potential to human Aβ-42 in its C-terminal end but not in its mid-peptide region, suggesting that this peptide may not be as aggregation prone as human Aβ-42.

Thesis Declaration Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Name of Candidate: Tanya Jayne

Signature: Date: 18/11/2019

Acknowledgements

I wish to extend my deepest gratitude to all who have contributed to the successful completion of this thesis.

Firstly, to my supervisors, without whose expertise and guidance I would not be the scientist I am today. To Dr Michael Lardelli, for allowing me the opportunity to work in the ADGL, for sharing your wisdom and for always challenging me. To Dr Morgan Newman, for your constant support and commitment to helping me see this PhD through. Thank you for reminding me that I deserve it when I needed it most. To Dr Joan Kelly, for your supervision and advice in my early days.

To all the members of the ADGL past and present, but especially to Dr Haowei Jiang, Dr Anne Lim, Dr Seyyed Hani Moussavi Nik, Karissa Barthelson and Yang Dong. Thank you for exchanging ideas, for excellent conversation, and for sharing in my failures, (occasional) successes and everything in between.

To my Mum, who gave up everything to make my life's journey up to this moment a reality. Thank you for your constant encouragement, positivity and for the frequent, gentle, loving reminder that I am worth it. To my sister, for listening to me vent when I needed another pair of ears. To my brother, for always telling me that you believe in me, and for inspiring me to be the strongest version of myself.

Finally, to Adrian, my partner in all things in life, for your love, encouragement and steadfast belief in me and my ability to achieve whatever I set my mind to. Thank you for listening to me vent and share my ideas, for comforting me through tears of frustration, and for always being my greatest source of strength. This thesis would not have been possible without you.

Chapter 1 – Literature Review

1.1 γ-secretase is a protease complex that is responsible for the cleavage of many substrates

γ-secretase is an important multi-subunit membrane-bound aspartyl protease complex that is responsible for the cleavage of more than 100 substrates, including; Amyloid Precursor Protein (APP), Notch and the p75 neurotrophin receptor (p75^{NTR}) [reviewed in 1, 2]. γ-secretase is a member of the intramembrane cleaving protease family (I-CLiP), which cleave type 1 membrane proteins enzymatically via a process termed regulated intramembrane proteolysis (RIP) [reviewed in 1, 2].

The most studied function of γ-secretase amongst the literature is its processing of APP. This is due to the seemingly critical role of APP in Alzheimer's disease (AD) etiology. γ-secretase's involvement in the sequential cleavage of APP can lead to increased Amyloid β-42 residue (Aβ-42) deposition. Aβ-42 is a neurotoxic peptide that aggregates and eventually forms senile/neuritic plaques (extracellular deposits of Aβ) in the central nervous system (CNS) [3]. This intensive research into γ-secretase processing of APP has been performed in the hopes of devising a potential method of treatment for AD [1, 4].

p75^{NTR}, also known as the 'low-affinity nerve growth factor receptor' (LNGFR), is another of the many substrates that is subject to cleavage within its transmembrane domain by γ-secretase [5]. p75^{NTR} is involved in development and maintenance of the nervous system through its interaction with the neurotrophins and Trk receptors [6, 7]. p75^{NTR} has also been implicated in AD [8], where it is hypothesised to play a role in cholinergic neuron loss via the Aβ peptide, which acts as a ligand for p75^{NTR} [8]. γ-secretase cleavage also has a crucial role in the processing of Notch. The Notch

receptor is a very well characterised signalling molecule. Notch signalling is vital for the determination of cell fate during embryogenesis and development [9, 10].

Components of the γ-secretase complex

The structure and mechanism of proteolysis by γ-secretase is still not completely characterised [11]. The multi-subunit γ-secretase complex is comprised of four major components (Figure 1.1), originally identified in *Caenorhabditis elegans* [12, 13]. They include; PRESENILINs (either *PSEN1* or *PSEN2*), Nicastrin (*NCSTN*), Anterior pharynx defective-1 (*Aph-1a* or *Aph-1b*) and Presenilin enhancer-2 (PSENEN or Pen-2). These integral membrane proteins provide the minimal requirement for a functional γ-secretase complex and exist in 1:1:1:1 stoichiometry [reviewed in 1, 2, 11, 14]. Previous studies using *Saccharomyces cerevisiae* confirmed these four components to be necessary and sufficient for γ-secretase activity [reviewed in 2]. The γ-secretase complex can exist in at least eight different possible compositions, as both PSEN and APH-1 can exists in two forms (PSEN1 and PSEN2 and APH-1a and APH-1b) and Aph-1a is also able to generate two individual isoforms due to alternative splicing events (APH1aS and APH-1aL) [15]. These multiple compositions may have a number of different possible functions, including; substrate affinity, diversity in activities and/or implications for different subcellular localisations [2, 14].

NCSTN is a type 1 membrane protein which has a large, heavily glycosylated extracellular domain. Originally, NCSTN was thought to recruit γ-secretase substrates through interaction between a conserved glutamate residue in its ectodomain and the N-terminal stub of the substrate [16]. However, experiments demonstrated that mutation of this glutamate residue does not result in loss of substrate attraction and binding by γ-secretase [17]. Also, a γ-secretase complex of

PSEN1/PSENEN/Aph1a was found to cleave APP and Notch in the absence of NCSTN [18]. Such findings suggest that while NCSTN may be involved in stabilisation of the complex, it is not indispensable for recognition of substrates. More recent findings suggest that NCSTN may operate as a gatekeeper for substrates of γ-secretase by sterically barring substrates that have not undergone ectodomain shedding from accessing the PRESENILIN active site [19]
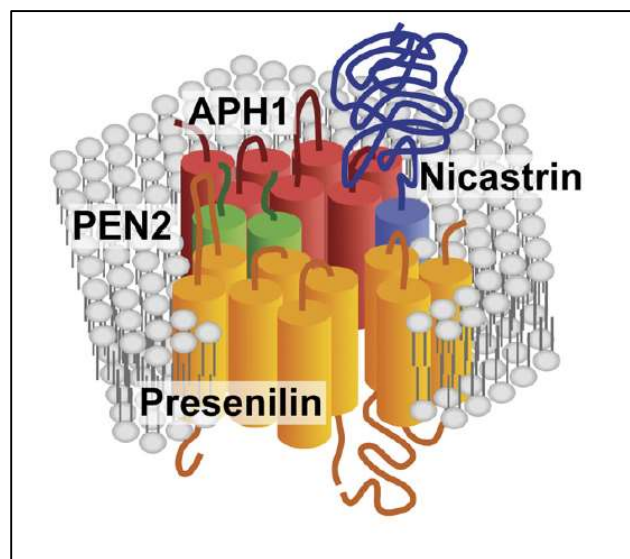
Figure 1.1. The multi-subunit γ-secretase complex. Presenilin provides catalytic activity of the complex. Presenilin, Nicastrin, APH-1 and PEN2 together are necessary and sufficient for γ–secretase activity [Adapted from 27].

PSENEN is a small (~12kDa) membrane protein with a hairpin-like structure. It is thought to be integral in stabilisation of the complex, which has been determined by mutagenesis studies [20, 21]. PSENEN directly interacts with the fourth transmembrane domain (TMD4) of PSEN1 [22, 23], and is also thought to assist PSENs in endoproteolysis [20, 24]. Aph-1 is a ~30kDa

transmembrane protein that associates with Nct and PSENs [12]. Aph-1 binds to immature Nct via its GxxxG motif and traffics this protein to the immature PSEN holoprotein [25]. Aph-1 is also thought to stabilise the γ-secretase complex [26].

The role of PRESENILINS in γ-secretase activity

The PSEN1 and PSEN2 proteins both undergo endoproteolysis to produce two fragments: a C-Terminal fragment (CTF) and an N-terminal fragment (NTF). This self-cleavage of PSENs has been shown to be promoted by *PSENEN* [28]. The two PSEN fragments associate to form a functional heterodimer, the active form of PSEN [reviewed in 14]. This heterodimer exists in the γ-secretase complex and makes up the catalytic core of this membrane anchored protease [16, 29].

Research into the PSENs has a complex history and for many years there was doubt as to whether the PSENs were responsible for γ-secretase cleavage activity. Early investigations in the lower eukaryote *Pichia pastoris* revealed that a lack of PSENs does not impede γ-secretase-like proteolytic processing of APP, suggesting that PSENs were not necessary for this function [30]. A study involving the processing of another γ-secretase substrate Notch, suggested that while PSENs do play a role in Notch processing they are not necessarily responsible for the direct cleavage of this substrate [31]. Another study found that small changes in pH conditions of an *in vitro* assay for γ-secretase can result in changes in the ability to inhibit a γ-secretase-like cleavage of APP [32]. A paper published by Ahn *et al* (2010) demonstrated that PSEN1 does have γ-secretase catalytic function [28]. They demonstrated that recombinant PSEN (rPS) exhibits γ-secretase activity and that rPS mutants lacking the catalytic aspartate residues lose this activity [28]. Furthermore, rPS harbouring familial Alzheimer's disease (fAD) causing mutations led to elevated

Aβ42-40 ratio produced from a wild type APP [28]. Supplementing this are PSEN1 knockout studies conducted in mice by De Strooper *et al* (1998), in which secretion of Aβ peptide was strongly inhibited. There was also accumulation of APP C-terminal fragments in PSEN1 deficient mouse neurons. Taken together, these findings provide a firm argument that PSEN is indeed responsible for the catalytic activity observed in γ-secretase [33]. Lending further support to this argument is the finding by Wolfe *et al* (1999), that two aspartate residues in PSENs are vital, both for endoproteolysis of PSEN and γ-secretase cleavage of substrates [34].

The mystery of substrate recognition and selection by γ-secretase

The specific nature of substrate selection by γ-secretase is still not clearly defined. γ-secretase substrates are typically derived from large precursor proteins that undergo a prerequisite removal of their ectodomain prior to γ-secretase cleavage, in a process termed "ectodomain shedding" [2]. Indeed, γ-secretase substrate recognition and cleavage are much more efficient for ectodomains with fewer than ~50 remaining amino acid residues [9]. In early studies, the only other observed prerequisite was that the substrate must be a type 1 membrane protein [16]. Later studies suggested that dimerization of substrates and/or the structure of substrate α-helices may regulate γ-secretase activity [35]. More recently, cryo-electron microscopy (cryo-EM) has been utilised to solve structures of the γ-secretase complex [36]. Unfortunately, the findings of this study did not reveal further details of how the γ-secretase complex interacts with its substrates. However, subsequent studies utilising cryo-EM improved the resolution of the human γ-secretase complex and revealed conformational flexibility upon substrate binding [37, 38]. Furthermore, a cryo-EM study by Bai *et al* (2015) seemingly supports that the previously mentioned glutamate residue (Glu333), along with a tyrosine residue (Try337), of NCSTN likely play a role in substrate recruitment [38, 39].

Other recent findings suggested a mechanism whereby γ-secretase trims substrates within their transmembrane domains (TMD) from within three substrate binding pockets (S1′, S2′ and S3′). These three substrate binding pockets trim peptides in intervals of three amino acids [19]. Even more recently, there have been two studies published that utilised cryo-EM to resolve structures of the γ-secretase complex bound to either Notch or APP [40, 41]. Both studies reported similar confirmations in binding of their respective substrate, where, the initial two-thirds of the TMD formed a helical conformation and the final third was unwound into a β-strand to engage the active site. Moreover, specific TMDs of PRESENILIN changed conformationally in order to interact with both substrates [19]. These findings are a great step forward in our understanding of γ-secretase substrate recognition, however, there remains much unknown about this process.

1.2 The p75 neurotrophin receptor (p75^{NTR})

The p75 neurotrophin receptor (p75^{NTR}), also known as the 'low-affinity nerve growth factor receptor' (LNGFR), is one of the 100+ substrates that is subject to cleavage within its transmembrane domain by γ-secretase. p75^{NTR} is a member of the tumour necrosis factor (TNF)/nerve growth factor (NGF) superfamily [5]. As its name suggests, p75^{NTR} acts as a receptor for the neurotrophins (NT's), both in their pro-NT (immature) and mature forms [reviewed in 6]. Neurotrophins are a family of protein growth factors, vital in the maintenance and development of the nervous system. The four neurotrophins are; 1. Nerve growth factor (NGF), 2. Brain derived neurotrophic factor (BDNF), 3. Neurotrophin-3 (NT-3) and 4. Neurotrophin-4 (NT-4) [reviewed in 6, 42, 43].

P75^NTR has an ECD (extracellular domain) containing a signal peptide and a cysteine-rich motif of four repeats. It has a transmembrane domain which is the site of γ-secretase cleavage and through which it has been found to form a homodimer at the cell surface [35]. This homodimer is thought to be involved in receptor-ligand interactions. It also has an ICD with a 'death domain' motif that shares similarity to the TNF (tumour necrosis factor) family of proteins. The ICD also contains a PEST sequence, an indicative character for proteins that undergo proteasome degradation [6, 35]. More recently, the p75ECD has been shown to be neuroprotective in mouse brain [44].

The ICD (intracellular domain) of p75^NTR can promote both apoptosis and cell survival upon release into the cytosol [43]. There are several described protein interactions which alter the final product of p75^NTR signalling. Promotion of neuronal survival is achieved through p75^NTR binding with TrkA in the presence of NGF (Figure 1.2) [reviewed in 6, 42]. Another interaction is p75^NTR with sortilin (SORT1) (Figure 1.2). Interaction of the p75^NTR/SORT1 heterodimeric complex with proNGF or proBDNF leads to apoptosis [reviewed in 6, 43].

p75^NTR proteolysis is carried out by α- and γ-secretases

The proteolysis of p75^NTR (Figure 1.3) is very similar to that of two of γ-secretase other substrates, APP and Notch. Initially p75^NTR is subject to cleavage 15 amino acids N-terminal to the transmembrane domain by α-secretase, leaving a membrane-anchored CTF [45]. This α-secretase cleavage of p75^NTR is mediated by a member of the 'A Disintegrin and Metalloproteinase' (ADAM) family [45]. Whether this cleavage is mediated by ADAM10 or ADAM17 has not been firmly established. Experiments by Kanning *et al* (2003) suggest that it is ADAM10 which provides the primary cleavage event [5]. However, evidence provided by Weskamp *et al* (2003)

and Forsyth *et al* (2014), suggests that ADAM17 that is responsible for ECD shedding [46, 47]. It is possible that both ADAM10 and ADAM17 can cleave p75^{NTR}. Following the shedding of the p75^{NTR} ectodomain there is a firmly established cleavage event by γ-secretase [5, 6]. This occurs within the transmembrane domain of p75^{NTR} although the specific position has not been well defined. This cleavage releases the ICD into the cytoplasm, which is responsible for downstream signalling leading to either cell survival or apoptosis [42, 43]. The processing of p75^{NTR} by γ-secretase has been demonstrated as highly important in respect to its downstream signalling responsibilities [42].
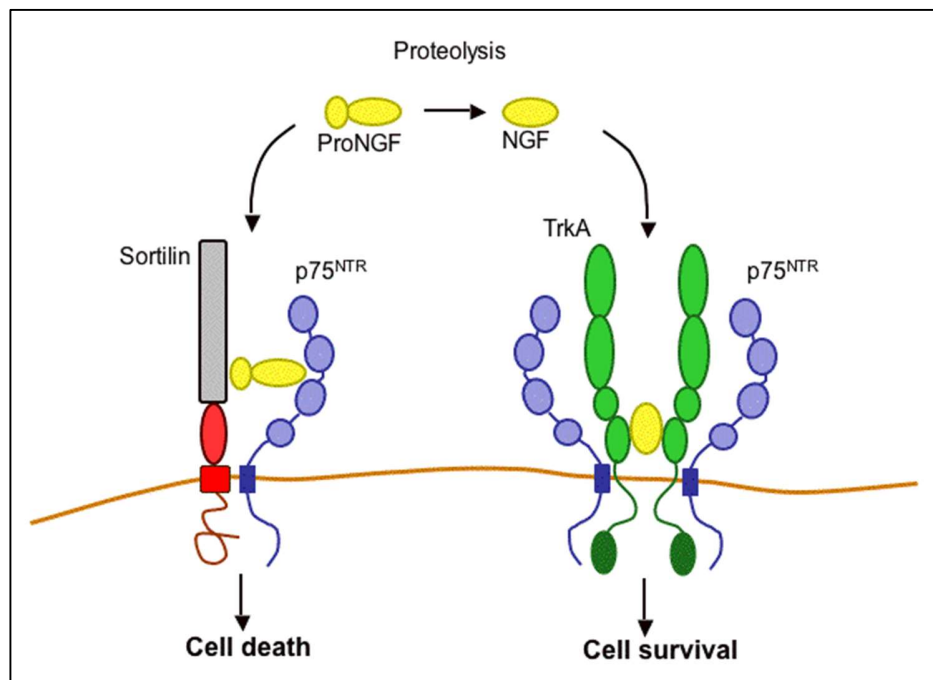
Figure 1.2. p75^{NTR}can interact with both Sortilin and TrkA receptors to promote cell death or survival respectively. (Image from http://www.newmind-center.dk)

p75^NTR has two homologues with distinct cleavage properties

Our interest in the issue of γ-secretase-like cleavage activities was piqued by a paper published by Kanning *et al* (2003) investigating the proteolytic processing of p75^{NTR}, along with what they described as the "neurotrophin receptor homologs" (NRH), NRH1 and NRH2 [5]. Database and EST searches have established that these two genes show greater sequence similarity to p75^{NTR} than its previously known closest relative sequences. Kanning *et al* (2003) found that NRH subfamily members share approximately 30% identity in their ICD across species, whereas only 13% identity is shared with the ICD of p75^{NTR} next nearest known relative sequence, human *TNFR1* [5]. p75^{NTR} and the NRH genes are described as a subfamily of the nerve growth factor (NGF)/tumour necrosis factor (TNF) family. NRH2 exists only in mammals and NRH1 exists only in amphibians, fish and chickens. Alignment of p75^{NTR} with both homologs revealed a high degree of similarity in ICD and transmembrane domains [5].

Experiments investigating the proteolytic processing of p75^{NTR} confirmed its processing by α- and γ-secretase. Results from Western blot analysis of NRH1 and NRH2 indicated that they are both cleaved within the transmembrane domain [5]. Using proteasome inhibitors, particularly the commonly used γ-secretase inhibitor DAPT, Kanning *et al* (2003) observed that DAPT had no effect on the cleavage of NRH1 or NRH2, suggesting that transmembrane cleavage of these proteins is not by γ-secretase [5]. Inconsistencies within the literature pertaining to this conclusion do exist. For example, a paper published by Zampieri et al (2005) stated that NRH2 is a target of both α- and γ-secretase activities [45]. However, no references or experimental data were provided as evidence for this statement [45]. It is true that the presence or absence of α-secretase cleavage of both NRH1 and NRH2 was not clearly assessed by Kanning *et al* (2003) [5]. However, their experimental procedures and controls addressing γ-secretase cleavage were sound and convincing.

As such it is increasingly clear that this cleavage must arise from another, currently unknown processing event. It would be interesting to investigate further the events which are involved in the processing of NRH1 that eventually lead to the release of its ICD.
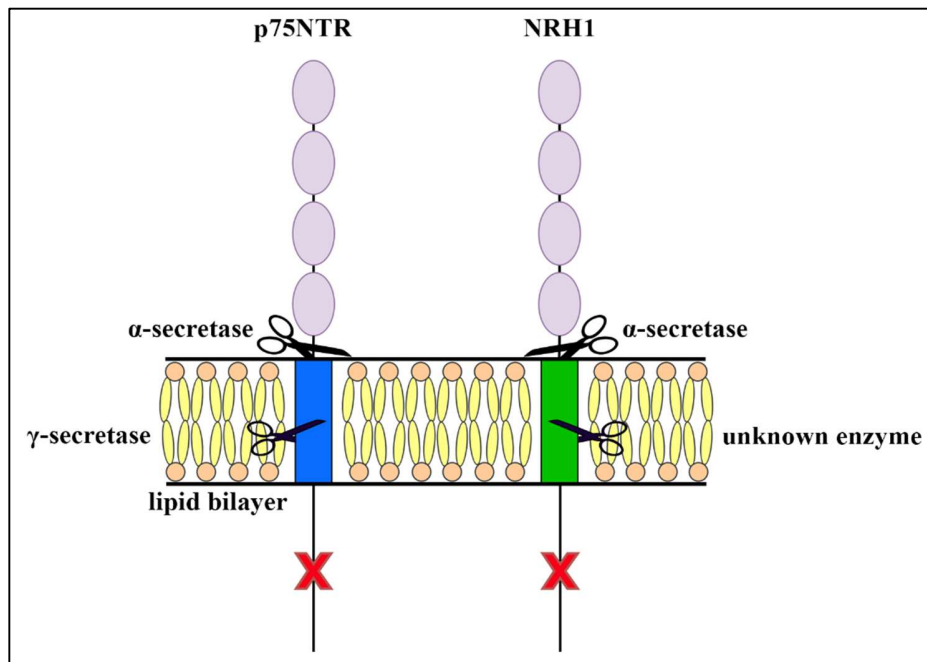
***Figure 1.3. Cleavage events of p75^NTR and its homolog NRH1.** Both p75^NTR and NRH1 are cleaved first in the extracellular domain by α-secretase and then in the transmembrane domain. While p75^NTR is cleaved in the TMD by γ-secretase, the enzyme that cleaves NRH1 at this position remains unknown.*

The p75^NTR homolog of greatest interest to this thesis is NRH1, as this is the p75^NTR homolog in zebrafish. We have previously observed that zebrafish NRH1 is not subject to γ-secretase-dependent cleavage [48]. However, the nature of primary cleavage events, although potentially mediated by α-secretase, remains uncertain [5]. The high sequence and structural similarity between NRH1 and p75^NTR [5] makes the lack of γ-secretase-dependent cleavage of NRH1

particularly intriguing, since the genes coding for these proteins must have arisen by duplication of a common ancestor. Further investigation of NRH1 processing events would be of interest, as this may give us some insight into; 1) the evolution of γ-secretase cleavage of p75^{NTR} and 2) substrate selection by γ-secretase in general.

1.3 Epidemiology, history and etiology of Alzheimer's disease

As previously mentioned, γ-secretase is most studied for its role in APP processing, through which production of the neurotoxic Aβ-42 peptide is thought to contribute to the observed neurodegeneration in AD. Alzheimer's is the most common form of dementia, with 50-70% of all dementia cases in Australia diagnosed as AD. More than 413,000 Australians are living with some form of dementia and it has been estimated that, by 2050, 1,000,000 Australians will be living with dementia. There is currently no cure for AD, even though it is the second leading cause of death in Australia (https://alzheimers.com.au/about-alzheimers/). As there is currently no cure, there is an intense research effort into all aspects of AD biology to attempt to identify one.

In 1901 at the Frankfurt Psychiatric Hospital, Alois Alzheimer began treating a 51 year old female patient, Auguste Deter. She had been admitted to hospital for symptoms that are associated with AD to this day; paranoia, sleep disorder, memory disturbance, aggressive behaviour and confusion [49]. After her death in 1906, Alzheimer performed post-mortem analysis on brain tissue from Auguste. His published observations described the presence of senile plaques and neurofibrillary tangles (NFT) within her cerebral cortex [50]. The presence of these two types of brain lesions has become widely accepted as the basis of characterisation of the disease. More than a century later, patients are still not confirmed as having had AD without displaying a similar brain pathology.

The characteristic senile plaques and NFT of AD are located in both the temporal neocortex and hippocampus [51, 52]. Plaques and NFT are a downstream effect of abnormal deposition of the Aβ peptide and hyperphosphorylated tau protein respectively. The deposition of these proteins results in neuronal loss and neurotoxicity by unknown mechanisms [53]. Aβ peptides are believed to be capable of triggering oxidative stress, which is extensively observed in AD [54, 55]. This stimulation can occur both directly and indirectly. Aβ peptides are able to produce hydrogen peroxide directly and also generate free radicals via the reduction of metal ions [56, 57]. Aβ peptides can also influence the generation of free radicals through binding Aβ-binding alcohol dehydrogenase (ABAD), which forms a direct link between Aβ and mitochondria [58]. Neuroinflammation is another mechanism through which Aβ peptides produce oxidative stress. Many studies have supported the implication of neuroinflammation's role in AD pathogenesis [59]. The neuronal death observed in AD is, in part, accredited to microglial and astrocyte activation. Both of these brain cell types are involved in the inflammatory and immune response to deposition of Aβ, and both have neuroprotective and neurodegenerative roles [60]. Studies have indicated that astrocytes are able to degrade Aβ deposits both *in vitro* and *in situ*, which suggests they may be involved in the clearance of Aβ aggregates [61]. Clustering of astrocytes at amyloid deposition sites lends further support to this proposed role. When activated, both microglia and astrocytes are known to release pro-inflammatory molecules. Microglia in particular have been shown to release proteases, cytokines (chemokines) and reactive oxygen species (ROS) in response to activation by Aβ, all of which are potentially neurotoxic [62].

There are two recognised forms of AD. Familial AD (fAD) also known as 'early onset' AD (EOAD) occurs in patients between the ages of 40 and 65 years of age. fAD is thought to represent approximately 2% of all AD. Sporadic AD (SAD) or 'late onset' AD (LOAD) occurs in patients

over the age of 65 and accounts for the majority of AD cases. Approximately 10% of fAD cases present a clear autosomal dominant inheritance pattern [63]. However, a more complex pattern of inheritance is observed in most cases of both forms of AD [64].

Genes implicated in familial Alzheimer's disease

There are currently three loci for autosomal dominant mutations associated with the rare familial form of AD (fAD). The first of these loci to be identified was APP, which encodes the Aβ peptide that is accumulated into amyloid plaques. 59 fAD contributing mutations have been described for the *APP* gene (https://www.alzforum.org/mutations/app). The methods by which APP mutations contribute to fAD will be discussed in more detail later. Linkage analysis identified that the two related genes known as *PSEN1* and *PSEN2* are also associated with fAD [65]. Mutations in *PSEN1* are the major contributor to fAD, with 300+ fAD mutations in this gene currently described. There are approximately 45 fAD-causing mutations in the gene encoding *PSEN2* (https://www.alzforum.org/mutations).

Most fAD causing mutations are found in the *PSEN1* gene. *PSEN1* mutations are spread throughout the open reading frame of the gene but are more concentrated in sequences coding for trans-membrane (TM) domains 1-6 [66]. Both *in vitro* and *in vivo* studies have shown that PSEN mutations increase Aβ-42 or decrease Aβ-40, which leads to an altered Aβ-42/Aβ-40 ratio [67-69]. However, studies using cell culture systems, while confirming an increase in Aβ42/Aβ40, showed that this was due to a decrease in Aβ-40 peptide levels [70]. Furthermore, Cacquevel *et al* (2012) confirmed that fAD-linked pathogenic mutations in *PSEN1* result in a loss of γ-secretase activity that reduces Aβ-40 levels, thereby increasing the Aβ-42/Aβ-40 ratio [71]. Another study showed that the cerebral spinal fluid (CSF) of PSEN1 mutation carriers had much lower levels of

Aβ-42 than sporadic early onset AD patients [72]. These observations and others suggest that *PSEN1* mutations can result in the observed altered ratio of Aβ peptides in a wide variety of ways. One question that remains is what the increased Aβ-42/Aβ-40 ratio observed for these *PSEN1* mutations mean in the context of AD development. Interestingly, Sun *et al* (2016) used an *in vitro* method to measure the production of Aβ-40 and Aβ-42 peptides from γ-secretase variants, where each variant carried 1 of 138 individual *PSEN1* AD mutations [69]. Their study aimed to address the question of whether the Aβ-42/Aβ-40 ratio resulting from these different *PSEN1* mutations, correlated with the mean age of onset of AD observed for each mutation. While they observed no such correlation, they also observed that some of the *PSEN1* mutations increased γ-secretase activity rather than decreasing it (as had been observed in the studies above), further complicating the question of the role of *PSEN1* mutations in AD [69].

Genes implicated in Late Onset Alzheimer's Disease

The gene with the strongest confirmed association to LOAD is apolipoprotein E (APOE). The polymorphic APOE gene has three allelic variants; ε2, ε3 and ε4. These isoforms are defined by variations in the amino acid residues at two positions in the 299 amino acid residue APOE protein, (112 and 158) that can alter its structure; ε2 allele (112 Cys, 158 Cys), ε3 allele (112 Cys, 158 Arg) and ε4 allele (112 Arg, 158 Arg) [reviewed in 73]. While the ε4 allele contributes increased susceptibility to AD, it is not necessary nor sufficient to cause this disease [74]. There is, however, a gene dosage effect associated with ε4. Carriers who are homozygous for the ε4 allele develop AD earlier than those who are heterozygous at this locus and may also have a greater burden of Amyloid [75]. Interestingly, while the ε4 allele seems to promote AD pathology, the ε2 allele is thought to be neuroprotective [76].

Regarding APOE's role in AD risk there are several observations to consider. Firstly, APOE has been described to have a direct interaction with Aβ protein *in vitro* [77]. Furthermore, stimulation of APP endocytosis and metabolism by the ε4 allele, both *in vitro* and *in vivo*, can potentially enhance Aβ production [78, 79] and APOE can also potentially modulate Aβ toxicity [80]. Finally, there is strong evidence suggesting that extracellular clearance of in Aβ the brain is regulated by APOE [81] and that the ε4 allele has a reduced ability to clear aggregated amyloid in plaques [82].

LOAD risk factors identified during the GWAS era

GWAS studies have been successfully used in the past to identify risk factors associated with complex disorders such as diabetes and Macular degeneration [83]. Since 2009, several Genome-wide association studies (GWAS) have identified more than 40 novel genetic loci associated with AD risk, including; *CLU*, *PICALM,* and *SORL1* [84].

Early GWAS identified variants of clusterin (*CLU*) and two polymorphisms in Phosphatidylinositol binding clathrin assembly protein (*PICALM*) associated with LOAD [85]. *CLU* is expressed in cerebrospinal fluid, amyloid plaques and is up regulated in the brains of AD affected patients [86]. Interaction of CLU with soluble Aβ has been implicated by studies in guinea pig, which suggest the complexes formed are able to cross the blood brain barrier [87]. The function of CLU binding to Aβ and chaperoning it across the blood brain barrier may be to prevent formation of Aβ oligomers and fibrils [83]. Furthermore, there is a potential role for CLU in protecting cells against apoptosis and oxidative stress [83]. *PICALM* is expressed ubiquitously in all tissue types but most prominently in neurons [85]. PICALM has also been implicated in endocytosis of APP and generation of Aβ and increased cleavage and aggregation of this peptide

may also result from underexpression of *PICALM* [88]. Recently, evidence from *in vitro* and *in vivo* studies has implicated PICALM in Aβ transcytosis and clearance at the BBB [89].

Other potential AD susceptibility genes identified by GWAS analyses can be found on the "AlzGene" database (http://www.alzgene.org) [90]. These genes have a range of roles in AD neuropathology progression through regulation of Aβ. Other roles related to AD pathology include suppression of Tau phosphorylation and NFT formation, involvement in lipid metabolism, impaired inflammation and other cellular pathways [83]. Other AD risk genes, such as TREM2, have also been identified by both next generation sequencing (NGS) and the Alzheimer's disease sequencing project (ADSP) [84].

The SORL1 gene potentially associates with both forms of AD

In the preceding sections we discussed genes known to associate specifically with either fAD or LOAD, however, SORL1 is associated with both forms of AD. SORL1 was first identified to have an association with AD through microarray screening of AD patient lymphoblasts [91]. Since then it has been implicated in both LOAD through GWAS analysis [92, 93] and fAD through exome sequencing [94]) and a candidate gene approach [95]. Rare missense variants of SORL1 have now been identified that co-segregate with autosomal dominantly inherited AD [96]. Interestingly, these rare variants SORL1 are also often observed to co-occur with an APOE ε4 homozygous genotype, suggesting that, in these families, the cause of AD might be oligogenic [96].

SORL1 is thought to act as an AD risk factor through trafficking and processing of APP [97]. Studies of mutant SORL1 variants demonstrated its ability to bind and traffic APP in two ways that potentially impair Aβ production. One of SORL1's normal functions is to retrogradely sort

APP into the retromer recycling pathway, shuttling it from endosomes to the trans-Golgi network (TGN) [98]. Consequently, overexpressing SORL1 has been confirmed to decrease Aβ production in cell lines [95]. Conversely, suppression of SORL1 increased the levels of Aβ with subsequent increase in senile plaques in AD mouse models [99]. This suggests a second role for SORL1 in moving APP anterogradely into the endosome, which is where APP is processed to release Aβ. From there SORL1 has also been shown to direct Aβ to lysosomes to be degraded, thereby potentially protecting the brain by reducing Aβ levels [97, 100]. Another possible explanation for Aβ accumulation in response to SORL1 suppression might be that SORL1 competes with APP as a substrate for γ-secretase. Bohm *et al* (2006) previously demonstrated that SORL1 is also a substrate of γ-secretase, therefore it's possible that a reduction of SORL1 levels in the brain would lead to increased processing of APP, consequently increasing Aβ [101]

The Amyloid cascade hypothesis is the prevailing hypothesis for AD

A complete understanding of the etiology of AD continues to prove elusive. The most commonly accepted hypothesis for the development of AD pathology is the 'Amyloid Cascade', originally described by Hardy and Higgins in 1992 [102]. This hypothesis suggests that missense mutations in *APP*, *PSEN1* or *PSEN2* trigger Aβ plaque and neurofibrillary tangle formation, which eventually leads to neuronal cell death and dementia [102, 103]. The key factor in this hypothesis is the production of the Aβ peptide. Depending on the positions of cleavage by multiple enzymes associated with the processing of APP, Aβ peptides of between 39-42 amino acid residues in length can be generated. Soluble Aβ-40 peptide is the major species present in normal brain. The less soluble Aβ-42 isoform of the Aβ peptide is thought to be the most prone to aggregation in the brain. Aβ-42 has been implicated as the major component of senile/neuritic plaques which are

deposited in brain parenchyma. As previously mentioned, these plaques are considered to be a histological hallmark of AD [104].
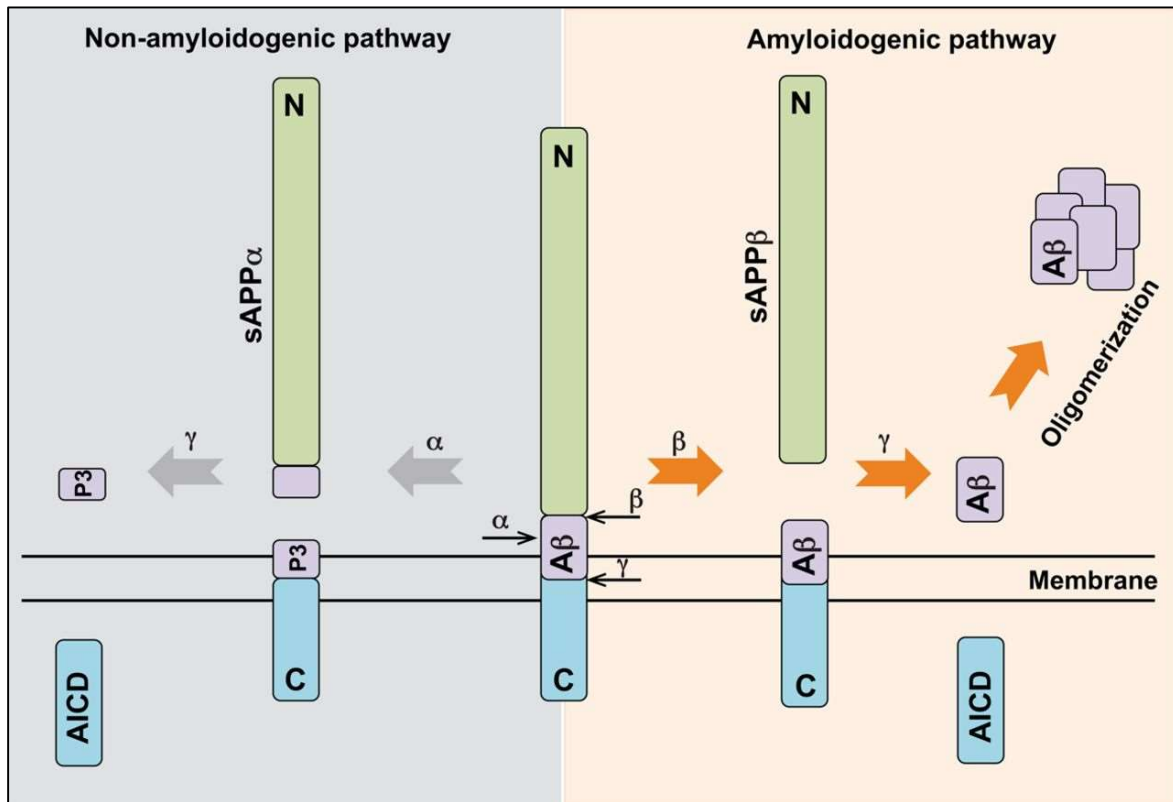
Figure 1.4. The proteolytic processing of APP. *APP can be cleaved by α-, β- and γ-secretases; the cleavage sites of these proteases are indicated in the full-length APP shown in the centre of the figure. Oligomerisation of Aβ is proposed to make up the senile/neuritic plaques in the brains of AD patients (reproduced without permission from [107]).*

Processing of APP to form aggregation-prone Aβ-42 peptide occurs in a sequential manner (Figure 1.4). Firstly, there must be removal of the extracellular domain (ECD) of the protein, termed 'shedding'. This is required in order for it to be processed by γ-secretase [9]. There are two pathways in which APP can be processed. In the amyloidogenic pathway the APP ECD is shed by β-secretase (BACE1) before cleavage by γ-secretase to produce Aβ. In the non-amyloidogenic

pathway, shedding is performed by α-secretase (ADAM10 or ADAM17) and, after γ-secretase cleavage, produces the p3 peptide instead of Aβ [2, 10, 66]. Shedding initially leaves a membrane bound C-terminal fragment. Shedding via α-secretase leaves C83. Shedding via β-secretase leaves C99. The C99 fragment then undergoes subsequent cleavage by γ-secretase to release the neurotoxic Aβ-42 peptide [105, 106]. Downstream effects of the increased production of Aβ-42 peptide and its subsequent deposition as diffuse plaques include the activation of microglia and astrocytes as discussed above.

Alternative hypotheses for Alzheimer's disease

Although the Amyloid cascade hypothesis remains the most popular hypothesis for AD, there are many who doubt its veracity. There is a limit to which this hypothesis can accurately explain AD progression. Reviews by David Drachman and Jack de la Torre discuss this issue in detail [108, 109]. Many AD sufferers never develop Aβ plaques and/or NFTs while many non-AD elderly individuals do exhibit these histological features. These among other observations call into question the ability of the amyloid cascade hypothesis to accurately describe AD pathogenesis. There are a great number of alternative hypotheses that attempt to describe the pathogenic nature of AD [reviewed in 110, 111]. Here I will discuss in detail a select few that are most relevant to this review's theme. If the reader is interested in other alternative hypotheses not discussed in this review, such as; the cholinergic hypothesis, the cell cycle hypothesis, the aluminium hypothesis, the inflammatory hypothesis, or the calcium hypothesis, they might be interested in the following book chapter [111].

The Presenilin hypothesis

The presenilin hypothesis (Figure 1.5) posits that the many pathogenic missense mutations in *PSEN* that are distributed throughout the protein result in a loss of protein function, potentially due to destabilisation of its overall structure. It is well documented that not all cases of neurodegenerative dementias have amyloid pathology as is observed in AD (eg. FTD). There are also mouse models that overexpress mutant APP and in which there is overproduction of Aβ peptides and subsequent deposition of this peptide but no apparent neurodegeneration [112]. These two observations led Shen and Kelleher to challenge the Amyloid cascade hypothesis and present their own alternative hypothesis, "the presenilin hypothesis" [113]. As mentioned above, the presenilins are involved in the γ-secretase-mediated cleavage of many substrates. There is a substantial body of evidence supporting the opinion that mutations in PSEN that are pathogenic may cause partial impairment of PSEN function, disturbing both γ-secretase-independent and -dependent actions. For example, the first study in a vertebrate system to show reduction of PSEN activity via pathogenic mutations, examined the effects of six FAD-linked PSEN mutations on generation of the Notch intracellular domain (NICD). All six mutations resulted in reduced proteolytic release of NICD [114]. PSENs are involved in γ-secretase-independent down-regulation of the Wnt signalling pathway by interacting with and destabilising β-catenin [115]. FAD-linked PSEN mutations have been found to impair this function of PSENs, leading to increased β-catenin stability and β-catenin-dependent signalling [116, 117].

Shen and Kelleher proposed that loss of PSEN protein function occurs via a dominant negative mechanism, whereby mutant PSEN interferes with the function of wildtype PSEN. This is supported by the nature and distribution of these mutations [113]. In order to incorporate the observation that *APP* mutations alone are sufficient to cause AD into this "PSEN loss of function"

hypothesis, the effects of Aβ peptide on PSEN function must be considered. Shen and Kelleher suggest that Aβ-42 overproduction in both *PSEN* and *APP* FAD mutations may enhance the PSEN impairment observed by way of a "product-based negative feedback mechanism" [113]. As γ-secretase has such a relaxed specificity for its substrates, longer forms of Aβ (42/43) containing potential cleavage sites (that would generate shorter forms), may fill the active site. It is probable that these Aβ species would be ineffective substrates for cleavage, as they lack downstream residues important for substrate-active site interactions. Due to this they would block the γ-secretase active site, preventing it from cleaving its other important substrates.
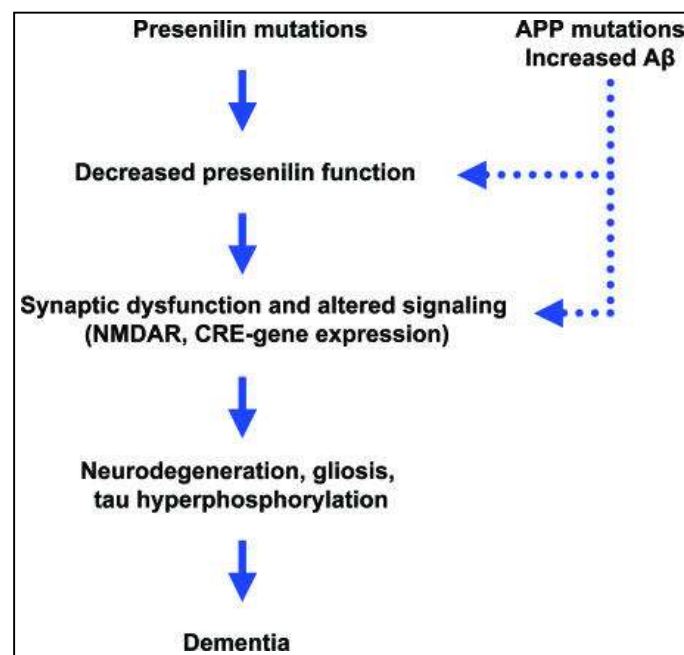
Figure 1.5. The Presenilin Hypothesis of AD. *(Reproduced without permission from [113]).*

A combination of PSEN inactivation and increased Aβ result in reduced cAMP-response element (CRE)-dependent gene expression and number of synaptic NMDA receptors [113]. Furthermore,

expression of both the PSEN1 and PSENEN encoding genes are CRE-dependent [113]. When PSEN function in the adult mammalian brain is lost this results in a pathogenic cascade, which ultimately triggers widespread and progressive neuronal atrophy and death (aka neurodegeneration resembling AD neuropathology) [113].

The vascular hypothesis of AD

First and foremost, the vascular hypothesis of AD attempts to step away from the amyloid cascade hypothesis. This hypothesis postulates that Alzheimer's is primarily a vascular disorder, and that the neurodegeneration observed in the ageing brain is a downstream effect of vascular dysfunction and subsequent cerebral hypoperfusion [118]. This hypothesis is supported by the discovery of vascular AD risk factors in the elderly, including; diabetes 2, hypertension, heart disease, smoking, obesity, atherosclerosis. These risk factors coupled with the ageing process can introduce a mild ischemia/hypoxia state that is sustained in the brain [118]. Hypoxia/ischemia refers to a state in which there is insufficient blood flow to a tissue and therefore insufficient oxygen supply. The cell responds to hypoxia by activating the transcriptional regulator, hypoxia-inducible factor 1-alpha (HIF-1α) [119]. HIF-1α has been demonstrated to activate both β-secretase (BACE1) and γ-secretase (PSEN1), which leads to this increased production of Aβ in the brain and cerebrovasculature [120]. Furthermore, hypoxia also stimulates oxidative stress that can produce both reactive oxygen species (ROS) and reactive nitrogen species (RNS). Peroxynitrite, one of these RNS, promotes nitrotyrosination of PSEN1, consequently shifting production of Aβ to favour Aβ-42, thus increasing the Aβ-42/Aβ-40 ratio [reviewed in 121].

1.4 The Amyloid Precursor Protein

The γ-secretase substrate APP is a highly conserved type 1 transmembrane protein (Figure 1.6) encoded by a gene on chromosome 21 [122]. There are orthologues in all mammals that have been studied as well as some non-mammals, such as zebrafish and *Drosophila*. APP is a member of a large protein family, along with transmembrane and secreted proteins (APLP's) [123].
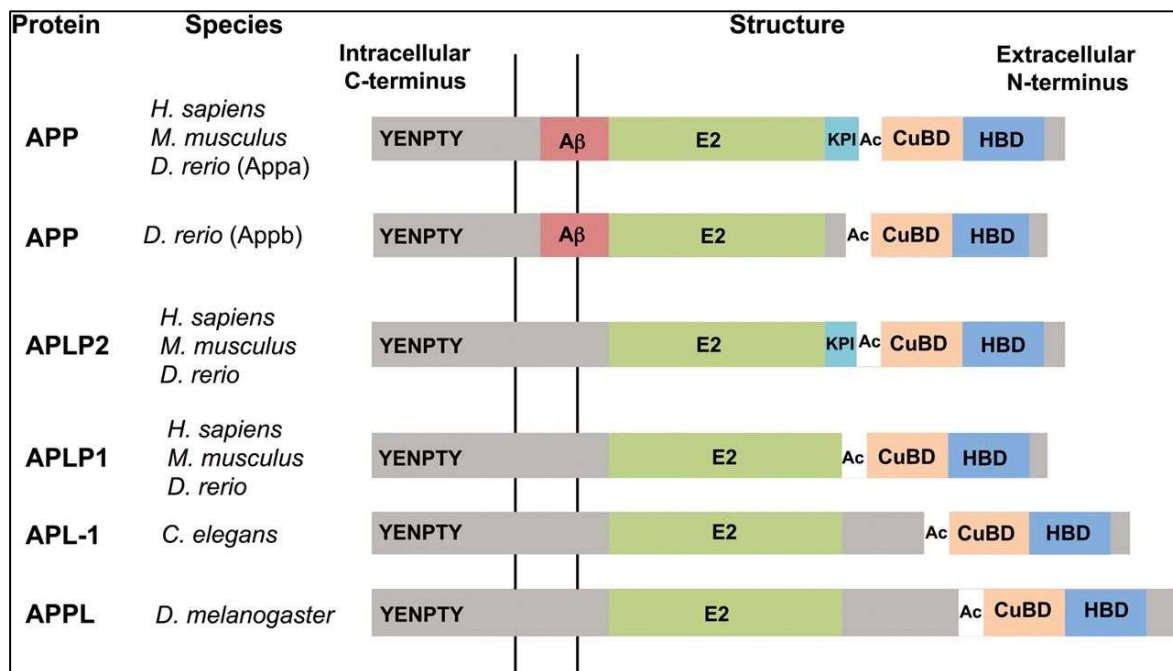
Figure 1.6. The domain structure of APP family members is conserved across species. *The domain structure of human (*H. sapiens*) APP and its homologs in the mouse (*M. musculus*), zebrafish (*D. rerio*), worm (*C. elegans*) and fruit fly (*D. melanogaster*) is shown. The extracellular region contains an E2 domain, an acidic (Ac) domain, a copper-binding domain (CuBD) and a heparin-binding domain (HBD), all of which are conserved across species. A Kunitz protease inhibitor (KPI) domain, which is subject to alternative splicing, is also found in APP and APLP2. The intracellular domain shows the highest homology and contains the YENPTY motif that is conserved across homologs. The Aβ sequence is only present in APP (reproduced without permission from [107]).*

The structure and heterogeneity of APP

There are three major APP isoforms (Figure 1.7), which are the result of alternative splicing of exons 7 and 8 of the APP gene [124]. Both the 770 and 751 amino acid residue splice variants are expressed in non-neuronal and CNS tissues. The 695 residue splice variant is expressed almost exclusively in the CNS [123]. APP770 represents the full length protein. APP751 results from exclusion of exon 8. APP695 results from exclusion of both exons 7 and 8. Exon 7 encodes a Kunitz type serine protease inhibitor (KPI) homologous 56-amino acid domain [125]. Those APP isoforms with a KPI domain found in human platelets are involved in the coagulation cascade via inhibition of IXa, X and Xia factors [126]. Exon 8 encodes the 19aa OX-2 domain, the function of which is largely unknown [127]. Post-translational modifications of APP also contribute to its heterogeneity, such modifications include addition on sulfates, phosphates and *N*- and *O*-linked sugars [reviewed in 128].
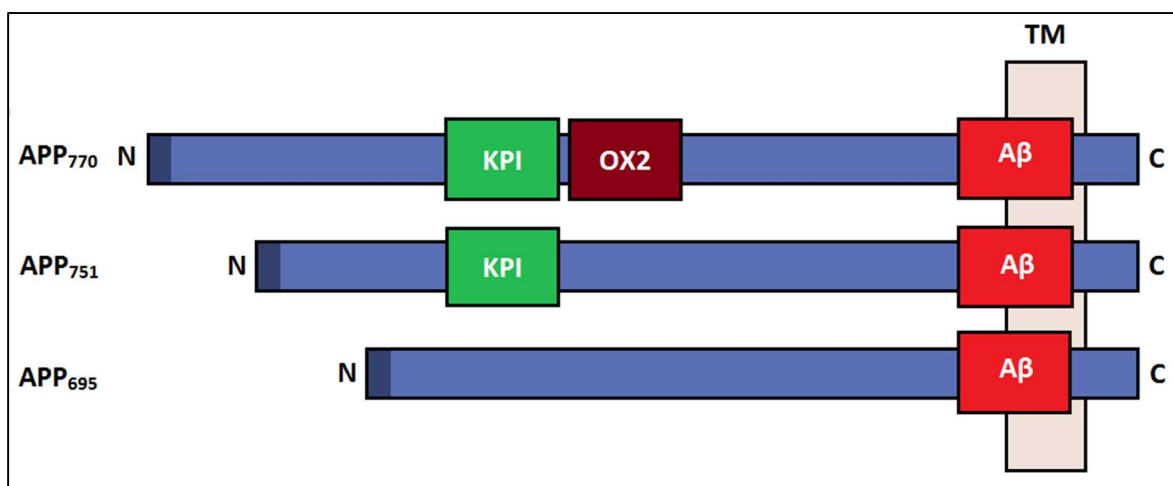
Figure 1.7. Diagram representing three proteins from the APP family and their main domains. *APP 770 contains all domains, APP 751 carries only the KPI domain and APP 695 lacks both the KPI and OX2 domains (inspired by [127]).*

Normal (non-pathogenic) functions of APP

As mentioned previously, research involving APP has been heavily focussed on its association with AD, hence its normal functions are much less well understood. APP is ubiquitously expressed and is highly evolutionarily conserved [129]. APP is thought to play roles in neurogenesis, neurite outgrowth and guidance, axonal transport and synaptogenesis and morphogenesis [107, 130]. Furthermore, the full length APP holoprotein has been proposed to have roles in cell-cell interactions [131] and neuronal precursor migration during brain development [132]. The many other suggested roles of APP are attributed to its cleavage products. APP has two well described processing fates, amyloidogenic and non-amyloidogenic. There are three fragments that can be generated by non-amyloidogenic processing of APP. Initial cleavage by α-secretase generates the soluble sAPPα fragment and a membrane bound C-terminal fragment (CTF). This CTF then undergoes cleavage by γ-secretase to release the p3 peptide and the APP intracellular domain (AICD) [2]. Each of these fragments have suggested roles in the brain.

sAPPα is both neurotrophic and neuroprotective

sAPPα is both neurotrophic and neuroprotective [129]. sAPPα carries a cysteine-rich domain resembling that of other growth factor domains, which suggests a function for sAPPα as a growth factor receptor ligand [107, 133]. Furthermore, sAPPα has more recently been found to inhibit collapse of the growth cone [130]. Early studies demonstrated stimulation of neural stem cells (NSCs) by sAPPα in rat [107]. Others found that a reduction of NSC proliferation in the brains of adult mice could be rescued by infusion of sAPPα [134]. Supplementing this was a study that showed inhibition of α-secretase decreased NSC proliferation *in vitro*, which was rescued by

adding sAPPα to the culture medium [135]. This same study found that sAPPα regulates proliferation of not only NSCs but also mesenchymal stem cells (MSCs) and human decidua parietalis placenta stem cells (hdpPSCs) [135]. Combined evidence suggests that sAPPα is a necessary proliferation factor for non-neural and neural adult stem cells. Findings of Hartl *et al* also support a role of sAPPα in regulating CDK5 signalling [136]. sAPPα was able to reduce the expression and activity of CDK5, which is a known inducer of neurodegeneration. Interestingly, SORL1 is a known sAPPα receptor. Hartl *et al* demonstrated that, in SORL1 deficient neurons, sAPPα was no longer able to regulate CDK5 [136]. This combined evidence suggests that sAPPα functions as a neuroprotective factor.

APP has also been shown to be neuroprotective in the immediate period following traumatic brain injury (TBI) [137]. The sAPPα fragment of the APP695 isoform has four domains. One of these domains, D1, spans residues APP96-110 and has been shown to be involved in recovery from TBI [138]. The D1 domain is a growth factor like domain that contains heparin binding sites (HSBs). These sites in APP96-110 confer ability to bind heparin sulphate proteoglycans (HSPGs). Binding of HSBs to HSPGs promotes neurite outgrowth, a neurotrophic property. Corrigan *et al* (2014) demonstrated that the neuroprotective activity of APP96-110 is exerted via its heparin binding properties [138].

The APP AICD is involved in nuclear signalling

Aside from a proposed role of the AICD in AD neurodegeneration this fragment is thought to play a role in nuclear signalling. Von-Rotz *et al* (2004) identified transactivation activity for AICD using a yeast GAL4 transcription factor fused to both APP and AICD independently [139]. This

was dependent on Fe65, an APP-adaptor protein [140] and the Fe65-binding nuclear histone acetylase, Tip60 [7]. It has been shown that Fe65 can stabilise AICD, which localises to the nucleus with Fe65 [141].

Fe65 contains a WW domain along with two phosphotyrosine interaction domains (PID) [142]. Only PID2 in Fe65 is involved in binding APP, leaving PID1 and the WW domain free for other binding partners [142]. AICD can also be bound by the XII protein though its PID [142]. XII has been shown to regulate APP by slowing cellular APP processing and reducing secretion of Aβ-40 and Aβ-42 [143] by preventing translocation of APP into β- and γ-secretase-rich lipid rafts [144].

Fe65 and XII may compete for binding, as they both bind to the same motif of AICD. Many lines of evidence suggest that it is important for APP regulation (and possibly function) that there is a functional balance between Fe65 and XII [142]. A third APP binding protein, mDAB1, also binds to the same region as Fe65 and XII. Fe65 and mDAB1 interact with mammalian Ena (mEna) [145], which is found in areas of actin remodelling. mEna's interaction with APP through its binding proteins forms a potential molecular link between APP and the cytoskeleton [146]. This link to mEna may also explain the proposed functions of APP in cell adhesion, growth cone outgrowth and axon guidance [146]. The interaction with mEna has been implicated in a role for AICD signalling, in synaptic plasticity and memory [147].

The non-amyloidogenic APP pathway product p3

There is little known of the normal functional role of the non-amyloidogenic APP pathway product p3. It has, however, recently been demonstrated by one group to be neuroprotective [148]. Han *et al* compared the proteolytic products of APP and phenotypes between two mouse models, BACE1-

deficient and PSENs conditional double knockout (cDKO). Their findings suggest that p3 may be a neurotrophic factor deficient in cDKO mice and may play a neuroprotective role [148].

APP is also processed by lesser known secretases

Although most research into APP proteolysis focuses on the well-established processing by α-, β- and γ-secretases (and on the fragments generated by these cleavages), several other secretases have been discovered that process APP [reviewed in 149]. Aparagine endopeptidase or δ-secretase cleaves APP within its ectodomain at two positions, N^{373}-E^{374} and N^{585}-I^{586}, to produce three different fragments that can then be further processed by β- and γ-secretases (Figure 1.8). Of the three possible fragments sAPP$_{1-373}$ alone was observed to be toxic to primary cultured neurons. Interestingly, when knocking out AED in AD mouse models, a reduced amount of Aβ was deposited along with a reduction of other AD features [reviewed in 149].

Another lesser-known APP processing secretase is the matrix metalloproteinase MT5-MMP aka η-secretase. η-secretase cleaves APP between position N^{504}-M^{505}, resulting in a ~80–95 kDa soluble fragment and a membrane bound CTFη. This CTFη fragment is further processed by either α- or β-secretase to release two additional fragments, termed Aη-α and Aη-β. These two fragments have been identified in mouse brain and human CSF at levels 5 times that of Aβ. Furthermore, knockout of MT5-MMP in AD mouse models reduces Aβ deposition within the brain [reviewed in 149]. Discovery of these additional secretase activities, along with our emerging understanding of other APP fragments (discussed above), further highlights the need for a greater understanding of the complex processing of APP.
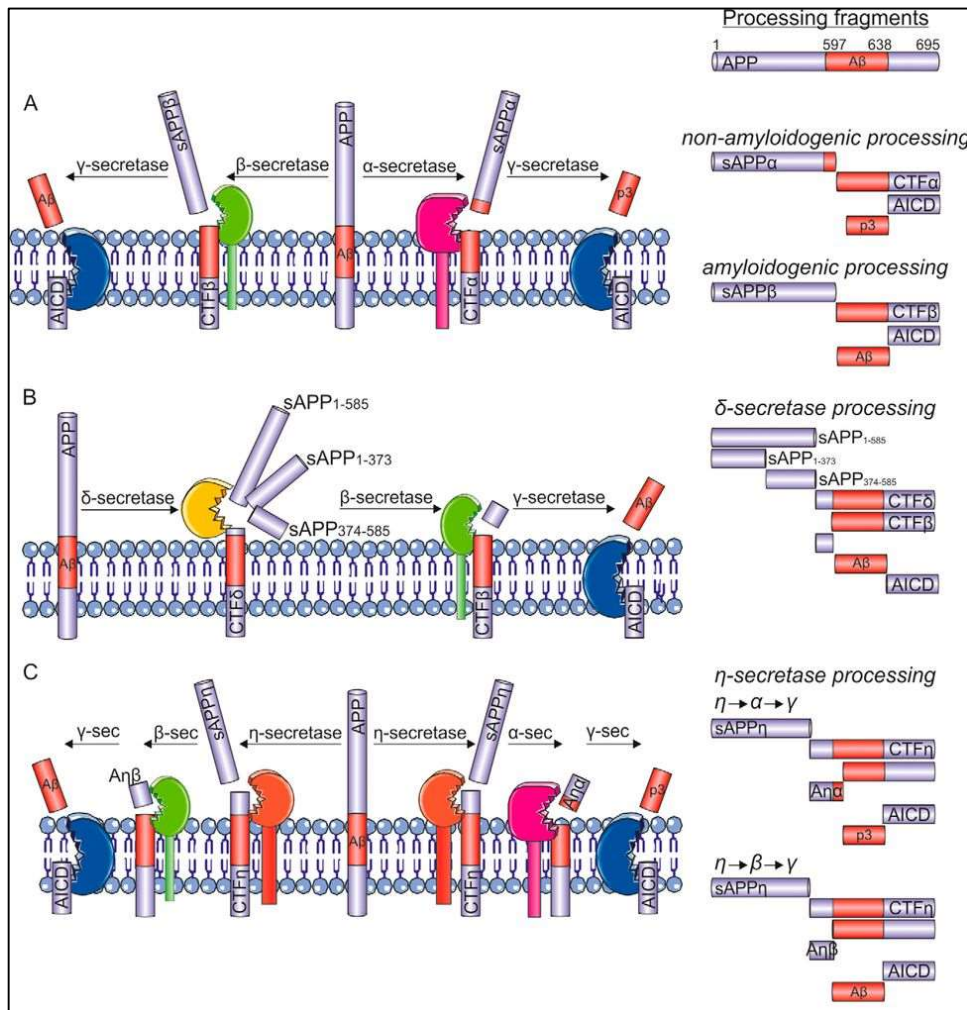
Figure 1.8. APP proteolysis by three distinct proteases. *A) both amyloidogenic and non-amyloidogenic γ-secretase processing of APP, B) δ-secretase processing of APP, and C) η-secretase processing of APP. Each of the fragments generated by these cleavage processes are presented in the right-hand section of the figure (reproduced without permission from [149]).*

Mutations in the APP gene contribute to fAD

APP has a proposed role in AD through its cleavage to form the Aβ peptide. Missense mutations in the APP gene can lead to dominantly inherited fAD. Currently 59 mutations are known and are

listed on Alzforum (https://www.alzforum.org/mutations/app). These mutations alter amino acid residues predominantly clustered around the Aβ peptide region of APP and are thought to cause approximately 15% of all cases of fAD. Furthermore, mutations in the APP gene are generally associated with an observed increase in the total levels of both Aβ-40 and Aβ-42, as well as an increase in the Aβ-42/Aβ-40 ratio. These factors lead to increased accumulation of the insoluble form of the Aβ peptide and, as a result, increased aggregation and formation of neurotoxic Aβ protofibrils. Generally, APP mutations are known to enhance AD pathology and cause a very early age of AD onset. In the next few sections we will discuss in detail mutations that are relevant to this thesis.

The Austrian mutation

One mutation of APP with an onset age as young as 36 years, is the "Austrian" mutation T714I. Kumar-Singh *et al* (2000) originally described an Austrian family with an autosomal dominant early-onset AD inheritance pattern [150]. The mutation exists in exon 17 of APP (T714I, according to APP770 isoform numbering) at the site of Aβ-42 cleavage. Analysis of brain sections of the proband revealed a huge load of non-neuritic/diffuse cotton-wool plaques (like patients with *PSEN1 Δ9* causing mutations that present with both AD and spastic paraparesis [151]), with N-truncated Aβ-42 composition [150]. Accumulation of hyperphosphorylated *tau* was also observed in neurites and there was a close association between blood vessels and fibrillar amyloid deposits [150]. Patients with this mutation presented with an early age of onset (approximately 36 years), rapid disease progression and early death (approximately 42 years), like PSEN1 AD mutations. Experiments found that Hek293T cells transiently transfected with T714I APP cDNA showed increased Aβ42 (3.5 fold) and decreased Aβ40 (68%) leading to a (10.8 fold) increase in the overall

ratio compared to wild type. Patient plasma had a 2.5 fold increase in ratio compared to the unaffected father and 1.7 fold compared to controls [150].

The Leuven mutation

The "Leuven" APP mutation was discovered in a Belgian patient presenting with early onset AD. This mutation (E682K), within exon 16 of the APP transcript, affects the alternative BACE1 cleavage site, β'-secretase [152]. BACE1 cleaves at both the β-secretase (Met671-Asp672) and β'-secretase (Tyr681-Glu682) sites in human APP [153], the latter of which is much less well-characterised. Cleavage at the β'-secretase site produces a C89 fragment which, when processed by the subsequent γ-secretase steps, generates truncated Aβ species [152]. Zhou *et al* (2011) found that the Aβ42/40 ratio was significantly increased in both mouse primary neurons and CHO cells harbouring this mutation [152]. Also, overall levels of both Aβ40 and Aβ42 were increased 2-3 fold. The E682K mutation prevents generation of Aβ40/Aβ42 species from the β' site, instead enhancing generation of these species from the β site [152]. Previous studies have indicated that low BACE1 levels result in an increase in cleavage at the β-site, whereas, at high BACE1 levels, cleavage at the β'-site becomes more predominant [154, 155]. Supplementing this is the finding that BACE1 has a much higher enzymatic efficiency towards peptides containing the β site than the β' site [156]. It is interesting to note that cleavage of endogenous APP at the β' site in rodents does not result in spontaneous development of amyloid plaques [152]. This may suggest that APP processing at the β' site may be important for normal APP metabolism [152]. Furthermore, these data suggest that processing of the β' site is anti-amyloidogenic, impeding amyloidogenic β site cleavage of APP. Processing of APP by BACE1 is disrupted when the β' cleavage site is blocked, which leads to increased generation of full length Aβ [152].

The Arctic mutation

The "Arctic" Alzheimer mutation, E693G, affects exon 17 of the APP open reading frame. Nilsberth *et al* (2001) found a decrease of 22-33% in the Aβ42/Aβ40 ratio in HEK293 cells transiently transfected with APP carrying the Arctic mutation [157]. In an Arctic mutation mouse model, enhanced accumulation of Aβ was observed inside neurons, strongly facilitating extracellular amyloid plaque deposition with age progression. Arctic Aβ peptides were observed to more easily form soluble protofibrils in vitro, leading to enhanced Aβ protofibril formation [158]. Experiments in HEK293 cells assessing the effects of the Arctic mutation on α-secretase cleavage revealed a reduction of APP abundance at the cell surface [157]. α-secretase cleavage is believed to occur at the cell surface but also in the trans-Golgi network [159, 160]. Hence, this mutation probably results in an inability of α-secretase to cleave APP, rather than representing a reduction in the enzymatic efficiency of APP processing by α-secretase. Due to reduced levels of fragments from α-secretase, there is a shift towards β-secretase cleavage, resulting in increased Aβ production [161]. β-secretase (BACE1) cleavage is thought to occur in acidified vesicles, primarily in the trans-Golgi network and in endosomal pathways [162].

1.5 Animal models of AD and related genes

The use of model organisms in experimental biology is an immensely informative research strategy. Model organisms provide effective *in vivo* systems in which the cellular biology involved in many complex diseases can be investigated. This is opposed to *in vitro* studies that are limited in their reproducibility of complex cellular system environments. In the following sections we will review the use of model animals specifically for AD research.

Invertebrate models of AD

Drosophila and *C.elegans* are well established model systems that have also been employed as tools for investigating AD neuropathology. Both model systems have orthologues of the *presenilin* gene [163]. *Drosophila* and *C.elegans* also have an APP-like protein (APPL) which has conserved function with vertebrate APP family members but lacks the Aβ domain [164].

As mentioned earlier *C.elegans* was instrumental in the identification of two of the γ-secretase components, APH-1 and PSENEN. *C.elegans* also contains three human presenilin orthologues, *sel-12*, *hop-1* and *spe-4* [165-167]. Human PSEN1 and PSEN2 are able to substitute for SEL-12 activity in *C.elegans,* suggesting SEL-12 protein has complete presenilin function [166]. As *C.elegans* APPL does not have an Aβ region, Link *et al* developed a transgenic model that expresses human Aβ (hAβ) in muscle cells of *C.elegans* [168]. Larvae of the hAβ transgenics presented with a muscle paralysis phenotype, which is likely due to increased toxicity in muscle cells [168]. Although Aβ expression is limited to muscle cells in this model, it can still be used to establish a link between Aβ expression and toxicity. Another study, using transgenic *C.elegans* lines, found that substituting Leu at position 17 for Pro and Met at positon 35 for Cys of Aβ resulted in no amyloid deposition, suggesting the importance of these single residues in the formation of amyloid [169]. *C.elegans* have also been used to confirm that the oxidative stress induced by Aβ (through generation of reactive oxygen species (ROS)) is related to Aβ toxicity [170].

Drosophila possess APPL (a member of the large family that includes APP and the APLP's (Figure 1.6)) and orthologues for all the major γ-secretase components [171]. Larval brains of transgenic *Drosophila*, expressing either wildtype (WT) or mutant human APP (hAPP), showed increased cell death [172]. Both Aβ and AICD of APP were required to induce this toxicity [172]. Crowther *et al* applied a more direct approach to look at Aβ-42 toxicity using *Drosophila* [173]. They found

that Aβ-42 expressed neuronally resulted in a reduced lifespan, locomotion defects and neurotoxicity [173]. Finelli *et al* generated *Drosophila* lines that express Aβ peptides. Overexpression of Aβ-40 and Aβ-42 was induced in neuronal cells of these lines [174]. Using similar transgenic *Drosophila* lines, Iijima *et al*, demonstrated that age-dependent learning deficiency, diffuse amyloid deposits and neurodegeneration were caused by Aβ-42 expression [175]. Similar learning deficits, without amyloid aggregation and neurodegeneration, were caused by Aβ-40 induction [175]. Furthermore, when *PSEN* mutations were introduced into *Drosophila PSEN*, the differences in activity in *Drosophila* of the mutant presenilins correlated with the age of onset of AD in humans [176]. *Drosophila* are also useful for screening potential AD drugs, such as γ-secretase inhibitors. In particular, the binding site of the γ-secretase inhibitor DAPT is conserved in *Drosophila* [177]. There have been numerous other studies utilising *Drosophila* to investigate the effects of Aβ in the context of AD. For a more comprehensive review of this topic please see the following book chapter [178].

Mouse models of AD

The mouse is the most commonly used animal model for the research into human neurodegenerative disease. Being a mammal, the mouse is more closely related to humans than invertebrate models. There are many transgenic mice that have been developed to model AD-related phenotypes [179]. Transgenic mice that express hAPP are some of the oldest and most widely used models. Transgenic hAPP mouse lines generally present with amyloid pathology and deficiency in memory [179]. These models of AD are complicated by the many factors that must be considered when using APP. One example of this is that hAPP lines can have different promoters to drive expression. There also are many hAPP isoforms, which further complicates

comparative analysis [179]. The different hAPP transgenic lines also express different AD associated mutations. A few lines have mutations that model those at the γ-secretase site. However, a majority of the mouse lines currently used express the Swedish double mutation (K670N/M671L) which exists at the β-secretase cleavage site [179]. Other lines either combine a γ-secretase cleavage site mutation with the Swedish mutation or add a mutation within the Aβ sequence (such as the E693G Arctic mutation discussed earlier).

Mouse models based on the FAD PSEN mutations have also been used for AD research. PSEN mutant transgenic mice show an increase in Aβ-42 levels, while Aβ-40 is unaffected [179]. However, these mice do not develop cognitive defects or AD pathology. This is probably due to the Aβ region in mouse APP (mAPP) being different to that of hAPP. Cleavage of mouse Aβ gives a shorter product than human Aβ and mouse Aβ peptides are much less efficient at aggregation [179]. It appears that development of AD phenotypes in mice requires human Aβ expression [179]. To test this, PSEN mutant mice were crossed to mice with either transgenic hAPP or mAPP [179]. When crossed to a mouse line overexpressing mAPP offspring have no AD pathology or cognitive deficits. When crossed to a hAPP transgenic mouse line offspring have extensive plaque deposition and behavioural deficits [179]. The hAPP/PSEN1 lines that are more commonly used today are generated via co-injection of PSEN and hAPP transgenes. A line containing the Swedish mutation and PSEN1 harbouring the ΔE9 mutation (APPswe/PS1ΔE9 mice) develops behavioural deficits and amyloid plaques at approximately 6-7 months of age. The 5XFAD line combines multiple FAD mutations, 2 in PSEN1 and 3 in APP. These mice express Aβ-42 at high levels, develop amyloid plaques and cognitive deficits at around 4 months. 5XFAD mice also develop neuron loss, which most other hAPP/PSEN1 and hAPP models do not [179]. Contrary to the mouse models overexpressing hAPP described above, Saito *et al* developed simple models that overproduce Aβ-

42 without overexpressing hAPP [180]. One of these mouse lines, carrying a combination of the Arctic, Swedish and Beyreuther/Iberian mutations, displayed aggressive Aβ amyloidosis in an age-dependent manner [180]. The amyloid deposition pattern of these mice was consistent with human Arctic mutation carrier pathology. These mice should provide an interesting tool for investigating select features of AD biology.

For a more extensive review of AD mouse models please see the following reference [178]. It is also worth mentioning here that, while transgenic mouse models have been useful in investigating certain aspects of the disease, the transcriptomic profiles of these mouse models are not consistent with those of human AD [181]. This observation suggests that these transgenic mouse models are not necessarily the best way to model AD moving forward.

Zebrafish as a model for AD

Danio rerio, commonly known as Zebrafish, are a useful model for the study of genetic diseases/conditions as they contain orthologues of many human genes. Zebrafish embryos can be readily manipulated, allowing for knockdown of gene expression with morpholino oligonucleotides (MO) or overexpression of genes by injection of sense RNA [182]. Zebrafish are also genetically malleable, allowing for more precise genome manipulations by employing genome engineering technologies, which will be discussed in more detail in below. Furthermore, the zebrafish genome has been sequenced completely, which makes it very accessible for genetic screens. In addition to this, short generation times and large numbers of offspring enable significantly larger screens over time frames much shorter than could be accomplished using rodent models. Zebrafish provide an additional benefit over mice due to low maintenance costs, as

well as the potential to visualise developmental processes through their transparent embryos. Most importantly, as a vertebrate, they provide many of the benefits of invertebrate models such as *C. elegans* (worm) and *D. melanogaster* (fruit fly) while maintaining much greater homology to humans.

Zebrafish contain orthologues of many of the human genes implicated in AD pathogenesis. Orthologues of Human *PSEN1* and *PSEN2* orthologues identified in zebrafish are *psen1* [183] and *psen2* [184] respectively. An orthologue of human SORL1 also exists in zebrafish, *sorl1* [185]. Zebrafish form a part of the Teleost infraclass of bony fishes and as such their genomes contain many duplicated genes [186]. The human APP gene has two co-orthologues in zebrafish, *appa* and *appb* [187]. Zebrafish *appa* is more closely related to longer isoforms of APP (770 and 752), while *appb* is closer to APP695 [187]. Zebrafish Appa and Appb have approximately 70% identity to human APP695 at the amino acid level, with 80% identity to the Aβ region and 95% identity within the transmembrane domain [187]. Orthologues of the other key components of γ-secretase also exist in the zebrafish genome; *PSENEN* (*psenen*) [13], *NCTN* (*ncstn*) [188, 189] and *APH1b* (*aph1b*) [16]. β-secretase orthologues have also been found in zebrafish, *bace1* [190] and *bace2* [191].

Zebrafish have been used extensively as a model for research into fAD and other neurodegenerative diseases [178, 192]. The two *PSEN* gene orthologues in zebrafish are expressed ubiquitously during embryo development [183, 184]. The residues that give PSEN1 catalytic function are also conserved in zebrafish [193]. Studies conducted *in vitro* demonstrated that *psen1* is able to functionally replace *PSEN1* and produces Aβ-42 from fAD Swedish mutant APP in an efficient manner [183]. A study in which a *psen1* translation-blocking MO was injected into fertilised embryos, revealed a reduction in the Notch target gene *hairy-related 1* (*her1*). This

indicated that loss of *psen1* caused a reduction of Notch signalling [193]. Hence zebrafish *psen1* shows conserved function with human *PSEN1*. Having only recently been identified as a genetic risk factor for fAD, *sorl1* has not been as extensively investigated using zebrafish. However, Lee *et al* (2017) found that zebrafish Sorl1 protein is highly conserved with human SORL1 [185]. They also placed *sorl1* expression in neural tissue in 72 hours post fertilisation (hpf) zebrafish [185]. Zebrafish orthologues of the other γ-secretase components have not been thoroughly analysed. However, Campbell *et al* (2006) demonstrated that blocking translation of both *psenen* and *aph1b* in zebrafish produced loss of Notch signalling phenotypes. Blockage of *psenen* also destabilised Psen1 protein and increased apoptotic induction [194]. Furthermore, Lim *et al* (2015) observed *ncstn* in the developing embryo and found it to be expressed most highly in ventricular cells of the developing brain at 24hpf [189].

Zebrafish co-orthologues of APP, *appa* and *appb*, are both widely expressed in the developing embryo from the mid-gastrulation stage [187]. 24 hour old embryos express both genes in the developing forebrain and elsewhere, *appb* alone is expressed in the spinal cord [187]. The function of Appa and Appb proteins have also been investigated using translation-blocking MOs [195]. Inhibition of Appb conferred a defect in convergent extension movements and reduced body length. This could be rescued by injection of mRNA encoding human APP into Appb deficient embryos [195]. Defective neural development including defective axonal outgrowth and synapse formation have also been demonstrated by loss of Appb activity [196]. Such studies as described above demonstrate the usefulness of zebrafish embryos in analysing different fAD genes.

Finally, while both brain-derived neurotrophic factor (BDNF) and TrkB have been analysed using zebrafish [197], studies investigating p75^{NTR} in the zebrafish are limited to a handful of papers where it has been utilised to investigate Nogo/Nogo receptor complex signalling [198, 199]. We

could find no studies to suggest that zebrafish NRH1 has been investigated previously using this animal model.

Our laboratory is the leading group using zebrafish to model AD. We have previously investigated extensively the roles of the zebrafish orthologues of PSEN1 and PSEN2 in various biological processes. The laboratory's current goal is to generate a catalogue of fAD-like and non-fAD mutations in the zebrafish *psen1*, *psen2*, *appa*, *appb* and *sorl1* genes for transcriptomic analyses in the hopes of identifying a molecular "signature" of AD. Thus far, mutations have been generated in *psen1*, *psen2* and *sorl1* (published [200-202] and unpublished) and some early analyses have revealed interesting effects on the hypoxic response and brain energy production in these model fish. In order to contribute to this wider laboratory project, the work presented in this thesis aimed to generate both fAD-like and null mutations in zebrafish *appa* and *appb*. This was attempted through generating mutations in endogenous zebrafish *appa* and *appb* genes with genome editing technologies, which will be introduced in the following section.

1.6 Techniques for genome engineering in the zebrafish

As we have just discussed, zebrafish are a useful genetic tool for studying molecular mechanisms and disease pathogenesis. Zebrafish embryos are easy to manipulate by injection with a variety of genome engineering tools. Morpholinos can be used for antisense gene knockdown [192] or to interfere with intron splicing [203] resulting in aberrant protein products. Conversely, direct injection of sense mRNA allows for overexpression of a gene of interest. While these two methods are successful in interfering with regular expression of a gene, their effects only persist for a short time. Targeted induced local lesions in genomes (TILLING) was the first approach for reverse genetic engineering in zebrafish that was successful in generating germline mutations in a specific

gene [204]. TILLING enabled the recovery of rare genomic mutations that can be analysed phenotypically at a later date [204]. Transposons are another useful reverse genetic approach. The Tol2 transposon system is used to generate transgenic zebrafish with germline mutations via transposition. A cell-type specific enhancer/promoter placed upstream of green fluorescence protein (*gfp*) on a Tol2 construct allows for the insertion of genes that can be visualised by expression of GFP in a cell-type specific manner [205]. The GAL4-UAS and Cre/loxP systems can be used to generate conditionally expressed transgenic zebrafish lines [204]. Both systems use a bipartite approach, in which two separate transgenic lines are generated and must be crossed in order to elicit the transgenes event. GAL4 can be driven in a temperature inducible manner using the heat shock protein 70 (hsp70) in zebrafish [204]. Research using zebrafish originally applied the Cre/LoxP system for the conditional expression of oncogenes. The system also used hsp70 promoter to induce Cre expression. Additionally, the temporal activity of Cre can be controlled through its fusion with a modified estrogen receptor that binds tamoxifen [204]. Early use of zebrafish for transgenic analyses was at a disadvantage relative to mice, due to the inability to generate targeted mutations. Recently, three techniques for targeted modification of genomes have been established.

Zinc finger nucleases (ZFNs) and Transcription activator-like effector nucleases (TALENs)

ZFNs and TALENs are both comprised of a nonspecific DNA cleavage segment and programmable, sequence-specific DNA-binding units [206]. The ZFNs and TALENs cause the targeted induction of double-stranded breaks (DSBs) that stimulate cellular DNA repair mechanisms such as Non Homologous End Joining (NHEJ) and Homology Directed Repair (HDR) [206]. ZFNs and TALENS are useful tools for the generation of targeted mutations. They

are both versatile due to the ability to engineer the DNA binding domains, which can be customised for the recognition of practically any sequence [206].

ZNFs are a chimeric fusion of a Cys2His2 zinc finger array and a Fok1 endonuclease cleavage domain [204]. The Cys2His2 array of a zinc finger contributes specificity for the DNA site of interest [204]. The target sites of ZFNs are comprised of two zinc finger binding sites separated by a 5-7bp spacer sequence, which is recognised by Fok1 for cleavage [206]. A pair of ZFNs is required, due to the necessity of Fok1 of dimerise in order to exert its cleavage ability [204, 206]. A pair of ZFNs must bind correctly to their target sequence in order to generate a DSB [204]. A ZFN normally has 3-6 fingers and each finger recognises 3bp of DNA. When combined, a pair of ZFNs can recognise 18-36bp of DNA sequence, which is dependent on the number of fingers. This large recognition site length means a ZFN has specificity within 68billion bp of DNA [204, 206]. Occasionally introduction of unwanted indels at the cleavage site occurs, due to repair by NHEJ. There can also be off target effects. ZFNs are limited by their inability to generate DNA changes with high *in vivo* activity [204].

Transcription activator-like effector nucleases (TALENs) have more recently been used as an alternative to ZFNs for editing genomes by the induction of DSBs. Like ZFNs, TALENs consist of a non-specific *Fok*I endonuclease domain, which is fused to a customisable and site-specific DNA-binding transcription activator-like effector (TALE) domain from the plant pathogenic bacterial species *Xanthomonas* [207]. This TALE domain is made up of a sequence of TALE repeats that each recognise one specific nucleotide of DNA [208]. Each highly conserved repeat is 32-35 amino acids long, with residues 12 and 13 (known as repeat variable di-residues or RVDs) enabling the TALEN to specifically bind a single DNA nucleotide [206-208]. Due to the nature of TALE proteins, a 5' T immediately upstream of the binding site is required for function [207, 209].

TALE repeat RVDs have been identified that can recognise each of the four DNA nucleotides. These can be assembled in such a way that virtually any DNA sequence that starts with a 5' T can be targeted and bound by TALENs [207]. Similarly to ZFNs, a pair of customised TALENs is needed to cleave a target site due to the dimerization requirements of Fok1. An optimal spacer length is required between a pair of TALENs so that Fok1 is in the correct position to dimerise and facilitate cleavage [208]. Spacer lengths for TALENs can vary between 12-20bp [206]. Hwang *et al* (2014) have found that spacer lengths of 16-18bps are effective in zebrafish [208]. Reyon *et al* suggest that TALENs comprised of 14.5-16.5-repeat arrays are most likely to be effective [210]. This length array would recognise 16-18 nucleotides in total (including the 5' T).

Compared to ZFNs, the method of target recognition and binding of TALENs is much simpler and more predictable, making it a good alternative to ZFNs [209]. TALENs have been used previously to successfully induce disruption of different zebrafish genes in a targeted manner [207]. However, there was only germline transmission for two of the genes [207]. Hwang *et al* (2014) managed to successfully mutate their target sites in zebrafish more than 80% of the time [208]. Somatic cells were found to be mutated by TALENs at rates of as high as 76%. Furthermore, they also found that TALENs efficiently induced heritable mutations in zebrafish [208].

Better sequence targeting specificity could be achieved by constructing TALENs that target more than 12bp of DNA. However, the highly repetitive nature of TALE repeats makes this practically impossible [209]. TALENs and ZFNs function in a similar way in that they both can induce indel mutations in the spacer region between binding sites. However, while ZFNs and TALENs can cleave DNA with similar efficiency when directed to the same target, TALENs usually have a

higher targeting efficiency and fewer off-target effects [209]. A study by Bedell *et al* (2012) used TALENs in conjunction with a short single stranded oligonucleotide donor to precisely insert a mutation into the genome of zebrafish [211]. TALENs are simple to design, have robust activity and are nearly limitless in their targeting range making them a very attractive platform for genome editing [207].

Clustered regulatory interspersed short palindromic repeats (CRISPRs)

A recently identified whole genome editing technique has also been adopted for use in the zebrafish, mouse, maize, drosophila and many others as an alternative to ZFNs and TALENS. This technique makes use of the CRISPR and CRISPR associated (Cas) proteins, which originally functioned as a form of protection for bacteria and archaea against invading viruses and plasmids [212]. In these organisms the CRISPR system acts as a form of "genetic memory" of the viral infection, allowing them to recognise and remove viruses that have attacked in the past [212]. Six different CRISPR–Cas types and at least 29 subtypes have now been discovered [213]. The type II CRISPR/Cas system is most commonly used for genome editing and consists of 4 *cas* genes originally identified in *Streptococcus pyogenes* [212]. The most heavily utilised of these, Cas9, is commonly employed to cleave the DNA target sequence [214, 215].

CRISPR RNA (crRNA) is an array of repeats interspaced by short variable sequences (protospacers) (Figure 1.9, 1). In bacteria and archaea, the first stage of the CRISPR process involves the acquisition of virus DNA and these "protospacers" represent short segments of DNA from the invading virus or plasmid [212]. It is thought that the protospacer adjacent motif (PAM) may be involved in protospacer acquisition, which results in PAM flanking the protospacers on the 3' side of this 'foreign DNA' within the crRNA in this type II system [212, 216]. PAM is

required for Cas9 activity, although the specific mechanism behind this remains unknown [215]. Due to the positioning of PAM (3' of each protospacer) it is possible that it's function is to help Cas9 determine the spacer orientation within the crRNA, providing a "polarity" for Cas9 to recognise substrates for cleavage (Figure 1.9, 4) [212].

The adapted crRNA that is used for genome editing encodes a single 20bp guide RNA (gRNA) sequence to direct Cas9 to a 20bp target site [214]. This gRNA immediately precedes the (type II) PAM sequence NGG [217]. Deep sequencing of RNA from the (type II) bacteria *S.pyogenes* crRNA revealed the presence of trans-acting antisense RNA (tracrRNA) [212]. This tracrRNA serves two functions in the type II CRISPR system; tracrRNA forms a duplex with pre-crRNA and recruits RNAse III an enzyme which processes the tracrRNA/pre-crRNA duplex to form mature crRNA (Figure 1.9, 2). tracrRNA also recruits Cas9 to the site of interest allowing for the invading virus/plasmid to be cleaved and degraded [215]. In the zebrafish editing system the tracrRNA is included with the gRNA in a single transcript which is given the abbreviation "sgRNA" [217]. sgRNA is already in the form of a mature crRNA and can direct cas9-mediated cleavage of target DNA [217].

Cas9 cleavage functions by generating double-stranded breaks (DSBs) (Figure 1.9, 6) [218]. These DSBs are repaired by either error prone non-homologous end-joining (NHEJ) resulting in insertions or deletions (indels) or, alternatively, homologous recombination (HR) [218]. Repair by HR requires co-injection of a single strand oligonucleotide (SSO) as a donor template, which encodes the desired mutation [218].
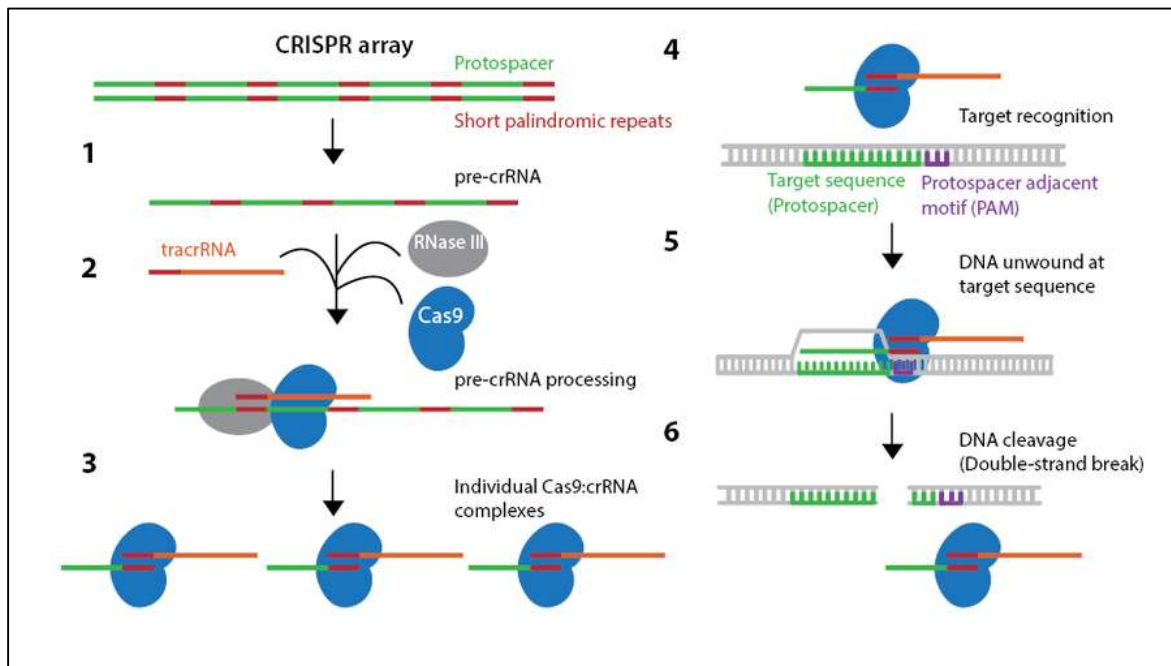
Figure 1.9. Summary of the bacterial type II CRISPR system. *1) generation of the pre-crRNA, 2) tracrRNA recruits RNase III and Cas9 to process pre-crRNA into crRNA, 3) tracrRNA:Cas9:crRNA complexes are formed, 4) tracrRNA, as a part of the crRNA:tracrRNA:Cas9 complex, directs Cas9 to the DNA region complementary to the crRNA, only when there is a directly adjacent PAM (3' of the crRNA binding site), 5&6) Induction of a DSB by Cas9. (reproduced without permission from* https://www.addgene.org/crispr/history/*).*

Using genome engineering tools to generate mutations of interest

There are both advantages and disadvantages for the use of CRISPR's over ZFN's and TALEN's for genome modifications. One such advantage is that a single customised gRNA is required to target a specific sequence, compared to the need to design and assemble two TALENs or ZFNs for each site [217]. One disadvantage is that there are additional constraints when designing and using CRISPRs. For successful utilisation of the Cas9 to recognise target sequence, a protospacer

adjacent motif (PAM) sequence is required 3' of the gRNA and upstream of the crRNA binding region [217]. Also, the use of the T7 promoter to make gRNAs means that GG is required at the 5' end of the transcript, which reduces the range of targetable sequences [217]. Taken together these constraints would mean that any sequence of the form 5' GG-N(18)-NGG 3' is available for targeting by CRISPR/Cas9, occurring once in every 128bps of random DNA sequence [217]. However, a later study by Hwang *et al* (2013) found that this targeting range could be expanded to once in every 8bps by relaxing the T7 promoter rule [219]. This was accomplished by the addition of GG to the 5' end of the 20nt gRNA sequence that were not complimentary to the genomic DNA target sequence [219]. Another common problem with the current system (and ZFNs and TALENs) is the occurrence and frequency of possible off target effects, due to binding at sites of less than complete homology [220]. Shen *et al* (2014) looked at the potential for using Cas9 nickase rather than endonuclease to reduce the prevalence of these off target effects *in vivo* using mouse embryos [221]. Nickase creates a "nick" in a single strand of the genomic DNA which allows it to be corrected by the endogenous base-excision repair pathway [221].

Many studies have attested to the usefulness and adaptability of CRISPRs as a genetic tool. Genome editing with CRISPR/Cas9 in two separate studies demonstrated that multiplex gene knockouts can be performed both simultaneously and highly efficiently in a single step using stem cells and mice [222, 223]. In this single step process, co-injection of two sgRNAs is performed into a target cell. Both of these studies resulted in biallelic gene conversions or inactivation's [222, 223]. Results presented by Wang *et al* (2013) suggest that injection with a single sgRNA results in up to 95% of mice carrying biallelic mutations in the targeted gene, whereas co-injection of two different sgRNAs, results in up to 80% of mice carrying biallelic mutations in both target genes [222]. Jao *et al* (2013) showed that mutations could be efficiently transmitted through the germline

into the F1 generation of zebrafish [223]. Another study, by Yang *et al* (2013), also demonstrated one step generation of mutants at multiple gene loci using mouse embryos [224]. Hruscha *et al* (2013) adapted the CRISPR protocol to synthesise gRNA without cloning and demonstrated efficient mutagenesis, knock-in (KI), and transmission to the germline that was comparable to ZFN and TALEN approaches [225]. They also provided evidence that off-target effects were limited, and that KI was efficient, particularly when they pre-selected founder fish [225].

One interesting utilisation of the CRISPR system in mice demonstrated CRISPR/Cas9-mediated correction of *Fah* mutation in hepatocytes in an adult mouse model of the human disease hereditary tyrosinemia (HTI) [226]. Adult FAH$^{mut/mut}$ mice were given hydrodynamic tail vein injections with ssDNA oligonucleotide, plus a pX330 vector expressing Cas9 and one of three individually cloned sgRNA's (FAH1-3) targeting the *Fah* locus. This study demonstrated that CRISPR/Cas9 could functionally rescue the Fah deficiency-induced liver damage in mice, demonstrating a potential for this system to correct disease genes *in vivo* in adult mouse liver using a CRISPR-Cas9 system [226]. Transient expression of Cas9, sgRNA and a co-injected ssDNA by non-viral hydrodynamic injection was sufficient to restore the weight loss of a mouse model of HTI [226]. This study demonstrated a previously unrealised potential of the CRISPR system to correct mutant gene phenotypes in adult specimens.

The discovery and application of CRISPR-Cas9 has transformed the field of genome engineering and further development of this technology is rapidly progressing. A recent review highlights further advances in the technology [213]. Some of these advances that can be applied to genome engineering in the zebrafish will be discussed further in subsequent chapters of this thesis.

1.7 Conclusion

γ-secretase is a protease complex that is important for the processing of more than 100 substrates. Processing of APP by γ-secretase is thought to be a critical component of AD pathogenesis. Although much research has been undertaken to identify the effects of γ-secretase in AD, the specific mechanism through which it recognises and selects its substrates remains elusive. The similar sequence and structure, but distinct transmembrane cleavage events, of p75[NTR] and its homolog NRH1 may allow us to further investigate this mechanism. Furthermore, 59 mutations in APP have been identified that are thought to cause fAD. However, the mechanism through which AD develops and progresses remains incompletely understood and there is still no single hypothesis that accurately predicts disease susceptibility. The zebrafish is an excellent model for us to further investigate the molecular effects of fAD-like mutations, in the hopes of better understanding what specific factors inevitably results in AD.

1.8 References

1. Beel, A.J. and C.R. Sanders, *Substrate specificity of gamma-secretase and other intramembrane proteases.* Cell Mol Life Sci, 2008. **65**(9): p. 1311-34.
2. Haapasalo, A. and D.M. Kovacs, *The many substrates of presenilin/gamma-secretase.* J Alzheimers Dis, 2011. **25**(1): p. 3-28.
3. Murphy, M.P. and H. LeVine, 3rd, *Alzheimer's disease and the amyloid-beta peptide.* Journal of Alzheimer's disease : JAD, 2010. **19**(1): p. 311-323.
4. D'Onofrio, G., et al., *Advances in the identification of gamma-secretase inhibitors for the treatment of Alzheimer's disease.* Expert Opin Drug Discov, 2012. **7**(1): p. 19-37.
5. Kanning, K.C., et al., *Proteolytic processing of the p75 neurotrophin receptor and two homologs generates C-terminal fragments with signaling capability.* J Neurosci, 2003. **23**(13): p. 5425-36.
6. Skeldal, S., et al., *Proteolytic processing of the p75 neurotrophin receptor: A prerequisite for signalling?: Neuronal life, growth and death signalling are crucially regulated by intra-membrane proteolysis and trafficking of p75(NTR).* Bioessays, 2011. **33**(8): p. 614-25.
7. Gentry, J.J., P.A. Barker, and B.D. Carter, *The p75 neurotrophin receptor: multiple interactors and numerous functions.* Prog Brain Res, 2004. **146**: p. 25-39.
8. Coulson, E.J., et al., *The role of the p75 neurotrophin receptor in cholinergic dysfunction in Alzheimer's disease.* Neuroscientist, 2009. **15**(4): p. 317-23.
9. Struhl, G. and A. Adachi, *Requirements for presenilin-dependent cleavage of notch and other transmembrane proteins.* Mol Cell, 2000. **6**(3): p. 625-36.

10.   Wolfe, M.S., *APP, Notch, and presenilin: molecular pieces in the puzzle of Alzheimer's disease.* Int Immunopharmacol, 2002. **2**(13-14): p. 1919-29.

11.   Brunkan, A.L. and A.M. Goate, *Presenilin function and gamma-secretase activity.* J Neurochem, 2005. **93**(4): p. 769-92.

12.   Goutte, C., et al., *APH-1 is a multipass membrane protein essential for the Notch signaling pathway in Caenorhabditis elegans embryos.* Proc Natl Acad Sci U S A, 2002. **99**(2): p. 775-9.

13.   Francis, R., et al., *aph-1 and pen-2 are required for Notch pathway signaling, gamma-secretase cleavage of betaAPP, and presenilin protein accumulation.* Dev Cell, 2002. **3**(1): p. 85-97.

14.   Wolfe, M.S., *The gamma-secretase complex: membrane-embedded proteolytic ensemble.* Biochemistry, 2006. **45**(26): p. 7931-9.

15.   Shirotani, K., et al., *Identification of distinct gamma-secretase complexes with different APH-1 variants.* J Biol Chem, 2004. **279**(40): p. 41340-5.

16.   Shah, S., et al., *Nicastrin functions as a gamma-secretase-substrate receptor.* Cell, 2005. **122**(3): p. 435-47.

17.   Chavez-Gutierrez, L., et al., *Glu(332) in the Nicastrin ectodomain is essential for gamma-secretase complex maturation but not for its activity.* J Biol Chem, 2008. **283**(29): p. 20096-105.

18.   Zhao, G., et al., *Gamma-secretase composed of PS1/Pen2/Aph1a can cleave notch and amyloid precursor protein in the absence of nicastrin.* J Neurosci, 2010. **30**(5): p. 1648-56.

19.   Wolfe, M.S., *Substrate recognition and processing by γ-secretase.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 2019: p. 1-7.

20.   Holmes, O., et al., *Pen-2 Is Essential for gamma-Secretase Complex Stability and Trafficking but Partially Dispensable for Endoproteolysis.* Biochemistry, 2014: p. 4393-4406.

21.   Prokop, S., et al., *Requirement of PEN-2 for stabilization of the presenilin N-/C-terminal fragment heterodimer within the gamma-secretase complex.* J Biol Chem, 2004. **279**(22): p. 23255-61.

22.   Watanabe, N., et al., *Pen-2 is incorporated into the gamma-secretase complex through binding to transmembrane domain 4 of presenilin 1.* J Biol Chem, 2005. **280**(51): p. 41967-75.

23.   Kim, S.H. and S.S. Sisodia, *Evidence that the "NF" motif in transmembrane domain 4 of presenilin 1 is critical for binding with PEN-2.* J Biol Chem, 2005. **280**(51): p. 41953-66.

24.   Kim, S.H. and S.S. Sisodia, *A sequence within the first transmembrane domain of PEN-2 is critical for PEN-2-mediated endoproteolysis of presenilin 1.* J Biol Chem, 2005. **280**(3): p. 1992-2001.

25.   Lee, S.F., et al., *A conserved GXXXG motif in APH-1 is critical for assembly and activity of the gamma-secretase complex.* J Biol Chem, 2004. **279**(6): p. 4144-52.

26.   Dries, D.R. and G. Yu, *Assembly, maturation, and trafficking of the gamma-secretase complex in Alzheimer's disease.* Curr Alzheimer Res, 2008. **5**(2): p. 132-46.

27.   Parks, A.L. and D. Curtis, *Presenilin diversifies its portfolio.* Trends Genet, 2007. **23**(3): p. 140-50.

28.   Ahn, K., et al., *Activation and intrinsic gamma-secretase activity of presenilin 1.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21435-40.

29.   Newman, M., et al., *Robust homeostasis of Presenilin1 protein levels by transcript regulation.* Neurosci Lett, 2012. **519**(1): p. 14-9.

30.   Le Brocque, D., et al., *Processing of the Alzheimer's disease amyloid precursor protein in Pichia pastoris: immunodetection of alpha-, beta-, and gamma-secretase products.* Biochemistry, 1998. **37**(42): p. 14958-65.

31.   Taniguchi, Y., et al., *Notch receptor cleavage depends on but is not directly executed by presenilins.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 4014-9.

32.   Yagishita, S., et al., *Abeta46 is processed to Abeta40 and Abeta43, but not to Abeta42, in the low density membrane domains.* J Biol Chem, 2008. **283**(2): p. 733-8.

33.   De Strooper, B., et al., *Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein.* Nature, 1998. **391**(6665): p. 387-90.

34.   Wolfe, M.S., et al., *Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity.* Nature, 1999. **398**(6727): p. 513-7.

35.	Sykes, A.M., et al., *The effects of transmembrane sequence and dimerization on cleavage of the p75 neurotrophin receptor by gamma-secretase.* J Biol Chem, 2012. **287**(52): p. 43810-24.

36.	Lu, P., et al., *Three-dimensional structure of human γ-secretase.* Nature, 2014. **512**: p. 166.

37.	Bai, X.-c., et al., *Sampling the conformational space of the catalytic subunit of human γ-secretase.* eLife, 2015. **4**: p. e11182.

38.	Sun, L., et al., *Structural basis of human γ-secretase assembly.* Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(19): p. 6003-6008.

39.	Bai, X.-C., et al., *An atomic structure of human γ-secretase.* Nature, 2015. **525**(7568): p. 212-217.

40.	Zhou, R., et al., *Recognition of the amyloid precursor protein by human γ-secretase.* Science, 2019. **363**(6428): p. eaaw0930.

41.	Yang, G., et al., *Structural basis of Notch recognition by human γ-secretase.* Nature, 2019. **565**(7738): p. 192-197.

42.	Urra, S., et al., *TrkA receptor activation by nerve growth factor induces shedding of the p75 neurotrophin receptor followed by endosomal gamma-secretase-mediated release of the p75 intracellular domain.* J Biol Chem, 2007. **282**(10): p. 7606-15.

43.	Skeldal, S., et al., *Mapping of the interaction site between sortilin and the p75 neurotrophin receptor reveals a regulatory role for the sortilin intracellular domain in p75 neurotrophin receptor shedding and apoptosis.* J Biol Chem, 2012. **287**(52): p. 43798-809.

44.	Yao, X.Q., et al., *p75NTR ectodomain is a physiological neuroprotective molecule against amyloid-beta toxicity in the brain of Alzheimer's disease.* Mol Psychiatry, 2015. **20**(11): p. 1301-10.

45.	Zampieri, N., et al., *Cleavage of p75 neurotrophin receptor by alpha-secretase and gamma-secretase requires specific receptor domains.* J Biol Chem, 2005. **280**(15): p. 14563-71.

46.	Weskamp, G., et al., *Evidence for a critical role of the tumor necrosis factor alpha convertase (TACE) in ectodomain shedding of the p75 neurotrophin receptor (p75NTR).* J Biol Chem, 2004. **279**(6): p. 4241-9.

47.	Forsyth, P.A., et al., *p75 neurotrophin receptor cleavage by alpha- and gamma-secretases is required for neurotrophin-mediated proliferation of brain tumor-initiating cells.* J Biol Chem, 2014. **289**(12): p. 8067-85.

48.	Jayne, T., *Developing in vivo assays for investigating p75NTR and NRH1 transmembrane cleavage events using zebrafish*, in *Biological Sciences*. 2013, The University of Adelaide: Adelaide, Australia.

49.	Hippius, H. and G. Neundorfer, *The discovery of Alzheimer's disease.* Dialogues Clin Neurosci, 2003. **5**(1): p. 101-8.

50.	Alzheimer, A., et al., *An English translation of Alzheimer's 1907 paper, "Uber eine eigenartige Erkankung der Hirnrinde".* Clin Anat, 1995. **8**(6): p. 429-31.

51.	Ball, M.J., *Neuronal loss, neurofibrillary tangles and granulovacuolar degeneration in the hippocampus with ageing and dementia.* Acta Neuropathologica, 1977. **37**(2): p. 111-118.

52.	Madsen, J.B., J. Folke, and B. Pakkenberg, *Stereological Quantification of Plaques and Tangles in Neocortex from Alzheimer's Disease Patients.* J Alzheimers Dis, 2018. **64**(3): p. 723-734.

53.	Cotman, C.W. and J.H. Su, *Mechanisms of neuronal death in Alzheimer's disease.* Brain Pathol, 1996. **6**(4): p. 493-506.

54.	Butterfield, D.A., et al., *Amyloid beta-peptide and amyloid pathology are central to the oxidative stress and inflammatory cascades under which Alzheimer's disease brain exists.* J Alzheimers Dis, 2002. **4**(3): p. 193-201.

55.	Aslan, M. and T. Ozben, *Reactive oxygen and nitrogen species in Alzheimer's disease.* Curr Alzheimer Res, 2004. **1**(2): p. 111-9.

56.	Huang, X., et al., *The A beta peptide of Alzheimer's disease directly produces hydrogen peroxide through metal ion reduction.* Biochemistry, 1999. **38**(24): p. 7609-16.

57.	Huang, X., et al., *Cu(II) potentiation of alzheimer abeta neurotoxicity. Correlation with cell-free hydrogen peroxide production and metal reduction.* J Biol Chem, 1999. **274**(52): p. 37111-6.

58.   Lustbader, J.W., et al., *ABAD directly links Abeta to mitochondrial toxicity in Alzheimer's disease.* Science, 2004. **304**(5669): p. 448-52.

59.   Sastre, M., T. Klockgether, and M.T. Heneka, *Contribution of inflammatory processes to Alzheimer's disease: molecular mechanisms.* Int J Dev Neurosci, 2006. **24**(2-3): p. 167-76.

60.   Cameron, B. and G.E. Landreth, *Inflammation, microglia, and Alzheimer's disease.* Neurobiol Dis, 2010. **37**(3): p. 503-9.

61.   Wyss-Coray, T., et al., *Adult mouse astrocytes degrade amyloid-beta in vitro and in situ.* Nat Med, 2003. **9**(4): p. 453-7.

62.   Tuppo, E.E. and H.R. Arias, *The role of inflammation in Alzheimer's disease.* Int J Biochem Cell Biol, 2005. **37**(2): p. 289-305.

63.   St George-Hyslop, P.H., *Molecular genetics of Alzheimer's disease.* Biol Psychiatry, 2000. **47**(3): p. 183-99.

64.   Farrer, L.A., et al., *Segregation analysis reveals evidence of a major gene for Alzheimer disease.* Am J Hum Genet, 1991. **48**(6): p. 1026-33.

65.   Yabu, T., et al., *Characterization of zebrafish caspase-3 and induction of apoptosis through ceramide generation in fish fathead minnow tailbud cells and zebrafish embryo.* Biochem J, 2001. **360**(Pt 1): p. 39-47.

66.   Kowalska, A., *Genetic basis of neurodegeneration in familial Alzheimer's disease.* Pol J Pharmacol, 2004. **56**(2): p. 171-8.

67.   Citron, M., et al., *Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice.* Nat Med, 1997. **3**(1): p. 67-72.

68.   Borchelt, D.R., et al., *Familial Alzheimer's disease-linked presenilin 1 variants elevate Abeta1-42/1-40 ratio in vitro and in vivo.* Neuron, 1996. **17**(5): p. 1005-13.

69.   Sun, L., et al., *Analysis of 138 pathogenic mutations in presenilin-1 on the in vitro production of Abeta42 and Abeta40 peptides by gamma-secretase.* Proc Natl Acad Sci U S A, 2017. **114**(4): p. E476-E485.

70.   De Strooper, B., *Loss-of-function presenilin mutations in Alzheimer disease. Talking Point on the role of presenilin mutations in Alzheimer disease.* EMBO Rep, 2007. **8**(2): p. 141-6.

71.   Cacquevel, M., et al., *Alzheimer's disease-linked mutations in presenilin-1 result in a drastic loss of activity in purified gamma-secretase complexes.* PLoS One, 2012. **7**(4): p. e35133.

72.   Balasa, M., et al., *PSEN1 mutation carriers present lower cerebrospinal fluid amyoid-beta42 levels than sporadic early-onset Alzheimer's disease patients but no differences in neuronal injury biomarkers.* J Alzheimers Dis, 2012. **30**(3): p. 605-16.

73.   Liu, C.C., et al., *Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy.* Nat Rev Neurol, 2013. **9**(2): p. 106-18.

74.   Bird, T.D., *Genetic factors in Alzheimer's disease.* N Engl J Med, 2005. **352**(9): p. 862-4.

75.   Corder, E.H., et al., *Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families.* Science, 1993. **261**(5123): p. 921-3.

76.   Corder, E.H., et al., *Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease.* Nat Genet, 1994. **7**(2): p. 180-4.

77.   Strittmatter, W.J., et al., *Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease.* Proc Natl Acad Sci U S A, 1993. **90**(5): p. 1977-81.

78.   Hoe, H.S., et al., *Effects of apoE on neuronal signaling and APP processing in rodent brain.* Brain Res, 2006. **1112**(1): p. 70-9.

79.   He, X., et al., *Apolipoprotein receptor 2 and X11 alpha/beta mediate apolipoprotein E-induced endocytosis of amyloid-beta precursor protein and beta-secretase, leading to amyloid-beta production.* J Neurosci, 2007. **27**(15): p. 4052-60.

80.   Holtzman, D.M., J. Herz, and G. Bu, *Apolipoprotein E and apolipoprotein E receptors: normal biology and roles in Alzheimer disease.* Cold Spring Harb Perspect Med, 2012. **2**(3): p. a006312.

81. Kim, J., J.M. Basak, and D.M. Holtzman, *The role of apolipoprotein E in Alzheimer's disease.* Neuron, 2009. **63**(3): p. 287-303.

82. Castellano, J.M., et al., *Human apoE isoforms differentially regulate brain amyloid-beta peptide clearance.* Sci Transl Med, 2011. **3**(89): p. 89ra57.

83. Bagyinszky, E., et al., *The genetics of Alzheimer's disease.* Clin Interv Aging, 2014. **9**: p. 535-51.

84. Dourlen, P., et al., *The new genetic landscape of Alzheimer's disease: from amyloid cascade to genetically driven synaptic failure hypothesis?* Acta Neuropathol, 2019. **138**(2): p. 221-236.

85. Harold, D., et al., *Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease.* Nat Genet, 2009. **41**(10): p. 1088-93.

86. Calero, M., et al., *Apolipoprotein J (clusterin) and Alzheimer's disease.* Microsc Res Tech, 2000. **50**(4): p. 305-15.

87. Zlokovic, B.V., et al., *Brain uptake of circulating apolipoproteins J and E complexed to Alzheimer's amyloid beta.* Biochem Biophys Res Commun, 1994. **205**(2): p. 1431-7.

88. Xiao, Q., et al., *Role of phosphatidylinositol clathrin assembly lymphoid-myeloid leukemia (PICALM) in intracellular amyloid precursor protein (APP) processing and amyloid plaque pathogenesis.* J Biol Chem, 2012. **287**(25): p. 21279-89.

89. Zhao, Z., et al., *Central role for PICALM in amyloid-beta blood-brain barrier transcytosis and clearance.* Nat Neurosci, 2015. **18**(7): p. 978-87.

90. Bertram, L., et al., *Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database.* Nat Genet, 2007. **39**(1): p. 17-23.

91. Scherzer, C., et al., *Loss of Apolipoprotein E Receptor LR11 in Alzheimer Disease.* Archives of Neurology, 2004. **61**(8): p. 1200-5.

92. Lambert, J.C., et al., *Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease.* Nat Genet, 2013. **45**(12): p. 1452-8.

93. Miyashita, A., et al., *SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians.* PLoS One, 2013. **8**(4): p. e58618.

94. Pottier, C., et al., *High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease.* Mol Psychiatry, 2012. **17**(9): p. 875-9.

95. Rogaeva, E., et al., *The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease.* Nat Genet, 2007. **39**(2): p. 168-77.

96. Campion, D., C. Charbonnier, and G. Nicolas, *SORL1 genetic variants and Alzheimer disease risk: a literature review and meta-analysis of sequencing data.* Acta Neuropathol, 2019. **138**(2): p. 173-186.

97. Yin, R.H., J.T. Yu, and L. Tan, *The Role of SORL1 in Alzheimer's Disease.* Mol Neurobiol, 2015. **51**(3): p. 909-18.

98. Offe, K., et al., *The lipoprotein receptor LR11 regulates amyloid beta production and amyloid precursor protein traffic in endosomal compartments.* J Neurosci, 2006. **26**(5): p. 1596-603.

99. Dodson, S.E., et al., *Loss of LR11/SORLA enhances early pathology in a mouse model of amyloidosis: evidence for a proximal role in Alzheimer's disease.* J Neurosci, 2008. **28**(48): p. 12877-86.

100. Andersen, O.M., I.M. Rudolph, and T.E. Willnow, *Risk factor SORL1: from genetic association to functional validation in Alzheimer's disease.* Acta Neuropathol, 2016. **132**(5): p. 653-665.

101. Bohm, C., et al., *SorLA signaling by regulated intramembrane proteolysis.* J Biol Chem, 2006. **281**(21): p. 14547-53.

102. Hardy, J.A. and G.A. Higgins, *Alzheimer's disease: the amyloid cascade hypothesis.* Science, 1992. **256**(5054): p. 184-5.

103. Hardy, J. and D.J. Selkoe, *The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics.* Science, 2002. **297**(5580): p. 353-6.

104. Perl, D.P., *Neuropathology of Alzheimer's disease.* The Mount Sinai journal of medicine, New York, 2010. **77**(1): p. 32-42.

105. Shoji, M., et al., *Production of the Alzheimer amyloid beta protein by normal proteolytic processing.* Science, 1992. **258**(5079): p. 126-129.
106. Qi-Takahara, Y., et al., *Longer forms of amyloid beta protein: implications for the mechanism of intramembrane cleavage by gamma-secretase.* J Neurosci, 2005. **25**(2): p. 436-45.
107. Nicolas, M. and B.A. Hassan, *Amyloid precursor protein and neural development.* Development, 2014. **141**(13): p. 2543-8.
108. Drachman, D.A., *The amyloid hypothesis, time to move on: Amyloid is the downstream result, not cause, of Alzheimer's disease.* Alzheimers Dement, 2014. **10**(3): p. 372-80.
109. de la Torre, J.C., *Unproven Hypotheses on the Cause of Alzheimer's*, in *Alzheimer's Turning Point: A Vascular Approach to Clinical Prevention*. 2016, Springer International Publishing: Cham. p. 39-48.
110. Armstrong, R.A., *What causes alzheimer's disease?* Folia Neuropathol, 2013. **51**(3): p. 169-88.
111. de la Torre, J.C., *Other Hypotheses on the Cause of Alzheimer's Disease*, in *Alzheimer's Turning Point: A Vascular Approach to Clinical Prevention*. 2016, Springer International Publishing: Cham. p. 49-60.
112. Irizarry, M.C., et al., *APPSw transgenic mice develop age-related A beta deposits and neuropil abnormalities, but no neuronal loss in CA1.* J Neuropathol Exp Neurol, 1997. **56**(9): p. 965-73.
113. Shen, J. and R.J. Kelleher, 3rd, *The presenilin hypothesis of Alzheimer's disease: evidence for a loss-of-function pathogenic mechanism.* Proc Natl Acad Sci U S A, 2007. **104**(2): p. 403-9.
114. Song, W., et al., *Proteolytic release and nuclear translocation of Notch-1 are induced by presenilin-1 and impaired by pathogenic presenilin-1 mutations.* Proc Natl Acad Sci U S A, 1999. **96**(12): p. 6959-63.
115. Zhang, S., et al., *Biological function of Presenilin and its role in AD pathogenesis.* Transl Neurodegener, 2013. **2**(1): p. 15.
116. Kang, D.E., et al., *Presenilin 1 facilitates the constitutive turnover of beta-catenin: differential activity of Alzheimer's disease-linked PS1 mutants in the beta-catenin-signaling pathway.* J Neurosci, 1999. **19**(11): p. 4229-37.
117. Killick, R., et al., *Presenilin 1 independently regulates beta-catenin stability and transcriptional activity.* J Biol Chem, 2001. **276**(51): p. 48554-61.
118. de la Torre, J., *The Vascular Hypothesis of Alzheimer's Disease: A Key to Preclinical Prediction of Dementia Using Neuroimaging.* J Alzheimers Dis, 2018. **63**(1): p. 35-52.
119. Iyer, N.V., et al., *Cellular and developmental control of O2 homeostasis by hypoxia-inducible factor 1 alpha.* Genes Dev, 1998. **12**(2): p. 149-62.
120. Salminen, A., A. Kauppinen, and K. Kaarniranta, *Hypoxia/ischemia activate processing of Amyloid Precursor Protein: impact of vascular dysfunction in the pathogenesis of Alzheimer's disease.* J Neurochem, 2017. **140**(4): p. 536-549.
121. Daulatzai, M.A., *Death by a thousand cuts in Alzheimer's disease: hypoxia--the prodrome.* Neurotox Res, 2013. **24**(2): p. 216-43.
122. Patterson, D., et al., *Mapping of the gene encoding the beta-amyloid precursor protein and its relationship to the Down syndrome region of chromosome 21.* Proc Natl Acad Sci U S A, 1988. **85**(21): p. 8266-70.
123. Newton, J.R., D. Parkinson, and M.R. Clench, *Strategies for examination of Alzheimer's disease amyloid precursor protein isoforms.* Anal Bioanal Chem, 2006. **385**(4): p. 692-9.
124. Yoshikai, S., et al., *Genomic organization of the human amyloid beta-protein precursor gene.* Gene, 1990. **87**(2): p. 257-63.
125. Schilling, J., et al., *Synthesis and characterization of the Kunitz protease-inhibitor domain of the beta-amyloid precursor protein.* Gene, 1991. **98**(2): p. 225-30.
126. Abe, K., R.E. Tanzi, and K. Kogure, *Selective induction of Kunitz-type protease inhibitor domain-containing amyloid precursor protein mRNA after persistent focal ischemia in rat cerebral cortex.* Neurosci Lett, 1991. **125**(2): p. 172-4.

127. Nalivaeva, N.N. and A.J. Turner, *The amyloid precursor protein: a biochemical enigma in brain development, function and disease.* FEBS Lett, 2013. **587**(13): p. 2046-54.

128. Jacobsen, K.T. and K. Iverfeldt, *Amyloid precursor protein and its homologues: a family of proteolysis-dependent receptors.* Cell Mol Life Sci, 2009. **66**(14): p. 2299-318.

129. Selkoe, D.J., *Alzheimer's disease: genes, proteins, and therapy.* Physiol Rev, 2001. **81**(2): p. 741-66.

130. Soldano, A. and B.A. Hassan, *Beyond pathology: APP, brain development and Alzheimer's disease.* Curr Opin Neurobiol, 2014. **27**: p. 61-7.

131. Qiu, W.Q., et al., *Cell-surface beta-amyloid precursor protein stimulates neurite outgrowth of hippocampal neurons in an isoform-dependent manner.* J Neurosci, 1995. **15**(3 Pt 2): p. 2157-67.

132. Young-Pearse, T.L., et al., *A critical function for beta-amyloid precursor protein in neuronal migration revealed by in utero RNA interference.* J Neurosci, 2007. **27**(52): p. 14459-69.

133. Demars, M.P., et al., *Soluble amyloid precursor protein: a novel proliferation factor of adult progenitor cells of ectodermal and mesodermal origin.* Stem Cell Res Ther, 2011. **2**(4): p. 36.

134. Caille, I., et al., *Soluble form of amyloid precursor protein regulates proliferation of progenitors in the adult subventricular zone.* Development, 2004. **131**(9): p. 2173-81.

135. Demars, M.P., et al., *Soluble amyloid precursor protein-alpha rescues age-linked decline in neural progenitor cell proliferation.* Neurobiol Aging, 2013. **34**(10): p. 2431-40.

136. Hartl, D., et al., *Soluble alpha-APP (sAPPalpha) regulates CDK5 expression and activity in neurons.* PLoS One, 2013. **8**(6): p. e65920.

137. Corrigan, F., et al., *Characterisation of the effect of knockout of the amyloid precursor protein on outcome following mild traumatic brain injury.* Brain Res, 2012. **1451**: p. 87-99.

138. Corrigan, F., et al., *The neuroprotective activity of the amyloid precursor protein against traumatic brain injury is mediated via the heparin binding site in residues 96-110.* J Neurochem, 2014. **128**(1): p. 196-204.

139. von Rotz, R.C., et al., *The APP intracellular domain forms nuclear multiprotein complexes and regulates the transcription of its own precursor.* J Cell Sci, 2004. **117**(Pt 19): p. 4435-48.

140. Sumioka, A., et al., *Role of 14-3-3gamma in FE65-dependent gene transactivation mediated by the amyloid beta-protein precursor cytoplasmic fragment.* J Biol Chem, 2005. **280**(51): p. 42364-74.

141. Kimberly, W.T., et al., *The intracellular domain of the beta-amyloid precursor protein is stabilized by Fe65 and translocates to the nucleus in a notch-like manner.* J Biol Chem, 2001. **276**(43): p. 40288-92.

142. De Strooper, B. and W. Annaert, *Proteolytic processing and cell biological functions of the amyloid precursor protein.* J Cell Sci, 2000. **113 (Pt 11)**: p. 1857-70.

143. Mueller, H.T., et al., *Modulation of amyloid precursor protein metabolism by X11alpha /Mint-1. A deletion analysis of protein-protein interaction domains.* J Biol Chem, 2000. **275**(50): p. 39302-6.

144. Saito, Y., et al., *X11 proteins regulate the translocation of amyloid beta-protein precursor (APP) into detergent-resistant membrane and suppress the amyloidogenic cleavage of APP by beta-site-cleaving enzyme in brain.* J Biol Chem, 2008. **283**(51): p. 35763-71.

145. Sabo, S.L., et al., *The Alzheimer amyloid precursor protein (APP) and FE65, an APP-binding protein, regulate cell movement.* J Cell Biol, 2001. **153**(7): p. 1403-14.

146. McLoughlin, D.M. and C.C. Miller, *The FE65 proteins and Alzheimer's disease.* J Neurosci Res, 2008. **86**(4): p. 744-54.

147. Ma, H., et al., *Involvement of beta-site APP cleaving enzyme 1 (BACE1) in amyloid precursor protein-mediated enhancement of memory and activity-dependent synaptic plasticity.* Proc Natl Acad Sci U S A, 2007. **104**(19): p. 8167-72.

148. Han, W., et al., *Peptide p3 may play a neuroprotective role in the brain.* Med Hypotheses, 2011. **76**(4): p. 543-6.

149. Andrew, R.J., et al., *A Greek Tragedy: The Growing Complexity of Alzheimer Amyloid Precursor Protein Proteolysis.* J Biol Chem, 2016. **291**(37): p. 19235-44.

150.   Kumar-Singh, S., et al., *Nonfibrillar diffuse amyloid deposition due to a gamma(42)-secretase site mutation points to an essential role for N-truncated A beta(42) in Alzheimer's disease.* Hum Mol Genet, 2000. **9**(18): p. 2589-98.

151.   Crook, R., et al., *A variant of Alzheimer's disease with spastic paraparesis and unusual plaques due to deletion of exon 9 of presenilin 1.* Nat Med, 1998. **4**(4): p. 452-5.

152.   Zhou, L., et al., *Amyloid precursor protein mutation E682K at the alternative beta-secretase cleavage beta'-site increases Abeta generation.* EMBO Mol Med, 2011. **3**(5): p. 291-302.

153.   Hunter, J.C., et al., *In situ selectivity profiling and crystal structure of SML-8-73-1, an active site inhibitor of oncogenic K-Ras G12C.* Proc Natl Acad Sci U S A, 2014. **111**(24): p. 8895-900.

154.   Qahwash, I., et al., *Processing amyloid precursor protein at the beta-site requires proper orientation to be accessed by BACE1.* J Biol Chem, 2004. **279**(37): p. 39010-6.

155.   Creemers, J.W., et al., *Processing of beta-secretase by furin and other members of the proprotein convertase family.* J Biol Chem, 2001. **276**(6): p. 4211-7.

156.   Yang, H.C., et al., *Biochemical and kinetic characterization of BACE1: investigation into the putative species-specificity for beta- and beta'-cleavage sites by human and murine BACE1.* J Neurochem, 2004. **91**(6): p. 1249-59.

157.   Nilsberth, C., et al., *The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Abeta protofibril formation.* Nat Neurosci, 2001. **4**(9): p. 887-93.

158.   Lord, A., et al., *The Arctic Alzheimer mutation facilitates early intraneuronal Abeta aggregation and senile plaque formation in transgenic mice.* Neurobiol Aging, 2006. **27**(1): p. 67-77.

159.   Parvathy, S., et al., *Cleavage of Alzheimer's amyloid precursor protein by alpha-secretase occurs at the surface of neuronal cells.* Biochemistry, 1999. **38**(30): p. 9728-34.

160.   Lammich, S., et al., *Constitutive and regulated alpha-secretase cleavage of Alzheimer's amyloid precursor protein by a disintegrin metalloprotease.* Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3922-7.

161.   Sahlin, C., et al., *The Arctic Alzheimer mutation favors intracellular amyloid-beta production by making amyloid precursor protein less available to alpha-secretase.* J Neurochem, 2007. **101**(3): p. 854-62.

162.   Vassar, R., et al., *Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.* Science, 1999. **286**(5440): p. 735-41.

163.   Fraser, P.E., et al., *Presenilin structure, function and role in Alzheimer disease.* Biochim Biophys Acta, 2000. **1502**(1): p. 1-15.

164.   Wasco, W., J.D. Brook, and R.E. Tanzi, *The amyloid precursor-like protein (APLP) gene maps to the long arm of human chromosome 19.* Genomics, 1993. **15**(1): p. 237-9.

165.   Arduengo, P.M., et al., *The presenilin protein family member SPE-4 localizes to an ER/Golgi derived organelle and is required for proper cytoplasmic partitioning during Caenorhabditis elegans spermatogenesis.* J Cell Sci, 1998. **111 (Pt 24)**: p. 3645-54.

166.   Levitan, D., et al., *Assessment of normal and mutant human presenilin function in Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14940-4.

167.   Li, X. and I. Greenwald, *HOP-1, a Caenorhabditis elegans presenilin, appears to be functionally redundant with SEL-12 presenilin and to facilitate LIN-12 and GLP-1 signaling.* Proc Natl Acad Sci U S A, 1997. **94**(22): p. 12204-9.

168.   Link, C.D., *Expression of human beta-amyloid peptide in transgenic Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 1995. **92**(20): p. 9368-72.

169.   Fay, D.S., et al., *In vivo aggregation of beta-amyloid peptide variants.* J Neurochem, 1998. **71**(4): p. 1616-25.

170.   Drake, J., C.D. Link, and D.A. Butterfield, *Oxidative stress precedes fibrillar deposition of Alzheimer's disease amyloid beta-peptide (1-42) in a transgenic Caenorhabditis elegans model.* Neurobiol Aging, 2003. **24**(3): p. 415-20.

171.   Takasugi, N., et al., *The role of presenilin cofactors in the gamma-secretase complex.* Nature, 2003. **422**(6930): p. 438-41.

172. Fossgreen, A., et al., *Transgenic Drosophila expressing human amyloid precursor protein show gamma-secretase activity and a blistered-wing phenotype.* Proc Natl Acad Sci U S A, 1998. **95**(23): p. 13703-8.

173. Crowther, D.C., et al., *Intraneuronal Abeta, non-amyloid aggregates and neurodegeneration in a Drosophila model of Alzheimer's disease.* Neuroscience, 2005. **132**(1): p. 123-35.

174. Finelli, A., et al., *A model for studying Alzheimer's Abeta42-induced toxicity in Drosophila melanogaster.* Mol Cell Neurosci, 2004. **26**(3): p. 365-75.

175. Iijima, K., et al., *Dissecting the pathological effects of human Abeta40 and Abeta42 in Drosophila: a potential model for Alzheimer's disease.* Proc Natl Acad Sci U S A, 2004. **101**(17): p. 6623-8.

176. Seidner, G.A., et al., *Modeling clinically heterogeneous presenilin mutations with transgenic Drosophila.* Curr Biol, 2006. **16**(10): p. 1026-33.

177. Micchelli, C.A., et al., *Gamma-secretase/presenilin inhibitors for Alzheimer's disease phenocopy Notch mutations in Drosophila.* FASEB J, 2003. **17**(1): p. 79-81.

178. Newman, M., et al., *Chapter 40 - Animal Models of Alzheimer's Disease*, in *Animal Models for the Study of Human Disease (Second Edition)*, P.M. Conn, Editor. 2017, Academic Press. p. 1031-1085.

179. Hall, A.M. and E.D. Roberson, *Mouse models of Alzheimer's disease.* Brain Res Bull, 2012. **88**(1): p. 3-12.

180. Saito, T., et al., *Single App knock-in mouse models of Alzheimer's disease.* Nature Neuroscience, 2014. **17**(5): p. 661-663.

181. Hargis, K.E. and E.M. Blalock, *Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us?* Behav Brain Res, 2017. **322**(Pt B): p. 311-328.

182. Kaslin, J.a.G., Y., *Using Zebrafish to Study Human Genetic Disease.* eLS, John Wiley & Sons, Ltd (Ed.), 2017: p. 1-8.

183. Leimer, U., et al., *Zebrafish (Danio rerio) presenilin promotes aberrant amyloid beta-peptide production and requires a critical aspartate residue for its function in amyloidogenesis.* Biochemistry, 1999. **38**(41): p. 13602-9.

184. Groth, C., et al., *Identification of a second presenilin gene in zebrafish with similarity to the human Alzheimer's disease gene presenilin2.* Dev Genes Evol, 2002. **212**(10): p. 486-90.

185. Lee, J., S.M. Peterson, and J.L. Freeman, *Sex-specific characterization and evaluation of the Alzheimer's disease genetic risk factor sorl1 in zebrafish during aging and in the adult brain following a 100 ppb embryonic lead exposure.* J Appl Toxicol, 2017. **37**(4): p. 400-407.

186. Catchen, J.M., I. Braasch, and J.H. Postlethwait, *Conserved synteny and the zebrafish genome.* Methods Cell Biol, 2011. **104**: p. 259-85.

187. Musa, A., H. Lehrach, and V.A. Russo, *Distinct expression patterns of two zebrafish homologues of the human APP gene during embryonic development.* Dev Genes Evol, 2001. **211**(11): p. 563-7.

188. Strausberg, R.L., et al., *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16899-903.

189. Lim, A., et al., *Analysis of nicastrin gene phylogeny and expression in zebrafish.* Development Genes and Evolution, 2015. **225**(3): p. 171-178.

190. Moussavi Nik, S.H., et al., *The BACE1-PSEN-AbetaPP regulatory axis has an ancient role in response to low oxygen/oxidative stress.* J Alzheimers Dis, 2012. **28**(3): p. 515-30.

191. van Bebber, F., et al., *Loss of Bace2 in zebrafish affects melanocyte migration and is distinct from Bace1 knock out phenotypes.* Journal of Neurochemistry, 2013. **127**(4): p. 471-481.

192. Newman, M., E. Ebrahimie, and M. Lardelli, *Using the zebrafish model for Alzheimer's disease research.* Front Genet, 2014. **5**: p. 189.

193. Nornes, S., et al., *Independent and cooperative action of Psen2 with Psen1 in zebrafish embryos.* Exp Cell Res, 2009. **315**(16): p. 2791-801.

194. Campbell, W.A., et al., *Zebrafish lacking Alzheimer presenilin enhancer 2 (Pen-2) demonstrate excessive p53-dependent apoptosis and neuronal loss.* J Neurochem, 2006. **96**(5): p. 1423-40.

195.	Joshi, P., et al., *Amyloid precursor protein is required for convergent-extension movements during Zebrafish development.* Dev Biol, 2009. **335**(1): p. 1-11.

196.	Abramsson, A., et al., *The zebrafish amyloid precursor protein-b is required for motor neuron guidance and synapse formation.* Dev Biol, 2013. **381**(2): p. 377-88.

197.	Lum, T., G. Huynh, and G. Heinrich, *Brain-derived neurotrophic factor and TrkB tyrosine kinase receptor gene expression in zebrafish embryo and larva.* Int J Dev Neurosci, 2001. **19**(6): p. 569-87.

198.	Han, H.W., et al., *The Nogo-C2/Nogo receptor complex regulates the morphogenesis of zebrafish lateral line primordium through modulating the expression of dkk1b, a Wnt signal inhibitor.* PLoS One, 2014. **9**(1): p. e86345.

199.	Brosamle, C. and M.E. Halpern, *Nogo-Nogo receptor signalling in PNS axon outgrowth and pathfinding.* Molecular and Cellular Neuroscience, 2009. **40**(4): p. 401.

200.	Newman, M., et al., *A familial Alzheimer's disease-like mutation in the zebrafish presenilin 1 gene affects brain energy production.* bioRxiv, 2019: p. 542134.

201.	Hin, N., et al., *Accelerated brain aging towards transcriptional inversion in a zebrafish model of familial Alzheimer's disease.* bioRxiv, 2018: p. 262162.

202.	Jiang, H., M. Newman, and M. Lardelli, *The zebrafish orthologue of familial Alzheimer's disease gene PRESENILIN 2 is required for normal adult melanotic skin pigmentation.* bioRxiv, 2018: p. 414144.

203.	Draper, B.W., P.A. Morcos, and C.B. Kimmel, *Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholino oligos: a quantifiable method for gene knockdown.* Genesis, 2001. **30**(3): p. 154-6.

204.	Lawson, N.D. and S.A. Wolfe, *Forward and reverse genetic approaches for the analysis of vertebrate development in the zebrafish.* Dev Cell, 2011. **21**(1): p. 48-64.

205.	Kawakami, K., *Tol2: a versatile gene transfer vector in vertebrates.* Genome Biol, 2007. **8 Suppl 1**: p. S7.

206.	Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, *ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.* Trends Biotechnol, 2013. **31**(7): p. 397-405.

207.	Cade, L., et al., *Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs.* Nucleic Acids Res, 2012. **40**(16): p. 8001-10.

208.	Hwang, W.Y., R.T. Peterson, and J.R. Yeh, *Methods for targeted mutagenesis in zebrafish using TALENs.* Methods, 2014. **69**(1): p. 76-84.

209.	Huang, P., et al., *Reverse genetic approaches in zebrafish.* J Genet Genomics, 2012. **39**(9): p. 421-33.

210.	Reyon, D., et al., *FLASH assembly of TALENs for high-throughput genome editing.* Nat Biotechnol, 2012. **30**(5): p. 460-5.

211.	Bedell, V.M., et al., *In vivo genome editing using a high-efficiency TALEN system.* Nature, 2012. **491**(7422): p. 114-8.

212.	Wiedenheft, B., S.H. Sternberg, and J.A. Doudna, *RNA-guided genetic silencing systems in bacteria and archaea.* Nature, 2012. **482**(7385): p. 331-8.

213.	Pickar-Oliver, A. and C.A. Gersbach, *The next generation of CRISPR-Cas technologies and applications.* Nat Rev Mol Cell Biol, 2019. **20**(8): p. 490-507.

214.	Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-308.

215.	Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.

216.	Shah, S.A., et al., *Protospacer recognition motifs: mixed identities and functional diversity.* RNA Biol, 2013. **10**(5): p. 891-9.

217.	Hwang, W.Y., et al., *Efficient genome editing in zebrafish using a CRISPR-Cas system.* Nat Biotechnol, 2013. **31**(3): p. 227-9.

218.	Chang, N., et al., *Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos.* Cell Res, 2013. **23**(4): p. 465-72.

219.   Hwang, W.Y., et al., *Heritable and precise zebrafish genome editing using a CRISPR-Cas system.* PLoS One, 2013. **8**(7): p. e68708.

220.   Fu, Y., et al., *High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells.* Nat Biotechnol, 2013. **31**(9): p. 822-6.

221.   Shen, B., et al., *Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects.* Nat Methods, 2014. **11**(4): p. 399-402.

222.   Wang, H., et al., *One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering.* Cell, 2013. **153**(4): p. 910-8.

223.   Jao, L.E., S.R. Wente, and W. Chen, *Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system.* Proc Natl Acad Sci U S A, 2013. **110**(34): p. 13904-9.

224.   Yang, H., et al., *One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering.* Cell, 2013. **154**(6): p. 1370-9.

225.   Hruscha, A., et al., *Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish.* Development, 2013. **140**(24): p. 4982-7.

226.   Yin, H., et al., *Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype.* Nat Biotechnol, 2014.

Chapter 2 – Developing *in vivo* assays for investigating p75^{NTR} and NRH1 transmembrane cleavage events using zebrafish embryos

2.0 Declaration

Within this chapter a portion of the results described in 2.2, namely, 1) the tblastn searches described in paragraph 1, below subheading *Identification of p75^{NTR} and NRH1gene orthologues in zebrafish*, 2) the mCherry construct designs described in paragraphs 2, 3 and 5, below subheading *Design of p75NTR and Nrh1 γ-secretase cleavage assay constructs*, 3) microinjections through to western immunoblots of mCherry constructs described in paragraph 6, below subheading *Design of p75NTR and Nrh1 γ-secretase cleavage assay constructs*, and 4) DAPT treatments and western immunoblots described in paragraph 1, below the subheading *Investigating differential cleavage properties of zebrafish p75^{NTF} and Nrh1* and presented in Figure 2.3, are from work that was previously completed and submitted by myself within a thesis for the Honours Degree of Bachelor of Science at the University of Adelaide.

As this chapter has been written with the intention to submit to a journal for publication, it was necessary to include the results and analysis of this previously described work to provide a complete narrative.

2.1 Introduction

γ-secretase is a multi-subunit membrane-bound aspartyl protease complex responsible for cleavage of over 100 substrates including Amyloid Precursor Protein (APP), Notch and the p75 neurotrophin receptor (p75^{NTR}) [1]. γ-secretase has been identified as a member of the intramembrane cleaving protease family (I-CLiP). I-CLiPs cleave type 1 membrane proteins

enzymatically via a process termed regulated intramembrane proteolysis (RIP) [2, 3]. The most studied function of γ-secretase is processing of APP. This is due to the seemingly critical role of APP in Alzheimer's disease (AD) etiology. Although numerous publications have discussed the perceived role of γ-secretase in AD, the specific nature of substrate selection by this protease is still not clearly defined.

γ-secretase substrates are typically derived from large precursor proteins that undergo a prerequisite removal of their ectodomain/lumenal domain prior to γ-secretase cleavage [3] Early research suggested that the only other prerequisite was that the substrate must be a type 1 transmembrane protein [4]. Further studies have now shown that additional factors may guide substrate selection by γ-secretase. It has previously been suggested that dimerisation of substrates and/or the structure of substrate α-helices may regulate γ-secretase activity [5]. γ-secretase substrate recognition and cleavage is also much more efficient for ectodomains with fewer than ~50 remaining amino acid residues [6]. Previous studies have attempted to elucidate Notch and APP residues that are required for γ-secretase cleavage [7]. However, a distinct cleavage site for γ-secretase within the transmembrane domains of its target proteins has not been defined [2].

p75^{NTR}, also known as the 'low-affinity nerve growth factor receptor' (LNGFR), is one of the many substrates of γ-secretase subject to cleavage within its transmembrane domain [8]. p75^{NTR} has been implicated in neuronal survival, myelination and neurite outgrowth among other pathways during vertebrate nervous system development, through its interactions with neurotrophins and Trk receptors [9]. Also, the Aβ peptide can act as a ligand for p75^{NTR} and is proposed to play a role in cholinergic neuron loss, implicating this protein in AD [10, 11]. Kanning et al (2003) investigated the proteolytic processing of p75^{NTR}, along with what they described as

the "neurotrophin receptor homologs" (NRH), NRH1 and NRH2 [8]. Database and EST searches have established that these two genes show greater sequence similarity to p75^{NTR} than any other homologous genes [8]. Experiments by Kanning *et al* (2003) confirmed processing of p75^{NTR} by α-secretase (ADAM10) and γ-secretase. Western blot analysis of NRH1 and NRH2 indicated that these are both cleaved within their transmembrane domains [8]. However, Kanning *et al* (2003) observed that the commonly used inhibitor of γ-secretase activity, DAPT, had no effect on the cleavage of NRH1 or NRH2, suggesting that transmembrane cleavage of these proteins is not by γ-secretase [8].

The possible lack of sensitivity to γ-secretase inhibitors of the p75^{NTR} homologs NRH1 and NRH2 may have interesting applications. p75^{NTR} and its homologs share high sequence similarity in their transmembrane domain, where γ-secretase cleavage occurs [8] and comparison of the transmembrane domains of p75^{NTR} and NRH1 might allow definition of transmembrane domain characteristics critical to permit cleavage by γ-secretase. We have developed an *in vivo* zebrafish assay that can be used to investigate the structural differences in the transmembrane domains of these two proteins that cause their differential sensitivity to γ-secretase.

2.2 Results and Discussion

Identification of p75^{NTR} and NRH1gene orthologues in zebrafish

Of the two *NRH* genes, *NRH1* has coding sequences more similar in length to *p75^{NTR}* and is the only know *NRH* gene with a putative orthologue in zebrafish. To investigate p75^{NTR} and NRH1 cleavage *in vivo* using zebrafish, we previously validated the orthology of these genes [12]. A tblastn search performed against the zebrafish genome using the entire putative protein sequence

of human $p75^{NTR}$, constrained to "RefSeq_RNA", returned candidate orthologues of the human $p75^{NTR}$ gene on zebrafish chromosomes 3, 12 and 16. At the time of the original analysis chromosome 12 appeared to hold two almost identical copies of the gene at different loci (supplementary file S1, Figure S1) which we suspected was due to a recent duplication event. The position of the duplicate appears to have been revised to chromosome 3 in the latest genome build (GRCz11) (supplementary file S1, Figure S2). The top tblastn hit, *nerve growth factor receptor b* (*ngfrb*), on chromosome 12, has the greatest query cover (percentage of the sequence aligned to a sequence in GenBank) to human $p75^{NTR}$ (93%) (supplementary file S1, Figure S2), so we tentatively named this "zebrafish p75". A tblastn search performed against the zebrafish genome using *NRH1* from *Xenopus laevis* (GenBank accession AF131890.1) returned the computer predicted sequence for *neurotrophin receptor associated death domain* (*nradd*) on chromosome 16 (also returned as a best hit in the human $p75^{NTR}$ tblastn search described above) with 100% query coverage to Xenopus *NRH1*. The only other strong zebrafish *Nrh1* candidate returned was *ngfrb*, which we had already established most likely represents $p75^{NTR}$ in zebrafish. Therefore, we predict that *nradd* is most likely a *Nrh1* orthologous gene in zebrafish.

NRH1 belongs to a subfamily of vertebrate p75^NTR-related proteins which also contains NRH2. NRH2 exists only in mammals while NRH1 exists only in amphibians, fish and birds [8]. The return of the predicted sequence *nradd* when searching for *NRH1* within the zebrafish genome is consistent with previous knowledge that *NRH2* is also known as *NRADD* in mouse and rat (mouse NCBI Gene ID: 67169, rat NCBI Gene ID: 246143). Therefore, we eliminated *nradd* as a zebrafish p75^NTR candidate gene and propose that it is the *Nrh1* orthologous gene in zebrafish.

To confirm our identification of zebrafish $p75^{NTR}$ and $Nrh1$ orthologues we next conducted phylogenetic analyses using the Geneious software suite [13]. Zebrafish p75^{NTR} and Nrh1 candidate amino acid sequences along with p75^{NTR} amino acid sequences from *Xenopus laevis*, *Gallus gallus* (chicken), mouse and human and NRH1 amino acid sequences from *Xenopus tropicalis*, *Xenopus laevis* and chicken were aligned using the Geneious alignment tool, and trees were built using both Bayesian and Maximum likelihood methods. We included the amino acid sequences of NRH2 from mouse and human in these analyses as these are the mammalian equivalents of NRH1 [8]. Accession numbers for all sequences used can be found in supplementary data File S2, Table S1. *Branchiostoma floridae* (lancelet) was used as an out-group as this was the most distant relative to zebrafish that returned a result when conducting tblastn searches using human $p75^{NTR}$ and *X. laevis NRH1*. Interestingly, tblastn searches of the lancelet genome using both $p75^{NTR}$ and *NRH1* returned the same gene in lancelet (supplementary data File S2, Table S1). As the chicken genome contains both $p75^{NTR}$ and *NRH1*-like sequences, tblastn searches of the lancelet genome using both full-length chicken sequences were performed to confirm the preliminary findings. These searches returned results identical to those using human $p75^{NTR}$ and *X. laevis NRH1*. This supports that there is only a single $p75^{NTR}$- and *NRH1*-like gene in this basal chordate and that $p75^{NTR}$ and *NRH1* arose from a gene duplication event early in vertebrate evolution. A dendrogram modelling the phylogenetic relationships of p75^{NTR} and its homologs demonstrated p75^{NTR} proteins clustering together and NRH1 proteins clustering together in separate clades (Figure 2.1). This supports that the sequence *nradd* on chromosome 16 of zebrafish is indeed the orthologue of *Nrh1*.
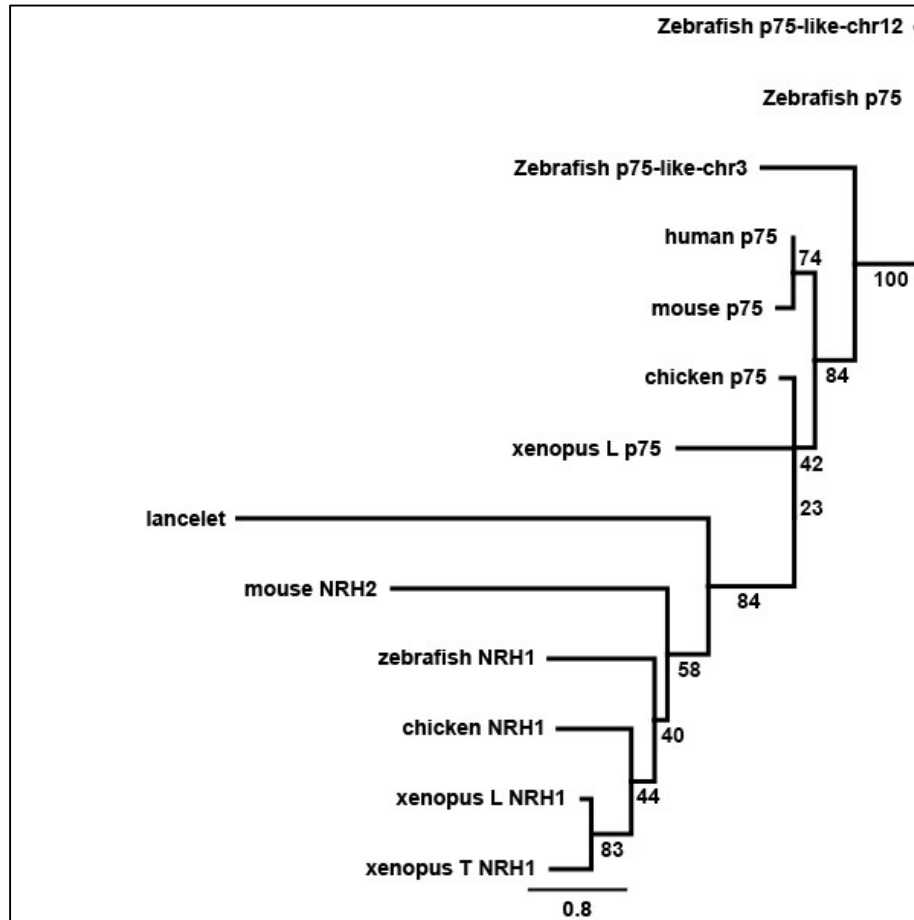
Figure 2.1. Phylogenetic tree of p75^NTR and NRH1 proteins. *This tree was generated using the Geneious software suite to perform Bayesian analysis (MrBayes). Numbers represent the aLRT branch-support values.*

It was difficult to discern which of the remaining three possible zebrafish p75$^{\text{NTR}}$ protein sequences was most likely to represent the true p75$^{\text{NTR}}$ in zebrafish from the phylogenetic analyses, although the sequence that we designated zebrafish p75-like-chr3 appeared to be marginally more similar to human p75$^{\text{NTR}}$ (Figure 2.1). The sequence we designated zebrafish p75, known as *ngfrb* in the NCBI database, is the only gene with a coding sequence not derived by computer-prediction from

genomic DNA sequence. It was thereby judged to be the most likely to be accurate and was selected for construction of an assay.

Design of p75^NTR and Nrh1 γ-secretase cleavage assay constructs

In order to be recognised as a cleavage substrate by γ-secretase, proteins must first undergo truncation of the extracellular domain [6]. This preliminary shedding of large extracellular domains by α- or β-secretase may be rate-limiting, creating problems for an *in vivo* assay as the rate of cleavage may be dependent on these events rather than γ-secretase itself. To overcome this in our assay, we proposed to produce artificially truncated forms of both proteins similar to a strategy we had adopted previously for zebrafish APP in another γ-secretase assay [14].

To truncate p75^{NTR} and Nrh1 to mimic approximately α-secretase-cleaved forms of these proteins, we first needed to identify the transmembrane domains of both p75^{NTR} and Nrh1. As the transmembrane domains of zebrafish p75^{NTR} and Nrh1 genes had not yet been defined, we had previously employed online prediction programs "TMHMM server, v. 2.0" (http://www.cbs.dtu.dk/services/TMHMM/) [15], "TMpred Server" (https://embnet.vital-it.ch/software/TMPRED_form.html) and "DAS" (https://tmdas.bioinfo.se/DAS/index.html, [15]) to determine the location of the transmembrane region within each protein (supplementary data File S3, Figure S3) [12]. The success of this approach was confirmed by comparing our predicted transmembrane domains to the defined transmembrane domain of human p75^{NTR} [5]. It has previously been established that γ-secretase cleavage of p75^{NTR} is dependent on shedding of the N-terminal extracellular domain by the α-secretase, A DISINTEGRIN AND METALLOPROTEINASE DOMAIN 17 (ADAM17). A previous study of p75^{NTR} suggested that

this cleavage occurs five amino acids before the N-terminal of the transmembrane [5]. Another study found that a large deletion of 15 amino acid residues within the juxtamembrane domain (before the N-terminal of the transmembrane) of p75^{NTR} led to substantial decrease in shedding, while a third study found that a stub of 15 amino acids N-terminal to the transmembrane domain was sufficient for γ-secretase cleavage [17, 18]. We elected to remove most of the extracellular domain, leaving only 15 amino acids immediately adjacent the N-terminal end of the membrane. The C-terminal intracellular domain was left at its original length as it has not been found to affect cleavage [17]. When attempting to visualise cleaved and uncleaved protein fragments on a western blot this small 15 residue N-terminal stub may have posed a problem. Post γ-secretase cleavage we would be left with a very short N-terminal fragment and a much larger C-terminal fragment, which would presumably be difficult to resolve on the same western immunoblot.

To overcome this, previous work in our laboratory fused 3 FLAG tags [19] in tandem to the N-terminals of both p75^{NTR} and Nrh1 [12]. The efficiency of processing by γ-secretase is not reduced unless the number of amino acids in the extracellular domain exceeds 50, so processing and cleavage should not be affected by the FLAG tags [6]. Single, silent, point mutations were introduced into both the second and third tandem FLAG tag repeats to inhibit recombination in bacteria during cloning [20]. The fourth codon in the second repeat was altered from GAT to GAC, and the sixth codon in the third from GAC to GAT respectively. The highly active HMM+38 secretory signal sequence was added N-terminal to the FLAG repeats to ensure insertion of the p75^{NTR} and Nrh1 proteins into lipid bilayers [21]. This signal sequence is cleaved off upon reaching the target site and is not involved in the metabolism of the final translated protein, hence does not alter the mature protein structure [21]. Destabilised green fluorescent protein (dGFP) was included

at the C-terminus of p75^{NTR} and Nrh1 to allow for visualisation of expression *in vivo* and via western immunoblot analysis. As these constructs contain a truncated, C-terminal fragment of the original p75^{NTR} and Nrh1 proteins we describe them as ssFLAG-p75^{NTR}C201-dGFP and ssFLAG-Nrh1C191-dGFP respectively.

Injection of transposon-based vector DNA into fertilised zebrafish eggs can give variable results in terms of the amount of DNA delivered and the subsequent degree of transposition and transgene expression. Therefore, when examining the stability of a protein expressed from an injected transgene, it can be useful to have an internal reference standard against which the quantity of the protein can be compared. In a previous study assaying γ-secretase activity by monitoring cleavage of zebrafish Appa fused to dGFP (also expressed from the Tol2 transgene vector, pT2AL200R150G), we co-injected a similar Tol2 vector expressing free GFP, as an internal reference standard [14]. However, this strategy might still be somewhat subject to variability as the free GFP vector might not transpose into the genome at a constant rate relative to the Appa construct.

The viral 2A (v2A) peptide ribosomal-skip mechanism allows for expression of two different proteins independently of one another. The skip mechanism occurs within the v2A sequence when a peptide bond fails to form between the penultimate (glycine) and final (proline) residue. Translation continues despite this failure, and tandem protein products are produced in a stoichiometric manner [22]. In order to express truncated p75NTR or Nrh1 simultaneously with an internal reference standard from the same expression vector, we previously included a v2a sequence at the C-terminal of ssFLAG-p75^{NTR}C201-dGFP and ssFLAG-Nrh1C201-dGFP

followed by coding sequence for the red fluorescence protein, mCherry [12]. We describe these Tol2-based expression constructs as pT2ALssFLAG-p75^{NTF}C201-dGFP-v2a-mCherry and pT2ALssFLAG-Nrh1C191-dGFP-v2a-mCherry (Figure 2.2, 1). For simplicity, we will henceforth refer to them as p75^{NTF}C201-dGFP and Nrh1C191-dGFP respectively. This design enables stoichiometric production of p75^{NTF}C201-dGFP or Nrh1C191-dGFP simultaneously with mCherry, allowing for normalisation of protein expression between successive batches of injected embryos.

Previous work in our laboratory investigating γ-secretase cleavage of both p75^{NTF}C201-dGFP and NRH1C191-dGFP, saw each expression vector co-injected with transposase mRNA into single-cell stage zebrafish embryos [12]. At 24 hours post fertilisation (hpf) embryos displaying GFP fluorescence under UV light were selected, their yolks removed, and protein extracted by lysis in SDS buffer. Protein lysates were separated by SDS polyacrylamide gel electrophoresis and then subjected to western immunoblotting to detect GFP. Uncleaved p75^{NTF}C201-dGFP and Nrh1C191-dGFP proteins were visible as bands of ~61kDa and ~57kDa respectively and could be assessed by densitometry, while their intra-membrane domain cleavage products could not be observed (presumably due to their instability, data not shown). Stripping and probing of the blot to detect mCherry followed by densitometry allowed normalisation of the GFP signals to facilitate comparison of p75NTFC201-dGFP and Nrh1C191-dGFP stability between samples.
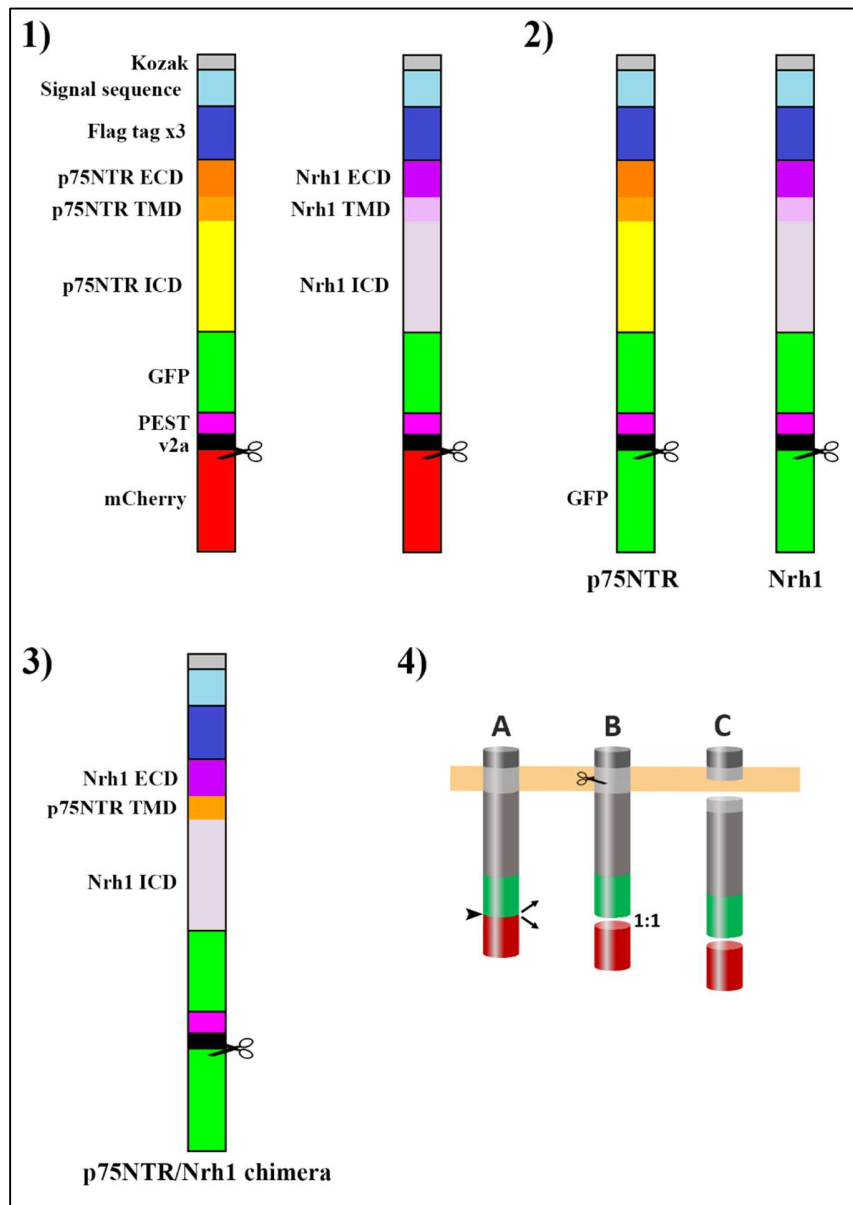
Figure 2.2 Overview of all construct designs and assay process. 1) Overview of pT2ALssFLAG-p75^{NTF}C201-dGFP-v2a-mCherry and pT2ALssFLAG-Nrh1C191-dGFP-v2a-mCherry constructs. 2) pT2ALssFLAG-p75^{NTF}C201-dGFP-v2a-GFP and pT2ALssFLAG-Nrh1C191-dGFP-v2a-GFP. 3) Overview of pT2ALssFLAG-A2C-dGFP-v2a-GFP. 4) Anticipated assay process. A) the position and mechanism of the ribosomal skip is indicated by arrows, B) TMD cleavage of constructs is represented by scissors. In all figures ECD = extracellular domain, TMD = transmembrane domain and ICD = intracellular domain.

Investigating differential cleavage properties of zebrafish p75^{NTF} and Nrh1

Previous experiments performed by Kanning *et al* (2003) indicated that both the p75^{NTR} and NRH1 proteins are cleaved within their transmembrane domains. However, treatment with the known γ-secretase inhibitor DAPT had no effect on NRH1 cleavage, suggesting that NRH1 is not processed by γ-secretase [8]. To test whether p75^{NTF}C201-dGFP and Nrh1C191-dGFP cleavage displayed differential sensitivity to inhibition of γ-secretase activity *in vivo*, batches of injected embryos were divided into two groups. Half of a batch injected with either p75^{NTF}C201-dGFP or Nrh1C191-dGFP was treated with 100µM of the γ-secretase inhibitor DAPT from 4hpf until 24hpf (as was previously optimised [14]), while the other half was left untreated. DAPT is a potent γ-secretase cleavage inhibitor, hence we expected that treatment with DAPT would result in an accumulation of uncleaved substrates of γ-secretase [18]. This was observed on p75^{NTF}C201-dGFP-injected embryos treated with DAPT (Figure 2.3, A), suggesting that zebrafish p75^{NTR} is, indeed, processed by γ-secretase. However, when this was normalised to mCherry across three replicates, the p-value (0.1697) of this observed increase did not support the likelihood that uncleaved zebrafish p75^{NTR} is consistently accumulated when treated with DAPT (Figure 2.3, B). This accumulation of substrate was not observed when Nrh1C191-dGFP was subjected to DAPT treatment (Figure 2.3, A), suggesting that zebrafish Nrh1 is not sensitive to the γ-secretase inhibitor DAPT. This result supports the observation made by Kanning *et al* (2003) that NRH1 is not a substrate of γ-secretase [8].

As was previously observed in a similar γ-secretase assay, both p75^{NTF}C201-dGFP and Nrh1C191-dGFP appeared to induce developmental abnormalities in embryos [14]. When injected, embryos displayed increased mortality (~46% vs. ~5-10% in uninjected controls) and a range of

developmental abnormalities, which were increased by DAPT treatment (~57% mortality) (data not shown). This toxicity may be due to excessive generation of p75^{NTF} and Nrh1 intracellular domains (ICDs), both of which carry a death domain [5, 8]. This increase in ICD carrying death domains may result in premature embryo death. p75^{NTF} is also known to form a heterodimer with the protein SORTILIN which then interacts with proNGF or proBDNF proteins leading to apoptosis [23, 24]. Treatment of embryos with DAPT may inadvertently facilitate this interaction by causing accumulation of uncleaved p75^{NTR}. Another factor to consider is the nature of the vectors used to express the assay constructs. Tol2 is a transposase vector that inserts randomly into the genome. If the Tol2-based constructs were to disrupt essential/housekeeping genes this would also affect the survival of embryos [25].

Zebrafish caspase 3 plays an important role in apoptosis signalling [26]. To overcome the toxicity of assay construct expression, we treated zebrafish embryos with 100μM of the caspase 3 inhibitor Z-DEVD-FMK (as was shown in a PhD dissertation [27]). However, treatment with this inhibitor did not reduce the degree of lethality observed (data not shown). Therefore, to overcome the problem of the lethality, an increased number of embryos were injected in each batch with only the most phenotypically normal embryos being selected for western blot analysis (after confirmation that they were expressing observable GFP).
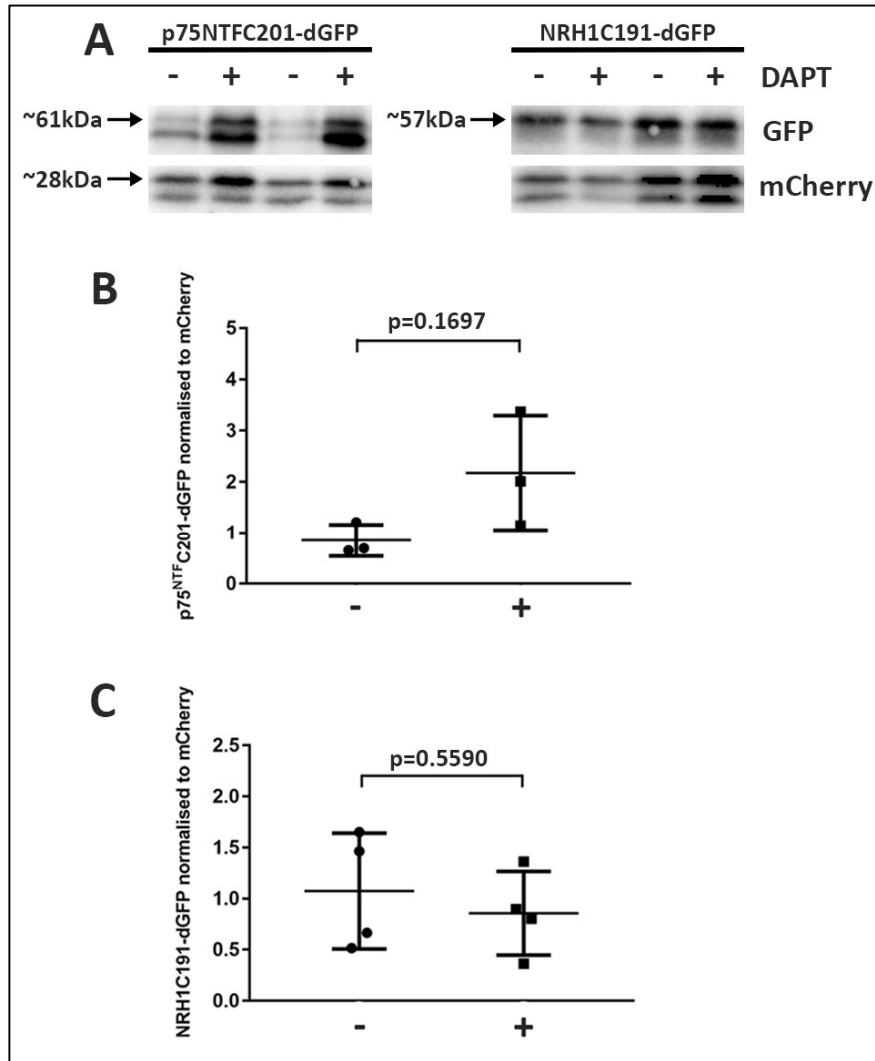
Figure 2.3. Western immunoblot analysis of p75^{NTF}C201-dGFP and Nrh1C191-dGFP. *A) Western immunoblots from p75^{NTF}C201-dGFP and Nrh1C191-dGFP injected embryos at 24hpf with and without DAPT treatment. + indicates embryos were treated with DAPT. The additional band observed below the bands indicated by arrows are most likely degradation products. B) Ratios of p75^{NTF}C201-dGFP/free mCherry in p75^{NTF}C201-dGFP injected embryos at 24hpf, with (+) (n = 3) or without (-) (n = 3) DAPT treatment. C) Ratios of Nrh1C191-dGFP /free mCherry in Nrh1C191-dGFP injected embryos at 24hpf, with (+) (n = 4) or without (-) (n = 4) DAPT treatment.*

Replacing mCherry with GFP reduces variability in the western immunoblot analyses

An apparent trend of accumulation of p75^NTFC201-dGFP due to γ-secretase inhibition was observed by western immunoblotting. However, statistical analysis of the densitometry measurements did not indicate significance due to the considerable variability between samples (Figure 2.3, B and C). A contributor to this variability may have been the necessity to strip and re-probe the western blot with the anti-mCherry antibody. To overcome this, the red fluorescence gene mCherry was excised from the constructs and replaced with a second GFP gene downstream of the C-terminal of v2a, producing vectors pT2ALssFLAG-p75^NTFC201-dGFP-v2a-GFP and pT2ALssFLAG-Nrh1C191-dGFP-v2a-GFP (Figure 2.2, 2). For simplicity, we will henceforth refer to them as p75^NTFC201-dGFPx2 and Nrh1C191-dGFPx2 respectively. This minor adjustment in construct design allows for the internal expression standard to be visualised using the same anti-GFP antibody as detects the p75^NTFC201-dGFP and Nrh1C191-dGFP fusions.

To evaluate the effectiveness of the modified assay constructs we performed injections on numerous batches of embryos and then ran protein samples on multiple western blots. Analysis using the new assay constructs consistently displayed an increase in accumulation of the p75^NTFC201-dGFPx2 substrate when treated with the γ-secretase inhibitor DAPT. This result was confirmed statistically by combining band intensity data from across all western blots and performing a two-tailed t-test assuming unequal variances, resulting in a p value of 0.0047 (figure 2.4, A).
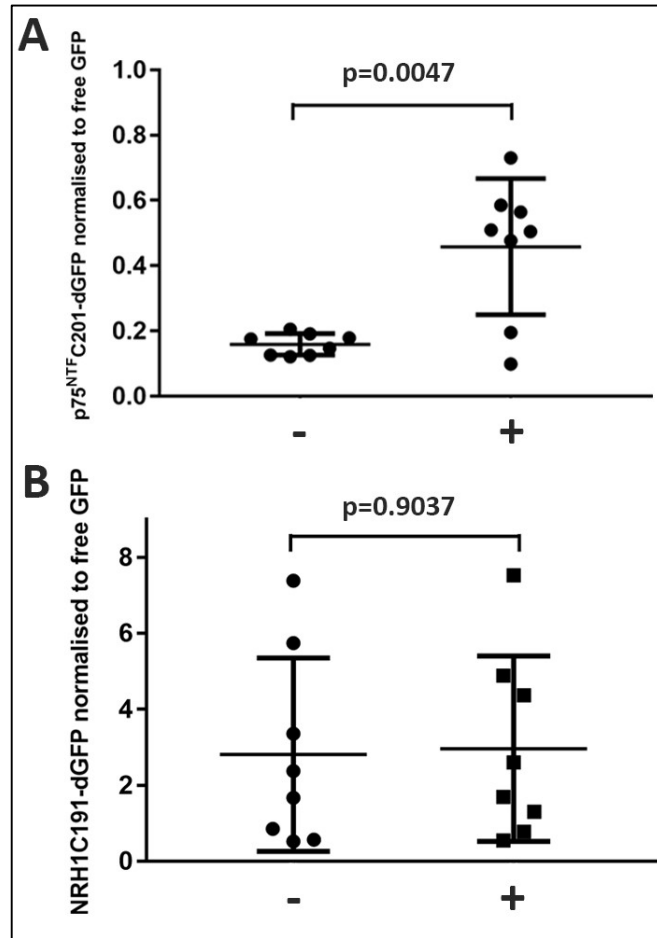
Figure 2.4. Western immunoblot analysis of p75^{NTF}C201-dGFPx2 and Nrh1C191-dGFPx2. *Ratios from western immunoblots of p75^{NTF}C201-dGFPx2/free GFP in p75^{NTF}C201-dGFPx2 injected embryos at 24hpf, with (+) (n = 8) or without (-) (n = 8) DAPT treatment. C) Ratios of Nrh1C191-dGFPx2 /free GFP in Nrh1C191-dGFPx2 injected embryos at 24hpf, with (+) (n = 8) or without (-) (n = 8) DAPT treatment. p value was calculated using an unpaired, two-tailed t-test.*

Conversely, there was no observed increase in Nrh1C191-dGFP substrate accumulation in response to DAPT treatment. Statistical analysis found no significant difference between the treated and untreated samples (p =0.9037) (Figure 2.4, B). Although the p75^{NTR} western immunoblot data was similar across numerous blots there was a high degree in variability in the

normalised values across Nrh1 immunoblots. This variability was unexpected and seems to be due to variation in the amount of free GFP on each blot. It is possible that GFP stability may play a role in this. Regardless of this we are confident that Nrh1 is not responsive to DAPT, as when looking at each blot individually there was consistently no accumulation in response to this inhibitor among replicates (data not shown).

γ-secretase cleavage is not conferred to Nrh1 by the p75^{NTF} transmembrane domain

The sequence similarity of p75^{NTR} and its homolog Nrh1 imply that these two genes share a relatively recent evolutionary origin through duplication of an ancestral p75^{NTR}/Nrh1-like gene. Therefore, if p75^{NTR} is a substrate of γ-secretase, then Nrh1 would most probably also be cleaved by it. The observed lack of γ-secretase-dependent cleavage of zebrafish Nrh1 in our assay is consistent with the results of Kanning *et al* (2003) [8]. The existence of this pair of closely related genes/proteins, one of which is cleaved by γ-secretase and one of which is not, presents us with a unique opportunity to dissect the structural basis of γ-secretase cleavage substrate specificity.

To begin dissection of γ-secretase cleavage substrate specificity using our assay, we designed a chimaeric construct in which Nrh1's transmembrane domain was replaced with the transmembrane domain from p75^{NTR}. The new construct, termed pT2ALssFLAG-A2C-dGFP-v2a-GFP (Figure 2.2 3), simplified to A2C-dGFPx2, was injected into one cell stage embryos which were subsequently treated as previously with or without DAPT. Protein samples were then collected at 24 hpf for analysis by western immunoblot. This did not reveal an accumulation of substrate when γ-secretase was inhibited (Figure 2.5). This suggests that the p75^{NTR} transmembrane domain alone is not sufficient to confer γ-secretase susceptibility and that structures outside of this domain are

also required. It is possible that altering the protein by swapping entire domains may disrupt its ability to form homo- or heterodimers, which have previously been found to be important for γ-secretase cleavage [5]. A detailed further analysis replacing Nrh1 amino acid residues with those not shared by p75^{NTR} in this chimaeric construct should allow definition of the structures critical to conferring γ-secretase susceptibility.
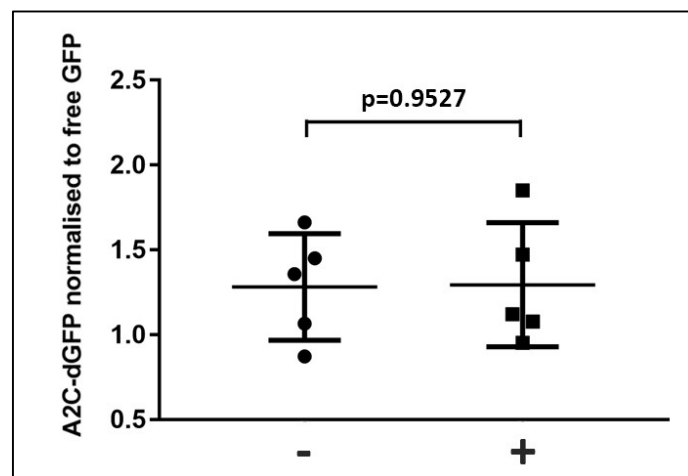
Figure 2.5. Western immunoblot analysis of A2C-dGFPx2. *Ratios from western immunoblots of A2C-dGFPx2 /free GFP in A2C-dGFPx2 injected embryos at 24hpf, with (+) (n = 5) or without (-) (n = 5) DAPT treatment.*

Further analysis of p75^{NTR} and Nrh1 transmembrane domains

Recent studies have shown that the α-helices of γ-secretase substrates Notch and APP unwind when they interact with the active site of PRESENILIN (the catalytic core of γ-secretase) [28-30]. A study investigating the conformation of the rhomboid substrate Gurken during cleavage by the archaeal homologue of PRESENILIN, MCMJR1, found that Gurken underwent a conformational change into a β-strand when interacting with MCMJR1 [31]. Proline residues have previously been found to disrupt transmembrane helices. Therefore, Brown *et al* (2018) altered the Pro252 TMD

residue of Gurken and found that it was no longer a substrate of MCMJR1. They suggested that the proline residue, through its perturbation of the α-helical conformation, effectively allows for this helix to unwind into the β-strand conformation that is preferred by MCMJR1 [31]. They also observed that so called "noncleavable" variants (i.e. those that cannot enter a β-strand conformation) could still bind to MCMJR1 with equal affinity to the cleavable substrates.

We aligned the TMDs of zebrafish p75^{NTR} and Nrh1 to investigate whether either of these sequences contain a proline residue that would allow them to unwind from an α-helix to a β-strand, as was observed for Gurken. Interestingly, while the TMD of zebrafish p75^{NTR} contains a proline close to the N-terminal end, zebrafish Nrh1 does not contain a proline within its TMD (Figure 2.6).

```
p75- 241 NLIPIYTSILAAVLLGLVAFIIF 263
Nrh1- 241 NNILVYVSVLAAVVLGLLLYVAY 263
```

Figure 2.6. Alignment of the zebrafish p75^{NTR} and Nrh1 transmembrane domains. *Red box indicates the proline residue in p75^{NTR} that is absent in Nrh1.*

2.3 Conclusions and Future directions

In this study we identified the *p75^{NTR}* and *Nrh1* orthologues in zebrafish suitable for design of an assay system in which to test γ-secretase cleavage of these proteins *in vivo*. We observed that, while cleavage of zebrafish p75^{NTR} by γ-secretase is sensitive to DAPT, zebrafish Nrh1 is not sensitive to this γ-secretase inhibitor. This finding is consistent with a previous study in which human p75^{NTR} and Xenopus NRH1 were investigated *in vitro* [8]. Furthermore, our analysis of a chimeric Nrh1 protein in which the Nrh1 transmembrane domain is replaced by that of p75^{NTR} revealed that this domain alone is not sufficient to confer γ-secretase cleavage susceptibility. This

is not completely unexpected as there is evidence to suggest that other factors are crucial for selection/cleavage by the γ-secretase complex [5]. A greater understanding of the specificity of γ-secretase substrate selection might be reached by extending the chimaeric approach by exchanging each of the domains of these proteins. Indeed, a study (that we discovered while writing this manuscript) using a similar domain swap approach but with two unrelated type 1 transmembrane proteins (the non γ-secretase substrate Itgβ1 and γ-secretase substrate vasorin) found that both a permissive transmembrane and a permissive intracellular domain were required for γ-secretase cleavage, confirming our findings [32].

Other than the prerequisites of being a type 1 membrane protein and of shedding the ectodomain it has previously been suggested that dimerisation of substrates and/or the structure of substrate α-helices may regulate γ-secretase activity [5]. This was not investigated in this study but is something that should be considered for future experiments using this assay system. Previous studies of the p75^{NTR} dimerisation domain AxxxGxxA found that, while this domain is not essential for dimerisation, altering its structure to LxxxLxxA via mutational analyses reduces γ-secretase cleavage [5, 33]. It was thought that this might be due to a stabilisation of the α-helix, inhibiting γ-secretase access by preventing unravelling of the transmembrane domain [5]. Zebrafish p75^{NTF} also contains the AxxxGxxA domain. However, in zebrafish Nrh1 the second alanine is replaced by leucine (AxxxGxxL) and we suggest this may have a similar effect to the LxxxLxxA mutant. It would be interesting to perform site directed mutagenesis on the L256 residue of the zebrafish Nrh1 dimerisation domain to assess whether this is a key feature preventing its cleavage by γ-secretase.

Another question that can potentially be addressed using this assay is the identity of the enzyme performing intramembranous cleavage of Nrh1. Since Kanning *et al* (2003) established that NRH1 does in fact produce cleavage products when treated with the PKC activator, PMA, a deeper investigation into what cleaves this protein may offer insights into alternative cleavage pathways for such substrates [8]. There are four main classes of proteases that perform intramembrane proteolysis. The first encompasses the aspartic proteases, including Presenilins (PSENs, the active subunit of γ-secretase), signal peptide peptidases (SPPs) and SPP-like proteases [34-36]. While γ-secretase cleaves proteins with type 1 membrane topology (C-termini towards the cytosol), the SPP and SPP-like proteases cleave proteins with type 2 membrane topology (N-termini towards the cytosol). Another class of I-CliPs consists of the Site-2 protease (S2P) and S2P-like proteases [37]. These are members of the metalloprotease family and cleave type 2 transmembrane proteases. Finally, there are the Rhomboid proteases, which are serine proteases [38]. Rhomboid proteases are highly specific in their substrate selection. The substrates selected are mostly type 1 transmembrane proteins, although there is also evidence to suggest they may cleave type 2 and multi pass membrane proteins in some cases [39, 40]. If we assume that NRH1 is a type 1 transmembrane protein like its homologue p75^{NTR}, then we can reasonably exclude two of the above classes of membrane cleaving proteases as candidates, namely, SPP (and SPP-like) and S2P (including S2P-like). However, the orientation of NRH1 within the membrane has not yet been investigated. Our assay could be used to test a range of protease inhibitors to identify which enzyme(s) cleave Nrh1.

The question of the effect of the α-helix structure of p75^{NTR} or NRH1 TMD on their cleavage susceptibility has not yet been investigated. Previous observations from a study of the effects of

the TMD of Gurken found that a β-sheet conformation was required for cleavage by the archaeal PRESENILIN homologue, MCMJR1. This raised the question of whether NRH1 can interact with PRESENILIN, but perhaps due to its TMD being in an α-helical conformation, cannot be cleaved by it. A simple amino acid sequence alignment of the zebrafish p75^{NTR} and Nrh1 TMDs (Figure 2.6) revealed that, while the p75^{NTR} TMD carries a proline residue that would supposedly allow it flexibility to conformationally change between an α-helix and β-sheet, the Nrh1 TMD lacks this residue. Interestingly, it has previously been observed that insertion of a single proline into the TMD can trigger cleavage of normally un-cleavable TMD's [31]. It may therefore be of interest in future to either insert or substitute a proline residue into the Nrh1 TMD to investigate its ensuing γ-secretase cleavage susceptibility. The results of such an experiment using our established assay might provide an answer for why Nrh1 is not naturally a γ-secretase substrate, while also contributing to the understanding of γ-secretase cleavage susceptibility. Regarding the previously observed cleavage of *Xenopus* NRH1 within its TMD and the question of what protease might be responsible for this cleavage, it has been observed that all intramembrane cleaving proteases (iCLIPs) prefer to cleave TMDs in their β-strand conformation [41]. If NRH1 is unable to enter this conformation, perhaps there is some other unknown enzyme responsible for this cleavage. If we wish to understand the cleavage properties of NRH1 it would be important to further investigate the conformational state of its TMD.

2.4 Materials and Methods

Ethics

This work was conducted under the auspices of the Animal Ethics Committee of the University of Adelaide and in accordance with EC Directive 86/609/EEC for animal experiments and the Uniform Requirements for manuscripts submitted to Biomedical journals.

Gene orthologue identification

Alignments and tree building were conducted using the Geneious software suite, version 5.6.7 (http://www.geneious.com, [13]). Alignments were performed with the following constraints; Cost matrix: Identity, Gap open penalty: 10, Gap extension penalty: 3, Alignment: Global. Bayesian trees were produced using the "Mr Bayes" program with the following constraints; Substitution model: GTR, Outgroup: Lancelet p75NTR, and the rest as default.

Constructs

ssFLAG-p75NTFC201-dGFP-v2a-mCherry, ssFLAG-Nrh1C191-dGFP-v2a-mCherry and ssFLAG-A2C-dGFP-v2a-mCherry DNA sequences were produced by Biomatik (complete DNA sequences are provided in supplemental file S4). These DNA sequences in the pBMH vector (provided by Biomaitk) were digested using *Bam*HI and *Cla*I in independent reactions and ligated into pT2AL200R150G between the *Bam*HI and *Cla*I sites.

Replacement of mCherry coding sequence by GFP coding sequence

GFP was amplified by polymerase chain reaction (PCR) using the following primers; Forward: 5'-GCTCTAGAATGGTGAGCAAGGGAGAGGA-3' and Reverse:

5'CCATCGATCTACTTGTACAGCTCGTCCATTCC-3'. The thermal cycling parameters were as follows: 98°C for 30 s, 15 cycles of 98°C for 10 s, 61°C for 30 s and 72°C for 30 s, followed by 72°C for 10 mins. mCherry was excised from pT2AL200R150GssFLAG-p75^{NTF}C201-dGFP-v2a-mCherry, pT2AL200R150G ssFLAG-Nrh1C191-dGFP-v2a-mCherry and pT2ALssFLAG-A2C-dGFP-v2a-mCherry by restriction digest with *Xba*1 and *Cla*1. The amplified GFP was then cloned between the *Xba*1 and *Cla*1 sites of pT2AL200R150GssFLAG-p75^{NTF}C201-dGFP-v2a-mCherry, pT2AL200R150G ssFLAG-Nrh1C191-dGFP-v2a-mCherry and pT2ALssFLAG-A2C-dGFP-v2a-mCherry.

DNA microinjection of zebrafish embryos and treatment with DAPT.

Tol2 transposase plasmid (pCS-TP) was linearised using Not1 (NEB) and mRNA was transcribed *in vitro* using the mMESSAGE mMACHINE SP6 Kit (Ambion Inc.). Fertilised embryos were injected with a solution containing 100ng/µl plasmid DNA and approximately 50ng/µl Tol2 transposase mRNA at the one cell stage. ~50 embryos injected with the injection solution above were placed into 35mm x 10mm petri dishes with 2ml E3 medium (15mM NaCl, 0.5mM KCl, 1mM MgSO$_4$, 0.15mM KH$_2$PO$_4$, 0.05mM Na$_2$HPO$_4$, 1mM CaCl$_2$, 0.7mM NaHCO$_3$). At 4hpf embryos were treated with 100µM DAPT (In solution^{TM} γ-secretase inhibitor IX, Calbiochem, San Diego, CA, USA) in 1% DMSO in E3 medium. Embryos were maintained at 28°C in a humid incubator. At 24hpf embryos were visualised under UV light for GFP expression. Embryos expressing GFP were selected for protein extraction.

Treatment of zebrafish embryos with Z-DEVD-FMK

Injected embryos were placed into 35mm x 10mm petri dishes with 2mL E3 medium (15mM NaCl, 0.5mM KCl, 1mM MgSO4, 0.15mM KH2PO4, 0.05mM Na2HPO4, 1mM CaCl2, 0.7mM NaHCO3). At 6hpf embryos were treated with 100µM Z-DEVD-FMK (ApexBio Technology LLC, Houston, TX, USA) in E3 medium. Embryos were maintained in standard temperature conditions in a humid incubator and checked for survival at 24hpf.

Western Immunoblot analyses

Dechorioned and de-yolked embryos were lysed by placement in sample buffer (2% sodium dodecyl sulfate (SDS), 5% β-mercaptoethanol, 25% v/v glycerol, 0.0625 M Tris-HCl (pH 6.8), and bromophenol blue) followed immediately by heating to 95°C for 10 min, before storage at -80°C prior to protein separation on 4-12% SDS polyacrylamide gels. Proteins were transferred to nitrocellulose membrane in buffer (25mM Tris, 192mM glycine, 0.1% sodium lauryl sulfate, 20% methanol in MilliQ H_2O) at 10V for 1hr. When immunoblotting, all membranes were blocked with 5% Western Blocking Reagent (Roche, Indianapolis, IN, USA) in TBST, incubated with primary antibodies in TBST containing 0.5% Western Blocking Reagent (Roche, Indianapolis, IN, USA), washed in TBST, and incubated in secondary antibody. Following secondary antibody incubations, all membranes were washed three times for 10 minutes in TBST and visualised with luminol reagents (Amresco, Ohio, USA or Thermo Scientific, Rockford, USA) by the ChemiDoc™ MP imaging system (Bio-Rad, Hercules, CA, USA). The p75[NTF]C201-dGFPx2, Nrh1C191-dGFPx2 and A2C-dGFPx2 protein bands were visualised at ~61kDa, 57kDa and 57kDa respectively (data not shown). Using Image Lab software (Bio-Rad), densitometry analyses were performed on the protein bands in each lane of the western immunoblot for each sample and free GFP internal

reference. An average value was obtained for the free GFP for each membrane. Each sample value was then normalised to the average free GFP value.

GFP immunoblots were incubated in a 1/5,000 dilution of anti-GFP antibodies (Rockland Immunochemicals Inc., Gilbertsville, PA, USA) and a 1/10,000 dilution of donkey anti-Goat IgG (Rockland Immunochemicals Inc., Gilbertsville, PA, USA).

mCherry immunoblots were incubated in a dilution of 1/2,000 of anti-mCherry antibody (Abcam, Cambridge, UK) and a 1/5,000 dilution of anti-Mouse IgG secondary antibodies (Rockland Immunochemicals Inc., Gilbertsville, PA, USA).

2.5 References

1.   Duggan, S.P. and J.V. McCarthy, *Beyond gamma-secretase activity: The multifunctional nature of presenilins in cell signalling pathways.* Cell Signal, 2016. **28**(1): p. 1-11.
2.   Beel, A.J. and C.R. Sanders, *Substrate specificity of gamma-secretase and other intramembrane proteases.* Cell Mol Life Sci, 2008. **65**(9): p. 1311-34.
3.   Haapasalo, A. and D.M. Kovacs, *The many substrates of presenilin/gamma-secretase.* J Alzheimers Dis, 2011. **25**(1): p. 3-28.
4.   Shah, S., et al., *Nicastrin functions as a gamma-secretase-substrate receptor.* Cell, 2005. **122**(3): p. 435-47.
5.   Sykes, A.M., et al., *The effects of transmembrane sequence and dimerization on cleavage of the p75 neurotrophin receptor by gamma-secretase.* J Biol Chem, 2012. **287**(52): p. 43810-24.
6.   Struhl, G. and A. Adachi, *Requirements for presenilin-dependent cleavage of notch and other transmembrane proteins.* Mol Cell, 2000. **6**(3): p. 625-36.
7.   Yan, Y., et al., *Defining the minimum substrate and charge recognition model of gamma-secretase.* Acta Pharmacol Sin, 2017. **38**(10): p. 1412-1424.
8.   Kanning, K.C., et al., *Proteolytic processing of the p75 neurotrophin receptor and two homologs generates C-terminal fragments with signaling capability.* J Neurosci, 2003. **23**(13): p. 5425-36.
9.   Gentry, J.J., P.A. Barker, and B.D. Carter, *The p75 neurotrophin receptor: multiple interactors and numerous functions.* Prog Brain Res, 2004. **146**: p. 25-39.
10.   Sotthibundhu, A., et al., *Beta-amyloid(1-42) induces neuronal death through the p75 neurotrophin receptor.* J Neurosci, 2008. **28**(15): p. 3941-6.
11.   Yaar, M., et al., *Binding of beta-amyloid to the p75 neurotrophin receptor induces apoptosis. A possible mechanism for Alzheimer's disease.* The Journal of clinical investigation, 1997. **100**(9): p. 2333-2340.

12. Jayne, T., *Developing in vivo assays for investigating p75NTR and NRH1 transmembrane cleavage events using zebrafish*, in *Biological Sciences*. 2013, The University of Adelaide: Adelaide, Australia.

13. Kearse, M., et al., *Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.* Bioinformatics, 2012. **28**(12): p. 1647-9.

14. Wilson, L. and M. Lardelli, *The development of an in vivo gamma-secretase assay using zebrafish embryos.* J Alzheimers Dis, 2013. **36**(3): p. 521-34.

15. Krogh, A., et al., *Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.* J Mol Biol, 2001. **305**(3): p. 567-80.

16. Cserzo, M., et al., *Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method.* Protein Eng, 1997. **10**(6): p. 673-6.

17. Weskamp, G., et al., *Evidence for a critical role of the tumor necrosis factor alpha convertase (TACE) in ectodomain shedding of the p75 neurotrophin receptor (p75NTR).* J Biol Chem, 2004. **279**(6): p. 4241-9.

18. Zampieri, N., et al., *Cleavage of p75 neurotrophin receptor by alpha-secretase and gamma-secretase requires specific receptor domains.* J Biol Chem, 2005. **280**(15): p. 14563-71.

19. Einhauer, A. and A. Jungbauer, *The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins.* J Biochem Biophys Methods, 2001. **49**(1-3): p. 455-65.

20. Bichara, M., J. Wagner, and I.B. Lambert, *Mechanisms of tandem repeat instability in bacteria.* Mutat Res, 2006. **598**(1-2): p. 144-63.

21. Barash, S., W. Wang, and Y. Shi, *Human secretory signal peptide description by hidden Markov model and generation of a strong artificial signal peptide for secreted protein expression.* Biochem Biophys Res Commun, 2002. **294**(4): p. 835-42.

22. Provost, E., J. Rhee, and S.D. Leach, *Viral 2A peptides allow expression of multiple proteins from a single ORF in transgenic zebrafish embryos.* Genesis, 2007. **45**(10): p. 625-9.

23. Urra, S., et al., *TrkA receptor activation by nerve growth factor induces shedding of the p75 neurotrophin receptor followed by endosomal gamma-secretase-mediated release of the p75 intracellular domain.* J Biol Chem, 2007. **282**(10): p. 7606-15.

24. Skeldal, S., et al., *Mapping of the interaction site between sortilin and the p75 neurotrophin receptor reveals a regulatory role for the sortilin intracellular domain in p75 neurotrophin receptor shedding and apoptosis.* J Biol Chem, 2012. **287**(52): p. 43798-809.

25. Kawakami, K., et al., *A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish.* Dev Cell, 2004. **7**(1): p. 133-44.

26. Yabu, T., et al., *Characterization of zebrafish caspase-3 and induction of apoptosis through ceramide generation in fish fathead minnow tailbud cells and zebrafish embryo.* Biochem J, 2001. **360**(Pt 1): p. 39-47.

27. Abraham, N., *Role of programmed cell death in defining zebrafish development [dissertation on the internet]*, in *Biological Sciences*. 2004, St. John's University: New York, NY.

28. Zhou, R., et al., *Recognition of the amyloid precursor protein by human γ-secretase.* Science, 2019. **363**(6428): p. eaaw0930.

29. Yang, G., et al., *Structural basis of Notch recognition by human γ-secretase.* Nature, 2019. **565**(7738): p. 192-197.

30. Clemente, N., et al., *Coupled Transmembrane Substrate Docking and Helical Unwinding in Intramembrane Proteolysis of Amyloid Precursor Protein.* Sci Rep, 2018. **8**(1): p. 12411.

31. Brown, M.C., et al., *Unwinding of the Substrate Transmembrane Helix in Intramembrane Proteolysis.* Biophysical Journal, 2018. **114**(7): p. 1579-1589.

32. Hemming, M.L., et al., *Proteomic profiling of gamma-secretase substrates and mapping of substrate requirements.* PLoS Biol, 2008. **6**(10): p. e257.

33. Nadezhdin, K.D., et al., *Structural Basis of p75 Transmembrane Domain Dimerization.* J Biol Chem, 2016. **291**(23): p. 12346-57.

34.   Wolfe, M.S., et al., *Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity.* Nature, 1999. **398**(6727): p. 513-7.

35.   Weihofen, A., et al., *Identification of signal peptide peptidase, a presenilin-type aspartic protease.* Science, 2002. **296**(5576): p. 2215-8.

36.   Friedmann, E., et al., *SPPL2a and SPPL2b promote intramembrane proteolysis of TNFalpha in activated dendritic cells to trigger IL-12 production.* Nat Cell Biol, 2006. **8**(8): p. 843-8.

37.   Rawson, R.B., et al., *Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs.* Mol Cell, 1997. **1**(1): p. 47-57.

38.   Urban, S., J.R. Lee, and M. Freeman, *Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases.* Cell, 2001. **107**(2): p. 173-82.

39.   Tsruya, R., et al., *Rhomboid cleaves Star to regulate the levels of secreted Spitz.* EMBO J, 2007. **26**(5): p. 1211-20.

40.   Fleig, L., et al., *Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins.* Mol Cell, 2012. **47**(4): p. 558-69.

41.   Madala, P.K., et al., *Update 1 of: Proteases Universally Recognize Beta Strands In Their Active Sites.* Chemical Reviews, 2010. **110**(6): p. PR1-PR31.

2.6 Supplementary information

This section is included in the thesis as information supplementary to Chapter 2. It contains additional information not included in the main text of the manuscript.

File S1. Screenshots of results from tblastn searches using the NCBI database

- The amino acid sequence of human p75^{NTR} was used in a tblastn search of the zebrafish genome (constrained to RefSeq_RNA).

Figure S1. Duplication of the p75^{NTR} gene on chromosome 12 of Danio rerio. *Red box indicates duplicate genes.*

Description	Max score	Total score	Query cover	E value	Ident	Accession
Danio rerio nerve growth factor receptor b (ngfrb), mRNA	332	332	93%	6e-109	45%	NM_001198660.1
PREDICTED: Danio rerio nerve growth factor receptor a (TNFR superfamily, member 16) (ngfra), transcript variant X1, mRNA	270	270	67%	5e-81	52%	XM_003198085.5
PREDICTED: Danio rerio nerve growth factor receptor a (TNFR superfamily, member 16) (ngfra), transcript variant X2, mRNA	223	223	56%	6e-65	52%	XM_021471410.1
PREDICTED: Danio rerio neurotrophin receptor associated death domain (nradd), mRNA	189	189	84%	6e-52	36%	XM_695893.8
PREDICTED: Danio rerio tumor necrosis factor receptor superfamily member 5-like (LOC561000), transcript variant X2, mRNA	62.4	99.7	34%	1e-09	30%	XM_005173526.4
PREDICTED: Danio rerio tumor necrosis factor receptor superfamily member 5-like (LOC561000), transcript variant X1, mRNA	62.0	99.4	34%	2e-09	30%	XM_684398.9

Figure S2. Duplication of p75^{NTR} gene on chromosome 3 of Danio rerio, as of GRCz11. *The two returned genes (indicated by the red box) represent the revised position of the p75^{NTR} duplicate gene on chromosome 3.*

95

File S2. Supporting table for phylogenetic analyses

Table S1. Names, NCBI gene names and NCBI accession numbers of all genes used in phylogenetic studies

Organism and gene name	NCBI name	NCBI Accession nucleotide
Zebrafish $p75^{NTR}$	ngfr-l	NM_001198660.1
Zebrafish $p75^{NTR}$ like 12	TNF 16-like	XM_003199576.1
Zebrafish $p75^{NTR}$ like 3	ngfr-a	XM_003198085.3
Zebrafish Nrh1	ngfr	XM_695893.4
Xenopus tropicalis NRH1	ngfr	NM_001007998.2
Xenopus $p75^{NTR}$	p75 NTR a	NM_001088466.1
Xenopus laevis NRH1	ngfr	NM_001091773.1
Mouse $p75^{NTR}$	Ngfr	NM_033217.3
Mouse NRH2	Nradd	NM_026012.2
Chicken $p75^{NTR}$	NGFR	NM_001146133.1
Chicken NRH1	TNFR 16-like	XM_418509.3
Human $p75^{NTR}$	TNFR superfamily 16	NM_002507.3
Human NRH2	NRADDP	NR_024046
Lancelet $p75^{NTR}$	Branchiostoma floridae hypothetical protein	XM_002588230.1
Drosophila $p75^{NTR}$ like	Drosophila melanogaster rumpelstiltskin (rump), transcript variant A	NM_141642.3
Human $p75^{NTR}$ like	TNFRSF11B	NM_002546.3

File S3. Transmembrane domain predictions for p75^{NTR} and Nrh1
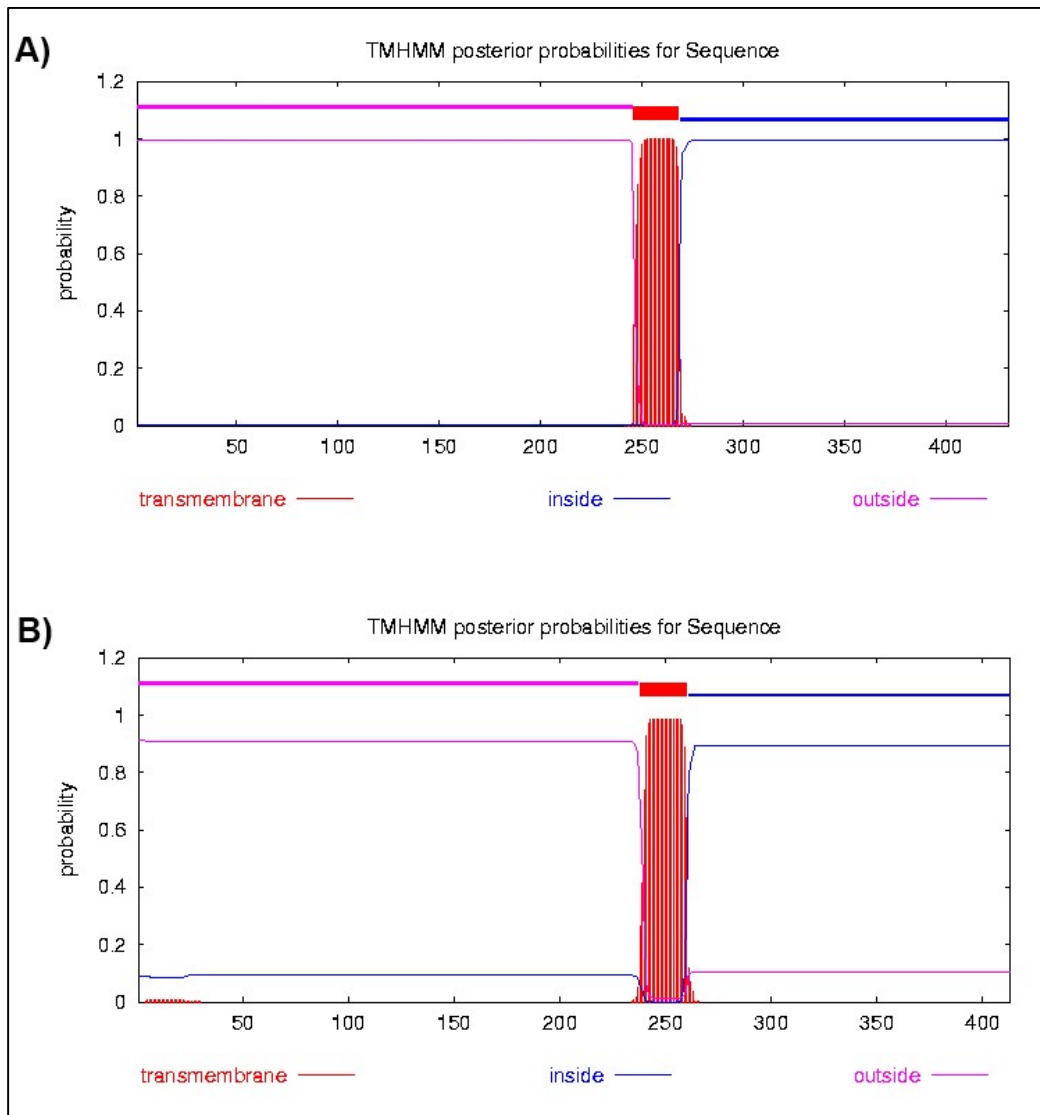
Figure S3. Transmembrane domain predictions for Nrh1 and p75^{NTR} using TMHMM. *Plot of probabilities generated by TMHMM 2.0. A) p75^{NTR} TMD prediction B) Nrh1 TMD prediction. "outside" refers to the prediction that the sequence sits on the cytosolic side of the membrane and "inside" refers to the prediction that the sequence sits on the non-cytosolic (lumenal) side of the membrane. "transmembrane" refers to predicted transmembrane helices in the sequence.*

File S4. Construct overviews and full sequences

(A) Sequence design for the pT2ALssFLAG-p75^{NTF}C201-dGFP-v2a-mCherry construct

Different regions of the sequence are indicated with colours corresponding to the legend below:

- BamHI-Kozak-START-SS-FlagTag1-FlagTag2-FlagTag3-p75(CTF)-eGFP-BspEI-MODCPEST-BspEI-v2a-Xba1-mCherry-Xba1-STOP-NotI-ClaI

- Capitalised C in FlagTag2 = silent mutation, Capitalised T in FlagTag3 = silent mutation

Sequence of pT2ALssFLAG-p75NTFC201-dGFP-v2a-mCherry construct:

```
5'-
cgggatccgccaccatgtggtggcgcctgtggtggctgctgctgctgctgctgctgctgtggcc
catggtgtgggccgattacaaggatgacgacgataaggattacaaggaCgacgacgataaggat
tacaaggatgacgaTgataagaccacgactgaccctgggcaaaagagactacatgggcttagcg
acaacctcattcccatctacacctccatactggctgcggtgctgttgggcctggtggccttcat
catatttaaacggtggaacagttgtaagcagaataagcaagccataaccgagcgtgcagcgca
aaccccagtcagactccttctcctgagggagagaaactccacagtgacagcggcatctcagtgg
acagtcagagcctgcaggacggccagggcctccacatacagtggtcaagatagatggaggttc
ggctctgtctttgcccttacacacacgggaggaggtggagaaactgctaaatcgtaccaatgaa
ggagaggagtcagcagccaatgaggagaccgactggtgcagcctagccgggctacttggataca
aagaagaacacattgctaatttcaagcaggaggaacggcccatccaggcacttctgtcgcactg
ggcgagccaggattcggctaacattgatacgctttgcacagctctgaagaagatcaacagagag
gatattgcgcaaagcatcattgtcaaaccaactgccacatctgccgtaatggtgagcaagggcg
aggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaa
gttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatc
tgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgc
agtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccga
aggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgag
gtgaagttcgagggcgacaccctggtgaaccgcatcgagctgaagggcatcgacttcaaggagg
acggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggc
cgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagc
gtgcagctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgcccg
acaaccactacctgagcacccagtccgccctgagcaaagaccccaacgagaagcgcgatcacat
ggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaagggc
tccggacatggcttccgccggcggtggcggcgcaggatgatggcacgctgcccatgtcttgtg
cccaggagagcgggatggaccgtcaccctgcagcctgtgcttctgctaggatcaatgtgtggctc
cggagctacaaatttctctctgttgaaacaggctggtgacgtcgaggagaatcctggcccatct
agaatggtttccaaaggagaggaggacaatatggcaatcatcaaagagttcatgcgttttaaag
ttcacatggaaggcagtgttaacggacatgaatttgagatagagggcgagggggagggcagacc
gtacgaaggaacacagaccgcaaaactgaaagtgaccaagggaggaccattgcctttcgcctgg
gatattctgtcaccacaattcatgtatggttcaaaagcctacgtcaagcacccggctgacatac
cggattatctgaagctaagctttcccgagggatttaaatgggagcgcgtgatgaatttcgaaga
```

```
tggcggcgtagtgaccgtgacacaggatagctctcttcaagatggggaatttatttacaaggtc
aagttacgcggcactaacttcccttctgatggaccagtgatgcagaaaaaaacaatgggttggg
aagctagctccgagaggatgtacccggaggacggtgcgcttaaaggtgagatcaagcaacgatt
gaaattgaaagacggcgggcattatgacgcagaggttaagaccacatacaaagccaagaagccc
gtccagcttcccggagcctataacgtcaacatcaaacttgatattacttctcacaacgaggatt
acaccatagttgagcagtacgaaagagcagagggcaggcattccaccggggggatggacgagct
gtacaagtctagatagcggccgcatcgatgg-3'
```

(B) Sequence design for the pT2ALssFLAG-Nrh1C191-dGFP-v2a-mCherry construct

Different regions of the sequence are indicated with colours corresponding to the legend below:

- BamHI-Kozak-START-SS-FlagTag1-FlagTag2-FlagTag3-Nrh1(CTF)-eGFP-BspEI-MODCPEST-BspEI-v2a-Xba1-mCherry-Xba1-STOP-NotI-ClaI

- Capitalised C in FlagTag2 = silent mutation, Capitalised T in FlagTag3 = silent mutation

Sequence of pT2ALssFLAG-Nrh1C191-dGFP-v2a-mCherry construct:

```
5'-
cgggatccgccaccatgtggtggcgcctgtggtggctgctgctgctgctgctgctgctgtggcc
catggtgtgtggccgattacaaggatgacgacgataaggattacaaggaCgacgacgataaggat
tacaaggatgacgaTgataagggctctggagcccctagactcacaccacaggaccaaggtggaa
acaacaacatcttggtgtatgtgtctgttttggctgctgtggtgcttggcctgctgctctacgt
tgcctacaaatgctggaaatcatgtcagcagaagcaggctctggtgaaggcccgggtcggagag
ctgaataatgcaggggaaggagagaaactgcacagcgacagtggtgttttcctggactctcaca
gccttcaggaaagccagcctagcaaaggcagtaaacgggacagcaaacaggatacacggctcta
cataaacctgcctccgcacagacaggaggaggtagagggacttctggctgagggggggcaatcga
agctggaaacagctagccgctacactaggctacgaacaggagcgcgtggacgtctttggacggg
gccaggaccccatccacaccctcatgaccgattggtcccagcaggaaggctccacattgggttt
gttgtgttcagctctgactcgcatcgagcgaccagacatcatcactgctctgaccgccccaacg
caaggagtatcagtggtcatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcc
tggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcga
tgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctgg
cccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatga
agcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttctt
caaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaac
cgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagt
acaactacaacagccacaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaa
cttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaac
acccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtccgccc
tgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgg
gatcactctcggcatggacgagctgtacaagggctccggacatggcttcccgccggcggtggcg
gcgcaggatgatggcacgctgcccatgtcttgtgcccaggagagcgggatggaccgtcaccctg
```

```
cagcctgtgcttctgctaggatcaatgtgggctccggagctacaaatttctctctgttgaaaca
ggctggtgacgtcgaggagaatcctggcccatctagaatggtttccaaaggagaggaggacaat
atggcaatcatcaaagagttcatgcgttttaaagttcacatggaaggcagtgttaacggacatg
aatttgagatagagggcgagggggagggcagaccgtacgaaggaacacagaccgcaaaactgaa
agtgaccaagggaggaccattgcctttcgcctgggatattctgtcaccacaattcatgtatggt
tcaaaagcctacgtcaagcacccggctgacataccggattatctgaagctaagctttcccgagg
gatttaaatgggagcgcgtgatgaatttcgaagatggcggcgtagtgaccgtgacacaggatag
ctctcttcaagatggggaatttatttacaaggtcaagttacgcggcactaacttcccttctgat
ggaccagtgatgcagaaaaaacaatgggttgggaagctagctccgagaggatgtacccggagg
acggtgcgcttaaggtgagatcaagcaacgattgaaattgaaagacggcgggcattatgacgc
agaggttaagaccacatacaaagccaagaagcccgtccagcttcccggagcctataacgtcaac
atcaaacttgatattacttctcacaacgaggattacaccatagttgagcagtacgaaagagcag
agggcaggcattccaccggggggatggacgagctgtacaagtctagatagcggccgcatcgat
gg-3'
```

(C) Sequence design for the pT2ALssFLAG-A2C-dGFP-v2a-mCherry construct

Different regions of the sequence are indicated with colours corresponding to the legend below:

- BamHI-Kozak-START-SS-FlagTag1-FlagTag2-FlagTag3- NRH1 ECD- P75 TMD-NRH1 ICD-eGFP-BspEI-MODCPEST-BspEI-v2a-XbaI-mCherry-XbaI-STOP-NotI-ClaI

- Capitalised C in FlagTag2 = silent mutation, Capitalised T in FlagTag3 = silent mutation

Sequence of pT2ALssFLAG-A2C-dGFP-v2a-mCherry construct:

```
5'-
cgggatccgccaccatgtggtggcgcctgtggtggctgctgctgctgctgctgctgctgtggcc
catgtgtgtgggccgattacaaggatgacgacgataaggattacaaggaCgacgacgataaggat
tacaaggatgacgaTgataaggctctggagcccctagactcacaccacaggaccaaggtggaa
acaacctcattcccatctacacctccatactggctgcggtgctgttgggcctggtggccttcat
catatttaaatgctggaaatcatgtcagcagaagcaggctctggtgaaggcccgggtcggagag
ctgaataatgcaggggaaggagagaaactgcacagcgacagtggtgtttcctggactctcaca
gccttcaggaaagccagcctagcaaaggcagtaaacgggacagcaaacaggatacacggctcta
cataaacctgcctccgcacagacaggaggaggtagagggacttctggctgagggggcaatcga
agctggaaacagctagccgctacactaggctacgaacaggagcgcgtggacgtctttggacggg
gccaggaccccatccacaccctcatgaccgattggtcccagcaggaaggctccacattgggtttt
gttgtgttcagctctgactcgcatcgagcgaccagacatcatcactgctctgaccgccccaacg
caaggagtatcagtggtcatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcc
tggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcga
tgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctgg
cccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatga
```

100

agcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttctt
caaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaac
cgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagt
acaactacaacagccacaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaa
cttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaac
acccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtccgccc
tgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgg
gatcactctcggcatggacgagctgtacaagggctccggacatggcttcccgccggcggtggcg
gcgcaggatgatggcacgctgcccatgtcttgtgcccaggagagcgggatggaccgtcaccctg
cagcctgtgcttctgctaggatcaatgtgggctccggagctacaaatttctctctgttgaaaca
ggctggtgacgtcgaggagaatcctggcccatctagaatggtttccaaaggagaggaggacaat
atggcaatcatcaaagagttcatgcgttttaaagttcacatggaaggcagtgttaacggacatg
aatttgagatagagggcgaggggggagggcagaccgtacgaaggaacacagaccgcaaaactgaa
agtgaccaagggaggaccattgcctttcgcctgggatattctgtcaccacaattcatgtatggt
tcaaaagcctacgtcaagcacccggctgacataccggattatctgaagctaagctttcccgagg
gatttaaatgggagcgcgtgatgaatttcgaagatggcggcgtagtgaccgtgacacaggatag
ctctcttcaagatggggaatttatttacaaggtcaagttacgcggcactaacttccccttctgat
ggaccagtgatgcagaaaaaaacaatggttgggaagctagctccgagaggatgtacccggagg
acggtgcgcttaaaggtgagatcaagcaacgattgaaattgaaagacggcgggcattatgacgc
agaggttaagaccacatacaaagccaagaagcccgtccagcttcccggagcctataacgtcaac
atcaaacttgatattacttctcacaacgaggattacaccatagttgagcagtacgaaagagcag
agggcaggcattccaccgggggggatggacgagctgtacaagtctagatagcggccgcatcgat
gg-3'

(4) Sequence of GFP that was PCR amplified and sub-cloned into constructs

5'-
atggtgagcaagggagaggagctgttcacaggagtggtgcctatcctggtggagctggacggag
acgtgaacggacacaagttcagcgtgagcggagagggagagggagacgctacatacggaaagct
gacactgaagttcatctgtacaacaggaaagctgcctgtgccttggcctacactggtgacaaca
ctgacatacggagtgcagtgtttcagcagataccctgaccacatgaagcagcacgacttcttca
agagcgctatgcctgagggatacgtgcaggagagaacaatcttcttcaaggacgacggaaacta
caagacaagagctgaggtgaagttcgagggagacacactggtgaacagaatcgagctgaaggga
atcgacttcaaggaggacggaaacatcctgggacacaagctggagtacaactacaacagccaca
acgtgtacatcatggctgacaagcagaagaacggaatcaaggtgaacttcaagatcagacacaa
catcgaggacggaagcgtgcagctggctgaccactaccagcagaacacacctatcggagacgga
cctgtgctgctgcctgacaaccactacctgagcacacagagcgctctgagcaaggaccctaacg
agaagagagaccacatggtgctgctggagttcgtgacagctgctggaatcacactgggaatgga
cgagctgtacaag-3'

File S5. Supporting tables for western blot analyses

Table S2. Intensity ratios from western immunoblots for Figure 2.3

p75^{NTF}C201-dGFP injected, without treatment and collected at 24hpf		
p75^{NTF}C201-dGFP	**free mCherry**	**p75^{NTF}C201-dGFP/free mCherry**
3355616	2776536	1.208562
1361151	2087967	0.651903
2030385	2921364	0.695013
p75^{NTF}C201-dGFP injected, with DAPT treatment and collected at 24hpf		
p75^{NTF}C201-dGFP	**free mCherry**	**p75^{NTF}C201-dGFP/free mCherry**
8523834	4229267	2.01544
10723713	3184398	3.367579
2570365	2241720	1.146604
Nrh1C191-dGFP injected, without treatment and collected at 24hpf		
Nrh1C191-dGFP	**free mCherry**	**Nrh1C191-dGFP/free mCherry**
4786670	2897640	1.65192
5352372	3658428	1.463025
2862144	4299228	0.665734
2482850	4824342	0.51465
Nrh1C191-dGFP injected, with DAPT treatment and collected at 24hpf		

Nrh1C191-dGFP	free mCherry	Nrh1C191-dGFP/free mCherry
3344274	2456604	1.36134
4696055	5235246	0.897008
3129192	3895992	0.803182
1659636	4562604	0.363748

Table S3. Intensity ratios from western immunoblots for Figure 2.4

p75[NTF]C201-dGFPx2 injected, without treatment and collected at 24hpf		
p75[NTF]C201-dGFPx2	free GFP	p75[NTF]C201-dGFPx2/free GFP
4099260	21408354	0.191479457
2265235	18651920	0.121447819
2760660	13464144	0.205037914
4012008	31961958	0.125524475
3141090	24936450	0.1259638
1371942	7807518	0.175720632
1161584	6491984	0.178925888
784560	5346448	0.146744156
p75[NTF]C201-dGFPx2 injected, with DAPT treatment and collected at 24hpf		
p75[NTF]C201-dGFPx2	free GFP	p75[NTF]C201-dGFPx2/free GFP
9307074	19496964	0.477360167
11522280	15773135	0.73050031

9746802	19327971	0.504284801
2743486	27871788	0.098432365
8096400	15898455	0.509257032
3262120	5782000	0.564185403
1073216	5489744	0.195494726
3258768	5566128	0.585464078

Nrh1C191-dGFPx2 injected, without treatment and collected at 24hpf

Nrh1C191-dGFPx2	free GFP	Nrh1C191-dGFPx2/free GFP
1491240	259845	5.738959764
3290364	445676	7.382861092
3774136	1123326	3.359786918
14486528	27954624	0.518215806
9071304	16119450	0.562755181
2069262	2441270	0.847617019
7176832	3022704	2.374308566
4717200	2825728	1.669375113

Nrh1C191-dGFPx2 injected, with DAPT treatment and collected at 24hpf

Nrh1C191-dGFPx2	free GFP	Nrh1C191-dGFPx2/free GFP
2157554	286594	7.528259489
2754024	563472	4.887596899
1687808	386528	4.366586638
16519654	30580320	0.5402054
15641976	20213136	0.773852014

6142512	3632768	1.690862725
4824704	3712496	1.299584969
6105872	2351728	2.59633427

Table S4. Intensity ratios from western immunoblots for Figure 2.5

A2C -dGFPx2 injected without treatment and collected at 24hpf		
A2C -dGFPx2	**free GFP**	**A2C -dGFPx2/free GFP**
5968380	4111428	1.451656213
3976478	2930257	1.357040696
3134032	1885752	1.661953428
2948385	2764267	1.066606446
5134025	5889052	0.871791419
A2C -dGFPx2 injected with DAPT treatment and collected at 24hpf		
A2C -dGFPx2	**free GFP**	**A2C -dGFPx2/free GFP**
4955500	2678104	1.850376236
2455871	1668183	1.472183208
3267396	3427666	0.953242235
3248465	2896872	1.121369878
4519932	4193688	1.077794056

Chapter 3 – Development and testing of genome editing tools for generation of fAD and null mutations in zebrafish *appa* and *appb*

3.1 Introduction

Missense mutations in the amyloid precursor protein (APP) can lead to an inheritable form of Alzheimer's disease known as familial Alzheimer's disease (fAD). The proposed mechanism through which mutations in APP lead to Alzheimer's disease involves its processing by the γ-secretase complex. APP is cleaved sequentially in one of two pathways. In the pathway that leads to the formation of the amyloid beta peptide, APP is first cleaved by β-secretase, then γ-secretase. fAD mutations in APP are mostly clustered around β- and γ-secretase cleavage sites. There are currently 59 known *APP* mutations involved in fAD (https://www.alzforum.org/mutations/app).

There are two co-orthologues of the human *APP* gene in zebrafish, *appa* and *appb* [1]. The zebrafish has been used previously to study the functions of *appa* and *appb* [2]. A search of the literature reveals only one paper describing an Alzheimer's disease zebrafish transgenic model of APP. Pu *et al*'s (2017) approach was to knock in mutant human *APP* driven by the zebrafish *appb* promoter [3]. This model was used to observe behaviour under mutant *APP* conditions as well as associated histopathology [3]. Other work on the zebrafish *app* genes has focused on investigating its roles in traumatic brain injury and development [4].

Although many previous attempts to study AD in mice have focussed on creating transgenic models of fAD genes, Hargis and Blalock (2017) observed inconsistent modelling of established transcriptomics-based changes detected in human AD in many of these models [5]. Considering this, we propose that the best approach to model the disease, and perform meaningful

transcriptomics analysis on these models, is to recreate (to the best of our ability) known fAD mutations in the equivalent endogenous zebrafish gene, via genome editing.

In recent years it has become possible to generate targeted mutations in the zebrafish genome. Two systems adapted for accomplishing this are: transcription activator-like effector nucleases (TALENs) and the clustered regulatory interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system. While both systems function by generating double-stranded breaks (DSB's) in the DNA, there are differences in their targeting mechanics and composition. TALENs initiate DNA cleavage through the binding of a pair of customisable DNA binding domains, both of which are fused to a non-specific *Fok*1 endonuclease domain, thus allowing *Fok*1 to dimerise and generate DSBs [6]. When used in zebrafish, the type II CRISPR/Cas9 system initiates cleavage through a single transcript called "sgRNA". This sgRNA is a combination of the single 20bp guide RNA (gRNA) and tracrRNA used in other systems and is responsible for recruiting the Cas9 enzyme to the site of interest for DNA cleavage [7].

DSBs generated by either the TALEN or CRISPR/Cas9 system result in the activation of one of two inherent cellular DNA repair mechanisms. Non-homologous end joining (NHEJ), which repairs DNA breaks by re-joining the cleaved ends, is known to randomly introduce insertions or deletions (indels) and is often exploited for gene editing [8]. Alternatively, homology directed repair (HDR) is a process where the cell repairs breaks in DNA using a template with homology to the original DNA sequence. This form of repair can be utilised to insert precise mutations using an oligonucleotide coding for the desired mutation [9, 10]. By utilising the cell's HDR machinery, both TALENs and CRISPRs can be used to engineer endogenous zebrafish genes to incorporate specific single nucleotide mutations. The mutations to be engineered into the zebrafish genome in

this study are human fAD *APP* missense mutations, as well as null mutations in zebrafish *appa* and *appb*, discussed in more detail below.

In this study, we designed and analysed sgRNAs targeting 3 fAD missense mutation sites from human APP in the equivalent positions of zebrafish *appa* and *appb*: the Austrian mutation (human APP T714I), found at the γ-secretase cleavage site, which presents with classic AD histopathology and increased Aβ-42:Aβ-40 ratio [11]; the Leuven mutation (human APP E682K), found at the (lesser known) β'-secretase cleavage site, reduces processing by this anti-amyloidogenic secretase and results in increased processing by β-secretase [12]; and the Arctic mutation (human APP E693G), occurring close to the α-secretase cleavage site, sees reduced levels of α-secretase processing thought to be due to altered APP localisation, with a subsequent shift to processing by β-secretase [13]. At the time of this study, we failed to design a sgRNA to target the most extensively studied fAD mutation at the γ-secretase cleavage position, the London mutation (human APP V717I) [14]. Fortunately, due to different design constraints, we were able to order a TALEN to target this residue's codon. We also designed sgRNAs to target near the start codons of both zebrafish *appa* and *appb* with the intention of generating null mutations. We found that, of the 7 sgRNAs and 1 TALEN tested, only 3 were able to induce DSBs. Unfortunately, with our knowledge at the time this study was being undertaken, we were unable to introduce fAD missense mutations through HDR. However, we have identified two sgRNAs that can be used to introduce fAD mutations, and we discuss ways in which this could be achieved in future studies. Finally, we identified a sgRNA that was able to generate DSBs near the start site of *appb*. This sgRNA was used successfully to generate a putative null mutation of *appb* and is discussed in detail in Chapter 4.
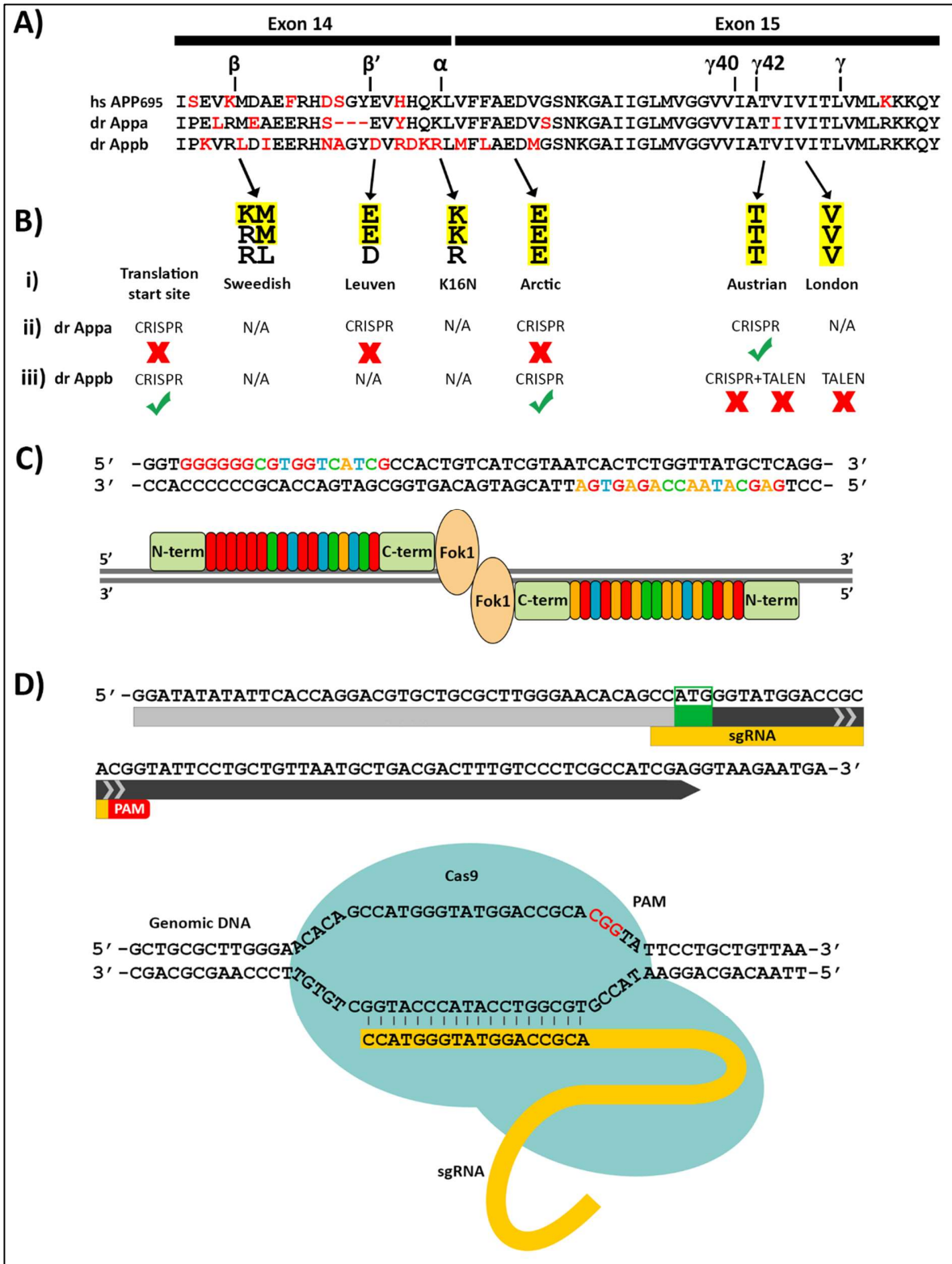
Figure 3.1 Overview of the CRISPR/TALEN mutational strategy employed in the work described in this chapter. A) Protein alignment of the Aβ peptide region of human APP (isoform hs APP695), zebrafish Appa (dr Appa) and zebrafish Appb (dr Appb). Secretase cleavage positions are indicated above the alignment. Red text indicates amino acid residues that are not conserved where there is a consensus of two between any two sequences. B) Overview of fAD mutations discussed in Chapter 3. Yellow highlighting indicates conserved residues. i) Names of potential mutation sites to be targeted by genome editing technology, ii) overview of mutational strategy for zebrafish Appa (dr Appa) and iii) zebrafish Appb (dr Appb). Specific genome editing technology used in this study to target each mutation position is indicated in text. Successful or unsuccessful generation of a DSB at each of the targeted positions is indicated by a green tick or a red cross respectively. C) Sequence and representation of TALEN arms designed by ZGeneBio targeting the London mutation equivalent position in zebrafish Appb. D) Example of CRISPR design using sgRNA targeting near the appb start codon. Grey labelled boxes under the sequence indicate the first Exon, with the pale grey box indicating specifically the 5'UTR. The green box indicates the start codon. Yellow and red boxes indicate the sgRNA binding site and PAM site respectively. Blue ovals indicate positioning of the Cas9 protein and yellow strand indicates the sgRNA position during DSB induction.

3.2 Results and Discussion

Assessment and selection of APP fAD mutations to target

The zebrafish has been used comprehensively as a model for human disease research [as reviewed in 15]. We aimed to generate several mutations in the zebrafish *appa* and *appb* genes, including null mutations and fAD-like mutations. We hoped to observe their effects on molecular/signalling events and to investigate whether fAD-like mutations in the zebrafish are appropriate models of the analogous mutations in humans. We intended to compare the effects of the *appa* and *appb* fAD-like mutations with the effects of fAD-like mutations in zebrafish *psen1*, *psen2* and *sorl1*.

We also intended to compare the null mutants with the effects of fAD-like mutations in zebrafish *psen1*, *psen2*, *appa*, *appb* and *sorl1*. Closer inspection of the protein alignment between zebrafish Appa, Appb and human APP revealed residues that are conserved within the α-, β- and γ-secretase cleavage regions (Figure 3.1, A). A selection of residues that are mutated in fAD have been highlighted and relevant cleavage positions marked.

When selecting human APP fAD mutations to be introduced, three criteria were considered. Firstly, the selected mutations should give an even coverage across the length of the Aβ peptide region of the APP protein. Secondly, the mutations should be at one of each of the α, β and γ-secretase cleavage positions within the Aβ peptide region (Figure 3.1, A). Lastly, in order to make use of the zebrafish system, the human APP fAD mutant amino acid residues needed to be conserved in the equivalent zebrafish App protein sequence. Following these criteria, the Arctic, Leuven, Austrian and London human APP fAD mutations were selected for introduction into the zebrafish genome by genome editing technologies (Figure 3.1, B).

The "Swedish" mutation is the most well characterised human APP fAD mutation. An initial alignment of the zebrafish Appa and Appb amino acid sequences against human APP revealed a lack of conservation of the residues affected by the "Swedish" mutation at the β-secretase sites of these proteins (Figure 3.1, B). Therefore, we decided to model the Leuven human APP fAD mutation at the alternative β'-secretase site, the equivalent residue of which only occurs in zebrafish Appa [12]. The APP fAD mutation at the γ-secretase cleavage position that has been most extensively studied is the "London" mutation, hence this is a good candidate [16]. The "Austrian" mutation occurs three residues upstream of the "London" mutation and has been quite intensively studied due to its effects on Aβ peptide production [11]. Of the reported fAD human APP mutations (Alzforum mutations database, https://www.alzforum.org/mutations/app) no

pathogenic mutations existed directly at the α-secretase cleavage site. The "Arctic" mutation was selected as the closest human APP fAD mutation to the α-secretase site that was also conserved in both zebrafish Appa and Appb (Figure 3.2). Later it was discovered that a mutation at the α-secretase cleavage site (K16N) does exist and is potentially pathogenic [17], however, this mutation is only conserved in Appa, hence we proceeded with the "Arctic" mutation (Figure 3.1, B).

Although this mutation is not directly positioned at the site of α-secretase cleavage it does affect this cleavage event [18]. It is thought to affect cleavage by reducing the abundance of APP at the cell surface, thereby reducing its availability for cleavage by α-secretase. Instead it is cleaved by the alternative β-secretase pathway [18]. The relative positions of each mutation of interest can be found in Table 3.1 A complete list of APP mutations can be found online at https://www.alzforum.org/mutations/app.

Table 3.1 The fAD mutations of the human *APP* gene to be edited into the zebrafish genome using CRISPR/Cas9 technology and TALENs				
Human APP mutation	Equivalent zebrafish Appa mutation	Equivalent zebrafish Appb mutation	Description and reference (https://www.alzforum.org/mutations/app)	System used to generate mutation
T714I Austrian	T682I	T638I	Single AA substitution at γ-sec cleavage site. MOA = 36.3yrs. Increased $A\beta_{42/40}$	CRISPR and TALEN

112

			ratio in HEK293 cells, mouse primary neurons and Plasma [11, 19].	
E693G Arctic	E661G	E617G	Single AA substitution between α-sec and γ-sec cleavage sites, closer to the α-sec site. MOA = 59.7yrs. Decreased $A\beta_{42/40}$ ratio in HEK293 cells [13].	CRISPR
E682K Leuven	E650K	N/A	Single AA substitution at the B'-sec cleavage site. MOA = 61yrs. Increased $A\beta_{42/40}$ ratio in CHO cells and mouse primary neurons [12].	CRISPR
V717I London	V685I	V641I	Single AA substitution at γ-sec cleavage site. MOA = 53.2yrs. Increased $A\beta_{42/40}$ ratio in Chinese hampster ovary cells, HEK293 cells and mouse primary neurons [11, 19, 20].	TALEN

Notes - TM = transmembrane domain, MOA = mean onset age of fAD. AA = amino acid. yrs = years. All work will be done with the Tuebingen (TU) strain of fish since this was the strain used to generate the zebrafish genome sequence.

Design of CRISPR sgRNAs for analysis

Transcription activator-like effector nucleases (TALENs) are a popular tool used for editing genomes by the induction of DSBs. Recently, the CRISPR/Cas9 system has grown in popularity for its ability to easily generate mutations of interest in many organisms. One disadvantage with

this system is the additional constraints when designing and using CRISPRs. For Cas9 to successfully recognise a target sequence, a protospacer adjacent motif (PAM) sequence is required 3' of the sgRNA and upstream of the crRNA binding region [7]. The use of the T7 promoter to transcribe gRNAs means that the sequence "GG" is required at the 5' end of the transcript, which reduces the range of targetable sequences [7]. Taken together these constraints would mean that any sequence of the form 5' GG-N(18)-NGG 3' is available for targeting by CRISPR/Cas9, occurring once in every 128bps of random DNA sequence [7]. However, a study by Hwang *et al* (2013) found that this targeting range could be expanded to once in every 8bps by relaxing the T7 promoter rule [21]. There are currently extensive options for designing and preparing CRISPR sgRNAs for generating mutations. During our design phase, we found a method described by Hwang *et al* (2013) which boasted a greater than 10% somatic mutation rate in the zebrafish [7].

The online "ZiFit targeter" tool was used to identify potential CRISPR/Cas target sites for each of the selected mutations (http://zifit.partners.org/ZiFiT/ChoiceMenu.aspx) [22, 23]. This targeting software was originally developed by the Zinc Finger Consortium for Zinc Finger engineering and was later updated to identify sequences that fit the standard CRISPR/Cas9 constraints discussed above. When using the "CRISPR/Cas9 nuclease" link of the software, the length (in nucleotides) of target site of "20" and "T7 promoter" options were selected. A region of approximately 2,000 nucleotides was run through the software to identify target sites close to the codon of interest (supplementary data file S1).

ZiFit identified oligonucleotides to generate an Arctic-like mutation in Appb but could not identify oligos for generating an Austrian-like mutation in Appb or a Leuven-like mutation in Appa. The previously mentioned study by Hwang *et al* (2013) discussed and demonstrated a method to overcome the 5' requirement of the software. Induction of site specific mutations was

accomplished by the addition of two guanines ("GG") to the 5' end of sgRNA sequences that are not complementary to genomic DNA target sequences [21]. This method was adopted when designing the Austrian-like Appb and Leuven-like Appa sgRNAs (supplementary data file S2, Table S1).

When designing the initial CRISPR sgRNAs in 2014, there was a limited number of published studies using this technology in the zebrafish. As a result, we failed in our attempts to design an sgRNA to target the fAD "London" missense mutation in *appb*. To overcome this, the London mutation position was targeted using TALEN technology, which has different targeting constraints. Approximately 200 nucleotides up- and down-stream of the London *appb* TALEN target sequence was supplied to ZGeneBio, who designed and constructed appropriate TALEN arms (Figure 1, C) (http://www.zgenebio.com/).

To corroborate the sgRNA target sites suggested by the "ZiFit" targeting software, sequences around the sgRNA sites were analysed using the online CRISPRscan software [http://www.crisprscan.org/ [24]]. CRISPRscan generates a score calculated using a mathematical model that predicts mutagenic activity based on the sgRNA target sequence [24]. Sequences containing approximately 50 nucleotides either side of the already designed sgRNAs (supplementary data file S2, Table S1) were submitted to the online software. Moreno-Mateos *et al* (2015) found that sgRNAs with a CRISPRscan score above 55 were able to efficiently cleave the DNA at the position of interest whereas sgRNAs that scored above 70 were highly efficient [24]. CRISPRscan returned sgRNA sequences that matched our previous designs and the scores of these are presented in supplementary data file S2, Table S1. Arctic *appb* sgRNA returned a score of 61, suggesting this should cleave efficiently. The *appa* and *appb* start site sgRNAs were also identified using CRISPRscan. The former returned a score of 70 suggesting it should be highly

efficient, while the latter returned a score of 66 suggesting it should also efficiently cleave the DNA. The Austrian *appb* sgRNA was not identified using CRISPRscan. This is perhaps unsurprising as this sgRNA required the addition of 5'GG in order to meet design requirements. Multiple attempts were made to identify the predicted Austrian sgRNA, including using 50 nucleotides of genomic DNA sequence either side of the sgRNA sequence without 5'GG, sgRNA sequence with additional 5'GG and also using 50 nucleotides of cDNA (exons only) either side of the sgRNA sequence equivalent from the NCBI database (one nucleotide difference). None of the attempts successfully identified the Austrian sgRNA as an option. Similarly, CRISPRscan was unable to identify the Leuven *appa* sgRNA designed through ZiFit. However, the program did return an sgRNA that closely resembled ours, with a score of 34, suggesting cleavage by this similar sgRNA would be inefficient.

The sgRNAs discussed above were designed to be constructed and transcribed via a method previously described in zebrafish (The "ADDGene method") [7]. We also designed an additional *appa* start site sgRNA, Austrian *appb* sgRNA and an sgRNA to target the Austrian mutation position in *appa*, to be constructed via an adapted "PCR method", rather than the method mentioned above [25] (supplementary data file S2, Table S2). When analysed with CRISPRscan, the *appa* start site sgRNA returned a score of 81 and the Austrian *appa* sgRNA returned a score of 66. Unlike the Austrian *appb* sgRNA designed for the first method [7], the Austrian *appb* designed at an alternative position returned a score of 79, suggesting it should be highly efficient in cleaving the DNA.

Identifying successful NHEJ events induced by CRISPR/TALENs

In order to generate mutations by non-homologous end joining (NHEJ) or to engineer mutations at a specific site by HDR, there must first be induction of a double-stranded break in the DNA at the position of interest. The cell has two major mechanisms for repairing DSBs, NHEJ and HDR. NHEJ re-connects DNA breaks, with nucleotides around the break point gained or lost as a part of this process. This often results in frameshifts that introduce a premature stop codon around the target site [8]. Therefore, NHEJ is often exploited in genome editing as a means of destroying gene function. In contrast, HDR allows for precise changes to be made using a template DNA sequence that is homologous to the sequence around the DSB [9].

Table 3.2 Successful or unsuccessful generation of mutations by CRISPRs or TALENs				
App isoform	**Technology**	**Intended mutation**	**sgRNA construction method**	**Mutations detected**
appb	TALEN	London/Austrian	Addgene	-
appa	CRISPR	Leuven	Addgene	-
appb	CRISPR	Arctic	Addgene	+
appb	CRISPR	start site (Geneart)	Addgene	+
appa	CRISPR	start site (Geneart)	Addgene	-
appa	CRISPR	start site (PCR)	PCR	-
appa	CRISPR	Austrian (PCR)	PCR	+
appb	CRISPR	Austrian (PCR)	PCR	-

Note: + indicates successful double-stranded break induction, - indicates that genome editing technology was unsuccessful in generating a double-stranded break. The "sgRNA construction method" refers to the methods used to generate sgRNAs for microinjection.

Each of the sgRNAs (presented in supplementary data file S2, Tables S1 and S2) along with the TALEN pair were first tested for whether they could successfully introduce a mutation via NHEJ. Mutations induced by injected CRISPR and TALENs were detected using a T7 endonuclease-based assay and gel electrophoresis. DNA strands of a region approximately 500-1000bp (spanning the sites of interest) containing either; 1) insertions or deletions (indels) from NHEJ following successful cleavage by CRISPR/Cas9 or TALENs, or, 2) uncleaved WT strands, were denatured and re-annealed at random. Non-perfectly matched (i.e. heteroduplex) DNA was cleaved by T7 endonuclease and visualised (Figure 3.2)

Analyses using PCR and T7 endonuclease revealed that the TALEN pair designed to target the Austrian and London mutation positions in zebrafish *appb* was unable to generate a mutation. Similarly, sgRNAs designed to target the Leuven mutation equivalent site in *appa*, Austrian mutation equivalent site in *appb* and the start codon of *appa* were unable to introduce mutations. sgRNAs designed to target the Arctic mutation equivalent site in *appb*, the Austrian mutation equivalent site in *appa* and the start codon of *appb* were successful in generating mutations. This data is summarised in Table 3.2.
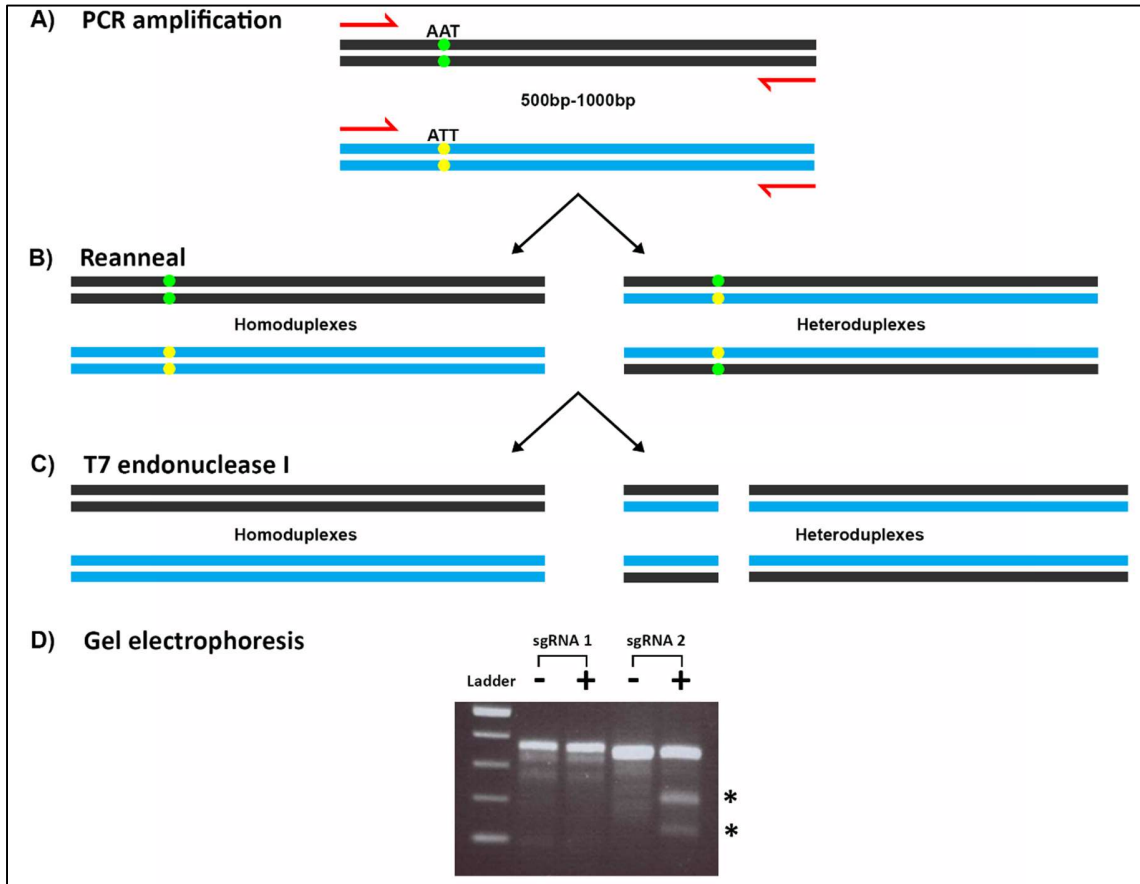
Figure 3.2 *Schematic depiction of the T7 endonuclease I mutation detection assay. (A) 500-1000bp of the genomic DNA surrounding the TALEN/CRSIPR target site is amplified via PCR. (B) The PCR amplified fragments are denatured and allowed to reanneal slowly, resulting in both homoduplex and heteroduplex formation. (C) T7 endonuclease 1 enzyme is applied, cleaving heteroduplexes and leaving homoduplexes uncleaved. (D) The T7 treated mix is separated on agarose gel. In this example, two sgRNAs have been tested for their ability to generate double-stranded breaks. An identical banding pattern is observed for both untreated and T7 endonuclease treated fragments for the example sgRNA 1, suggesting no cleavage (all homoduplexes). The * next to the T7 + lane of sgRNA2 represent alternative bands resulting from heteroduplex cleavage (suggesting a possible mutation). From top of ladder lane (in descending order), ladder sizes for each visible band are 500bp, 400bp, 300bp, 200bp and 100bp.*

Sequencing the APPa 5'UTR region

In order to test the efficacy of the sgRNA targeting just downstream of the start codon of *appa*, primers that could detect DSBs at this site needed to be designed. However, the 5'UTR region identified from database entries (ENSEMBL, NCBI and USCS), was incomplete. When searching through the available sequence for appropriate primer sites, highly repetitive regions and blocks of undetermined sequence were observed. Therefore, we needed to characterise the genomic sequence upstream of the start codon in order to be able to design primers that would bind in this region.

Using the available sequence, primers were first designed to amplify ~1000bp upstream of the start codon (supplementary data file S2, Table S3). PCR amplification of the region upstream of the start codon was performed on multiple Tübingen-strain fish from different families in order to identify common polymorphisms. The PCR-amplified strands were sent for Sanger sequencing (Australian Genome Research Facility, SA, AUS). The multiple sequences were combined to generate a consensus sequence. The sequence of genomic DNA upstream of the start codon (also including Exon 1 and part of Intron 1) can be found in supplementary data file S3.

Attempting to engineer precise fAD mutations into appa *and* appb

To engineer a specific mutation using genome editing technologies the appropriate system can be co-injected with an oligonucleotide template encoding the desired mutation [21]. The sequence of this single-stranded oligonucleotide (ssON) template can then be repaired into the genome via HDR. In human *APP*, the Austrian fAD mutation results from a single substitution of ACA to ATA [11]. As can be observed on the Aβ peptide region alignment (Figure 3.1, B, Austrian

residue), the Threonine (T) residue at position 714 in human APP695 is conserved in zebrafish Appa, while its nucleotide sequence at the equivalent codon is ACC. We designed a custom ssON, where the zebrafish *appa* nucleotide sequence was substituted from ACC to ATA, to emulate the substitution that occurs in the human *APP* Austrian fAD mutant nucleotide sequence (supplementary data file S2, Table S4). In order to increase the likelihood of detecting successful incorporation of the Austrian-like Appa mutation, silent substitutions were made in some codons so that mutation specific primers would be less likely to bind to wildtype sequence (supplementary data file S2, Table S4). The Arctic fAD mutation results from substitution of GAA to GGA in human *APP* [26]. The glutamic acid (E) residue at position 693 in human APP695 is conserved in zebrafish Appb (Figure 3.1, B). However, the nucleotide sequence corresponding to the Arctic residue in zebrafish *appb* (GAG) is not conserved with human *APP* (GAA). Therefore, we designed a custom ssON where the zebrafish *appb* nucleotide sequence was substituted from GAG to GGA, to emulate the substitution that occurs in the human *APP* Arctic fAD mutant nucleotide sequence (supplementary data file S2, Table S4).

Oligonucleotides coding the Austrian-like *appa* and Arctic-like *appb* fAD missense mutations were co-injected with the corresponding sgRNA that had been shown to successfully induce mutations via NHEJ. PCR was performed on groups of 10 embryos at 24hpf, using primers that had been designed to bind specifically DNA sequences containing the incorporated oligonucleotides (supplementary data file S2, Table S5). Neither the Austrian-like *appa* nor Arctic-like *appb* oligonucleotides were detected via PCR, suggesting they were not incorporated by HDR.

3.3 Conclusion and Future Directions

In this study, we designed and tested CRISPR/Cas9 sgRNAs and a TALEN for their ability to successfully generate mutations in the zebrafish *app* genes. The single TALEN designed to target the Austrian/London mutation positions in *appb* was unsuccessful. However, 1 of the 4 sgRNAs designed to target sites within *appa* (Austrian) and 2 of the 3 sgRNAs designed to target sites within *appb* (start codon and Arctic) were able to successfully generate DSBs. Single-stranded oligonucleotides coding for the Arctic-like and Austrian-like fAD missense mutations were co-injected with the successful sgRNAs. However, we could not detect incorporation of the oligonucleotides into the *appa* or *appb* genes by PCR. Therefore, we were unable to introduce the Artic-like and Austrian-like fAD missense mutations into the zebrafish genome.

During our sgRNA design stage, Hwang *et al* (2013) tested the efficiency of sgRNAs in zebrafish and reported successful generation of mutations in 8 of 10 sgRNAs (targeting 10 different endogenous genes) tested, with those 8 successful sgRNAs having high mutational frequencies [7]. However, our rate of successful double-stranded break induction in the *appa* and *appb* genes in zebrafish was much lower than that. Only 3 of the 7 CRISPR sgRNAs tested in this study were able to generate mutations sufficiently to be detected by our methods. Recent studies suggest that local chromatin structure may have an effect on the efficiency of genome editing by CRISPR/Cas, as there may be limited access to certain sites in the genome that are more tightly packaged when editing is attempted [6, 27, 28]. It is also possible that this limited activity may result from the differences between eukaryotic chromatin structures and DNA packaging in the prokaryotes, from which the CRSIPR system was originally adapted, as Cas9 was derived from a prokaryotic enzyme [28]. In a study investigating the effects of chromatin on efficient Cas9 mutagenesis, Uusi-Mäkelä *et al* (2018) observed that, while a lack of chromatin openness reduced the efficiency of DSB

induction, lower gene expression levels were also associated with the inability of CRISPR/Cas9 to effect mutagenesis [28]. Interestingly, their findings also suggested that genes that become active at the maternal to zygotic transition (MZT) are more accessible for Cas9 and therefore undergo more efficient mutagenesis. In the zebrafish embryo, MZT spans the period of just after fertilisation to the shield stage of gastrulation (i.e 6hpf) [29]. When looking at the expression levels of both *appa* and *appb* during early developmental stages using the EMBL-EBI "Expression atlas" (https://www.ebi.ac.uk/gxa/home, [30]) we observed that both genes had low expression during the recorded stages of early development aligning with MZT (blastula 128-cell to gastrula shield) (supplementary data file S4, Figure S1). This low expression during MZT may explain our limited success in targeting these genes with CRISPR sgRNAs.

Although two of the CRISPR sgRNAs targeting fAD missense mutation equivalent sites in zebrafish *appa* and *appb* in our study appeared to generate mutations (and, hence, could generate DSBs), sequences from the corresponding oligonucleotide template donor DNAs carrying the mutations of interest were unable to be incorporated. This could be due to several factors. One such factor is the nature of the HDR process, which has been observed to be inefficient in the zebrafish, averaging only 3-4% success when attempting to incorporate specific mutations by providing an ssON template [31]. Instead, cells generally prefer to repair breaks in their DNA by NHEJ, perhaps dues to the increased activity of NHEJ over HDR in the developing zebrafish embryo [32]. Options for overcoming these factors include inhibiting NHEJ and stimulating HDR, both of which are discussed in more detail below. Another factor is the optimal length of the HDR template molecule, which is still debated, as is the effect of this on HDR efficiency [31, 33].

Enhancing DSB formation

While CRISPR/Cas9 mediated genome editing has many benefits over alternative methods, there are also challenges. Site-specific genome editing can be restricted in some cases due to the absence of an NGG (PAM) at the precise position being targeted. When attempting to precisely engineer a mutation via HDR, the most efficient position at which to generate the DSB is within 10-20bp of the desired change. One group has recently attempted to overcome these challenges by engineering Cas9 PAMs with novel specificities [34]. A previous attempt to change the PAM recognition site had varying results. Anders *et al* (2014) mutated the two residues that contact the guanine nucleotides of the NGG to R1333Q/R1335Q respectively. However, this mutant Cas9 was unable to cleave a site at the newly expected NAA PAM *in vitro* [35]. It was later confirmed by Kleinstiver *et al* (2015) that this variant is unable to cleave NAA target sites efficiently [34]. They also tested both variants separately. However, the expected NAG and NGA PAM's were also ineffective [34]. They investigated the efficiency of two smaller Cas9 orthologues from different bacteria, termed St1Cas9 and SaCas9, which have alternative PAM specificities[1] with some success [34]. Feng *et al* (2016) have now shown that both SaCas9 and its variant, KKH SaCas9 (that recognises a relaxed NNNRRT PAM) can be utilised in zebrafish with a mutational frequency of up to 90% [36]. Such studies highlight the potential for finding alternatives for targeting challenging sites in zebrafish in the future.

Another recent development in the field of genome editing is the Cpf1-containing class 2 CRISPR system (recently renamed Cas12a [37]), which has been identified as a potential alternative to the commonly used type II CRISPR/Cas9 systems [38]. Some of the advantages of this new system

---

[1] Alternate PAM specificities are NNAGAA, NNGGAA, NNAGGA, NNAGCA, NNACCA, NNATAA, NNAAAA, NNGGGA (i.e. NNRRRA) for St1Cas9 and NNGGGT, NNGAAT, NNGAGT (i.e. NNGRRT) for SaCas9

are as follows: 1) being a single crRNA-guided endonuclease, Cas12a does not require tracrRNA to processes crRNA, potentially simplifying design and delivery of this system. 2) DSBs generated by Cas12a result in a staggered 5' overhang. This is advantageous compared to the blunt ends generated by Cas9 as it enables the design of DNA templates that can be inserted directionally by NHEJ rather than via HDR mechanisms. 3) The PAM site for Cas12a is T rich (5'-TTN) as opposed to the G rich Cas9 PAM (5'-GGN). This opens possibilities for targeting regions of the genome that are richer in A/T. These are only a few of the many potential advantages of this system [38]. Two Cas12a protein variants have been successfully used in mammals previously for genome editing, AsCas12a and LbCas12a [39]. One of the benefits of these variants observed thus far has been the extension of target recognition to T-rich sequences such as non-coding RNAs, 5'UTR and 3'UTR etc (PAM 5' TTTV), among other things [40]. Despite the benefits of CRISPR-Cas12a, previous studies reported reduced activity in *Drosophila* and plants [41-44]. Moreno-Mateos *et al* (2017) performed a study in which they characterised and optimised CRISPR-Cas12a for zebrafish and *Xenopus tropicalis* [39]. Their study found that the amount of Cas12a activity both *in vitro* and *in vivo* was temperature sensitive, which could help to explain the previously reported reduction of activity in some organisms [41, 44]. Interestingly, they found that incubating injected embryos at 34°C for up to 48hpf increases the mutagenic activity of Cas12a [39]. More importantly, their study also found that co-injection of CRISPR-LbCas12a along with single stranded donor DNA significantly increased the efficiency of homology directed repair (HDR) in zebrafish [39]. A recent study by Liu *et al* (2019) further investigated the use of Cas12a in the zebrafish by modifying crRNA in the direct repeat region and stem loop [45]. In the case of standard pre-crRNA processing, mature crRNAs are produced with a shorter Direct Repeat (DR) sequence. Liu *et al* (2019) explored the effects of modifying crRNA to have the pre-processing

125

full-length Direct Repeat (DRf-crRNA) and observed higher editing efficiency compared to mature crRNA produced by the native system. Interestingly, when they further modified DRf-crRNA by substituting two of the A:U pairs to G:C within the stem loop region of the DR (DRf-GC@13-crRNA) they found an even higher editing efficiency than DRf-crRNA [45]. Their study also investigated the use of another Cas12a variant, FnoCas12a, which has most recently been adapted for use in mammalian cells. FnoCas12a recognises a TTN PAM, which differs from the TTTV PAM of the previously discussed LbCas12a and AsCas12a. Their successful use of FnoCas12a in the zebrafish adds yet another option for gene editing, offering greater targeting flexibility and making it an attractive option for future studies. Cas12a opens new possibilities for targeting the zebrafish *appa* and *appb* sites that were unsuccessfully targeted in this study.

Enhancing HDR

While successful incorporation of the single stranded oligonucleotide template DNA carrying the mutation of interest was not achieved by co-injection in this study, it is now possible to suppress the NHEJ pathway by treating zebrafish embryos with the NHEJ inhibitor SCR7 [19]. Zhang *et al* (2018) saw a significant improvement in HDR mutation incorporation efficiency in zebrafish when treating with SCR7 [19]. Very recently, Askoy *et al* (2019) investigated a different NHEJ inhibitor, NU7441. They found that NU7441 improved their CRISPR-mediated genome editing rate by up to 13.4-fold [46]. Interestingly (and conversely to the study by Zhang *et al*), they saw minimal or no increase in HDR efficiency with SCR7. They suggested several possible explanations for this observation and ultimately concluded that successful HDR stimulation using SCR7 is likely context-specific [46].

The HDR pathway can also be stimulated by treatment with RS-1 [46]. Zhang *et al* (2018) also tested RS-1 in the zebrafish and, as with SCR7, found significant improvement in HDR efficiency [19]. Most importantly, their study found that a combination of these treatments greatly increases the efficiency of HDR, in the case of their target mutation, to around 74% [19]. Applying this combination of treatments to the sgRNAs tested in this study may enable successful generation of zebrafish carrying the human *APP* fAD Arctic mutation equivalent in *appb* and/or the human *APP* fAD Austrian mutation equivalent in *appa*.

Finally, other research attempting to improve HDR efficiency used an sgRNA targeting exon 6 of the zebrafish *albino* gene to attempt to revert the mutation. They then tested donor DNA templates of various lengths for HDR, finding that only the larger DNA fragments (between 986 bp and 3.8 kb) gave efficient sequence incorporation [20]. In order to overcome the toxicity sometimes observed when co-injecting large linear DNA fragments, they designed circular plasmid donor DNAs. These circular donors comprised the linear template DNA fragment flanked by two CRISPR target sites that allow the plasmid to be linearised within cells [20]. By employing this strategy, they achieved 46% efficiency in detection of reversion in mosaic fish, although only 10% of 3-day-old larvae reached adulthood. In this study the linear donor DNA templates tested were only 50 bp in length. At the time of design of our study there was no consensus on the most appropriate template length for HDR repair. The large circular donor DNA method could easily be adapted to carry the Arctic-like *appb* and Austrian-like *appa* mutations for delivery into the zebrafish in hopes of improving the chance of successful incorporation.

Conclusions

Generating *APP* fAD mutations in endogenous zebrafish *appa* and *appb* is an important step towards allowing us to make more meaningful transcriptomic analyses of fAD mutations in the *APP* gene. This study has identified and tested sgRNAs targeting a variety of *APP* fAD mutation loci for their ability generate DSBs. Of the 7 sgRNAs tested, 4 were unable to cleave the DNA at their target loci. Similarly, the single TALEN tested in this study failed to generate a DSB at the Austrian/London position in zebrafish *appb*. 3 sgRNAs with the ability to generate DSBs were also identified. However, introduction of the desired mutation through HDR was not achieved for the 2 sgRNAs targeting fAD *APP* mutation sites, namely the fAD Arctic (*appb*) and Austrian (*appa*) mutations. The combined use of SCR7 (or NU7441) and RS-1 treatments, after co-injection of sgRNAs and oligonucleotide templates, may allow us to successfully inhibit NHEJ, stimulate the HDR machinery and inevitably introduce these fAD *APP* mutations. The third successful sgRNA, targeting the *appb* start site, was used to generate an *appb* null mutation which will be discussed in detail in the following chapter (Chapter 4).

Future attempts to generate DSBs at the Austrian/London sites in *appb* as well as the Leuven site in *appa* should make use of the recently improved understanding of site accessibility and should be attempted with new technology, such as Cas12a. The collection of *APP* fAD mutation models produced would be an immensely valuable resource for further research on APP activities.

3.4 Methods

Ethics

This work was conducted under the auspices of the Animal Ethics Committee of the University of Adelaide and in accordance with EC Directive 86/609/EEC for animal experiments and the Uniform Requirements for manuscripts submitted to Biomedical journals.

Generation of TALEN coding sequences

TALEN coding sequences were designed by, and purchased from, Zgenebio (http://www.zgenebio.com/). The DNA binding sites for the TALEN pair targeting the London mutation equivalent site within exon 16 of zebrafish *appb* were (5' to 3'): left, GGGGGGCGTGGTCATCG and right, GAGCATAACCAGAGTGA (Figure 3.3).

Generation of sgRNA coding sequences

Oligonucleotide templates for each sgRNA to be constructed (supplementary data file S2, Table S1) were provided by SIGMA Aldrich Australia (Castle Hill, NSW). Due to the rapid development of CRISPR technology, two different methods were used to construct sgRNAs. In the first method ("ADDGene"), oligonucleotides encoding sgRNAs from initial designs (supplementary data file S2, Table S1) were cloned into the pDR274 plasmid (https://www.addgene.org/42250/) as previously described in [7]. In the second "PCR" based method (adapted from a method developed for *Drosophila*), forward primers for the template synthesis PCR consisted of target sites for sgRNAs (that were selected based on their CRISPRscan score) flanked by a T7 polymerase binding site and a region complementary to a common reverse primer [25]. These primers,

presented in supplementary data file S2, Table S2, were annealed to the common reverse primer with sequence: 5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAA CTTGCTATTTCTAGCTCTAAAAC-3', and amplified via PCR using Phusion HF DNA polymerase (New England BioLabs, Ipswich, MA, USA) with the following cycling conditions; 98°C for 30 seconds, 34 rounds of 98°C for 10 sec, 60°C for 30 sec, 72°C for 15 sec, then 72°C for 10 mins. The resulting template was purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, Wisconsin, USA, A9281).

In-vitro mRNA synthesis

The coding sequences of the TALENs were provided in the pZGB4 vector for *in-vitro* mRNA synthesis. TALEN-containing plasmids were linearised with *Not* I and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, Wisconsin, USA, A9281). Purified linearised DNA was used as a template for *in-vitro* mRNA synthesis using the mMESSAGE mMACHINE T7 transcription kit (Thermo Fisher, Waltham, USA) as per the manufacturer's instructions.

The sgRNA-containing pDR274 plasmid was linearised with *HindIII* and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, Wisconsin, USA, A9281). Purified linearised DNA or purified PCR reaction product was used as a template for *in-vitro* mRNA synthesis using the HiScribe™ T7 High Yield RNA Synthesis Kit (New England BioLabs, Ipswich, MA, USA) as per the manufacturer's instructions. The mRNA was precipitated using

ammonium acetate as described in the "MAXIscript" kit procedure (Thermo Fisher, Waltham, USA).

Microinjection of zebrafish embryos

sgRNA and Cas9 mRNA or Cas9 nuclease (protein) were co-injected into one-cell stage zebrafish embryos. Cas9 mRNA was synthesised using the mMESSAGE mMACHINE T7 transcription kit (Thermo Fisher, Waltham, USA) as per manufactures instructions. Cas9 protein was purchased from ToolGen (ToolGen, Geumcheon-gu, Seoul, South Korea).

In the first attempt to generate the Arctic, Leuven and Austrian mutations, each embryo was microinjected with a ~3 nl solution containing 12.5ng/µl of sgRNA and 300ng/µl of Cas9 mRNA **or** 12.5ng/µl of sgRNA and 500ng/µl of Cas9 nuclease (protein). If a single-stranded oligonucleotide was also co-injected with sgRNA and Cas9 (Arctic), it was at a final concentration of 50ng/µl. In the second attempt, embryos co-injected with approximately 3 nl of a solution containing 200ng/µl of sgRNA and 1.5µg/µl of Cas9 nuclease (protein) (Invitrogen, Carlsbad, California, USA). For TALEN injections, embryos were microinjected with an approximately 3 nl mixture of 100ng/µl of left and right TALEN mRNA. Embryos were inspected using a light microscope at 24 hpf and only embryos that had developed normally were selected for analysis of genome editing.

Genomic DNA extraction from zebrafish embryos and tail clips

A random selection of 10-25 injected embryos was collected at 24 hpf. Once grown into adults (~3 months old), a small section of the dorsal fin (~3 mm) was removed with a sterile scalpel blade under anaesthesia (Tricaine solution, 1.68 μg/mL).

Embryos or tail fin tissue were placed in 50μl of a 17mg/ml solution of Proteinase K (Roche, Basel, Switzerland) with 1XTE and then incubated at 55-65°C until digestion was complete (2-4 hours). The solution was then placed at 95°C for 5 mins to inactivate the Proteinase K. This was then centrifuged at maximum speed for 2 mins to sediment cellular debris. The supernatant was transferred into a clean microfuge tube ready for subsequent PCR [47].

Mutation detection assays

Mutation due to NHEJ of double-stranded DNA breaks was detected by either the GeneArt Genomic Cleavage Detection Kit or the T7 endonuclease I method adapted from Reyon *et al* (2012) [48]. Targeted genomic loci were amplified from the pooled genomic DNA of ten injected embryos using primers designed to anneal ~500 base pairs upstream and downstream from the expected cut site using the GeneArt Kit (Thermo Fisher, Waltham, USA) as per manufacturer's instructions. In the T7 endonuclease I method, loci were amplified with GoTaq polymerase (New England Biolabs, Ipswich, Massachusetts, USA) and the resulting DNA was re-annealed under the following conditions; 95°C for 5 mins then ramped down from 95°C-85°C at a rate of -2°C/sec, followed by ramping down from 85°C-25°C at a rate of -0.1°C/sec. After reannealing, 1μl of T7 endonuclease I (New England Biolabs, Ipswich, Massachusetts, USA) was added to the PCR

products and the entire mixture was incubated at 37°C for 15 mins. The mixture was then analysed by electrophoresis on a 2% agarose gel that was run at 90V for 1hr.

Sanger sequencing

The genomic region surrounding the Arctic mutation sgRNA target site was PCR amplified using "Arctic and Austrian *appb* F" and "Arctic and Austrian appb R" primers (supplementary data file S2, Table S5). The purified *appb* Arctic genomic DNA fragments were then used for Sanger sequencing with the "Arctic *appa*" sequencing primer (supplementary data file S2, Table S5) (Australian Genome Research Facility, SA, AUS). The genomic region surrounding the *appa* 5'UTR/Exon 1 region was PCR amplified using *appa* 5'UTR amplification forward and reverse primers (supplementary data file S2, Table S3). The *appa* 5'UTR PCR fragments were cloned into pGEM-T Easy (Promega, Madison, Wisconsin, USA) and sequenced with M13 forward and reverse primers (Australian Genome Research Facility, SA, AUS).

Multiple sequence alignment to identify an appa *consensus sequence upstream of the start codon*

The *appa* nucleotide sequences generated by Sanger sequencing were aligned to each other along with the *appa* genomic DNA sequence from the zebrafish SANGER database (SSS10789, https://www.sanger.ac.uk/) using the online tool "MUSCLE" (https://www.ebi.ac.uk/Tools/msa/muscle/). Differences between the Tübingen fish sequenced for this study and the available database sequence were identified by eye and a consensus sequence was compiled (supplementary data file S3).

3.5 References

1. Musa, A., H. Lehrach, and V.A. Russo, *Distinct expression patterns of two zebrafish homologues of the human APP gene during embryonic development.* Dev Genes Evol, 2001. **211**(11): p. 563-7.
2. Joshi, P., et al., *Amyloid precursor protein is required for convergent-extension movements during Zebrafish development.* Dev Biol, 2009. **335**(1): p. 1-11.
3. Pu, Y.Z., et al., *Generation of Alzheimer's Disease Transgenic Zebrafish Expressing Human APP Mutation Under Control of Zebrafish appb Promotor.* Curr Alzheimer Res, 2017. **14**(6): p. 668-679.
4. McCutcheon, V., et al., *A Novel Model of Traumatic Brain Injury in Adult Zebrafish Demonstrates Response to Injury and Treatment Comparable with Mammalian Models.* J Neurotrauma, 2017. **34**(7): p. 1382-1393.
5. Hargis, K.E. and E.M. Blalock, *Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us?* Behav Brain Res, 2017. **322**(Pt B): p. 311-328.
6. Chen, Y., et al., *Using local chromatin structure to improve CRISPR/Cas9 efficiency in zebrafish.* PLoS One, 2017. **12**(8): p. e0182528.
7. Hwang, W.Y., et al., *Efficient genome editing in zebrafish using a CRISPR-Cas system.* Nat Biotechnol, 2013. **31**(3): p. 227-9.
8. Pannunzio, N.R., G. Watanabe, and M.R. Lieber, *Nonhomologous DNA end-joining for repair of DNA double-strand breaks.* J Biol Chem, 2018. **293**(27): p. 10512-10523.
9. Wright, W.D., S.S. Shah, and W.D. Heyer, *Homologous recombination and the repair of DNA double-strand breaks.* J Biol Chem, 2018. **293**(27): p. 10524-10535.
10. Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, *ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.* Trends Biotechnol, 2013. **31**(7): p. 397-405.
11. Kumar-Singh, S., et al., *Nonfibrillar diffuse amyloid deposition due to a gamma(42)-secretase site mutation points to an essential role for N-truncated A beta(42) in Alzheimer's disease.* Hum Mol Genet, 2000. **9**(18): p. 2589-98.
12. Zhou, L., et al., *Amyloid precursor protein mutation E682K at the alternative beta-secretase cleavage beta'-site increases Abeta generation.* EMBO Mol Med, 2011. **3**(5): p. 291-302.
13. Nilsberth, C., et al., *The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Abeta protofibril formation.* Nat Neurosci, 2001. **4**(9): p. 887-93.
14. Goate, A., et al., *Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease.* Nature, 1991. **349**(6311): p. 704-706.
15. Newman, M., E. Ebrahimie, and M. Lardelli, *Using the zebrafish model for Alzheimer's disease research.* Front Genet, 2014. **5**: p. 189.
16. Talarico, G., et al., *The London APP mutation (Val717Ile) associated with early shifting abilities and behavioral changes in two Italian families with early-onset Alzheimer's disease.* Dement Geriatr Cogn Disord, 2010. **29**(6): p. 484-90.
17. Kaden, D., et al., *Novel APP/Abeta mutation K16N produces highly toxic heteromeric Abeta oligomers.* EMBO Mol Med, 2012. **4**(7): p. 647-59.
18. Sahlin, C., et al., *The Arctic Alzheimer mutation favors intracellular amyloid-beta production by making amyloid precursor protein less available to alpha-secretase.* J Neurochem, 2007. **101**(3): p. 854-62.
19. Zhang, Y., Z. Zhang, and W. Ge, *An Efficient Platform for Generating Somatic Point Mutations with Germline Transmission in the Zebrafish by CRISPR/Cas9-mediated Gene Editing.* J Biol Chem, 2018.
20. Irion, U., J. Krauss, and C. Nusslein-Volhard, *Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system.* Development, 2014. **141**(24): p. 4827-30.
21. Hwang, W.Y., et al., *Heritable and precise zebrafish genome editing using a CRISPR-Cas system.* PLoS One, 2013. **8**(7): p. e68708.

22.   Sander, J.D., et al., *ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W462-8.

23.   Sander, J.D., et al., *Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W599-605.

24.   Moreno-Mateos, M.A., et al., *CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo.* Nat Methods, 2015. **12**(10): p. 982-8.

25.   Bassett, A. and J.L. Liu, *CRISPR/Cas9 mediated genome engineering in Drosophila.* Methods, 2014.

26.   Lord, A., et al., *The Arctic Alzheimer mutation facilitates early intraneuronal Abeta aggregation and senile plaque formation in transgenic mice.* Neurobiol Aging, 2006. **27**(1): p. 67-77.

27.   Daer, R.M., et al., *The Impact of Chromatin Dynamics on Cas9-Mediated Genome Editing in Human Cells.* ACS Synth Biol, 2017. **6**(3): p. 428-438.

28.   Uusi-Makela, M.I.E., et al., *Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (Danio rerio).* PLoS One, 2018. **13**(4): p. e0196238.

29.   Langley, A.R., et al., *New insights into the maternal to zygotic transition.* Development, 2014. **141**(20): p. 3834-41.

30.   Petryszak, R., et al., *Expression Atlas update--an integrated database of gene and protein expression in humans, animals and plants.* Nucleic Acids Res, 2016. **44**(D1): p. D746-52.

31.   Cornet, C., V. Di Donato, and J. Terriente, *Combining Zebrafish and CRISPR/Cas9: Toward a More Efficient Drug Discovery Pipeline.* Front Pharmacol, 2018. **9**: p. 703.

32.   Hagmann, M., et al., *Homologous recombination and DNA-end joining reactions in zygotes and early embryos of zebrafish (Danio rerio) and Drosophila melanogaster.* Biol Chem, 1998. **379**(6): p. 673-81.

33.   Boel, A., et al., *CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments.* Dis Model Mech, 2018. **11**(10).

34.   Kleinstiver, B.P., et al., *Engineered CRISPR-Cas9 nucleases with altered PAM specificities.* Nature, 2015. **523**(7561): p. 481-5.

35.   Anders, C., et al., *Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease.* Nature, 2014. **513**(7519): p. 569-73.

36.   Feng, Y., et al., *Expanding CRISPR/Cas9 Genome Editing Capacity in Zebrafish Using SaCas9.* G3 (Bethesda), 2016. **6**(8): p. 2517-21.

37.   Shmakov, S., et al., *Diversity and evolution of class 2 CRISPR-Cas systems.* Nat Rev Microbiol, 2017. **15**(3): p. 169-182.

38.   Zetsche, B., et al., *Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system.* Cell, 2015. **163**(3): p. 759-71.

39.   Moreno-Mateos, M.A., et al., *CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing.* Nat Commun, 2017. **8**(1): p. 2024.

40.   Kim, H.K., et al., *In vivo high-throughput profiling of CRISPR-Cpf1 activity.* Nat Methods, 2017. **14**(2): p. 153-159.

41.   Kim, H., et al., *CRISPR/Cpf1-mediated DNA-free plant genome editing.* Nat Commun, 2017. **8**: p. 14406.

42.   Hu, X., et al., *Targeted mutagenesis in rice using CRISPR-Cpf1 system.* J Genet Genomics, 2017. **44**(1): p. 71-73.

43.   Tang, X., et al., *A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants.* Nat Plants, 2017. **3**: p. 17103.

44.   Port, F. and S.L. Bullock, *Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs.* Nat Methods, 2016. **13**(10): p. 852-4.

45.   Liu, P., et al., *Enhanced Cas12a editing in mammalian cells and zebrafish.* Nucleic Acids Res, 2019. **47**(8): p. 4169-4180.

46.   Aksoy, Y.A., et al., *Chemical reprogramming enhances homology-directed genome editing in zebrafish embryos.* Commun Biol, 2019. **2**: p. 198.

47.   Hruscha, A. and B. Schmid, *Generation of zebrafish models by CRISPR /Cas9 genome editing.* Methods Mol Biol, 2015. **1254**: p. 341-50.

48.   Reyon, D., et al., *FLASH assembly of TALENs for high-throughput genome editing.* Nat Biotechnol, 2012. **30**(5): p. 460-5.

2.3 Supplementary Information

This section is included in the thesis as information supplementary to Chapter 3. It contains additional information not included in the main text of the manuscript.

File S1. Sequences run through Zifit targeter software to identify sgRNA sequences

(A) The ~2000bp of the zebrafish *appb* sequence that was used to identify sgRNA sequences to target the fAD "Arctic" mutation is included below:

- Yellow highlight and brackets indicate the specific residue we wished to mutate

5'-

ATGGGTATAGACCGCACGGTATTCCTGCTGTTAATGCTGACGACTTTGTCCCTCGCCATCGA

GGTGCCGTCGGATGACTCGGTGGGCTTGTTGGCGGAGCCTCAGGTGGCCATGTTCTGTGGGA

AACTCAACATGCATATCAACGTCCAGAGTGGCAAGTGGGAGCCTGATCCAACTGGCACCAA

GAGCTGCATCAGCACCAAAGAGGGCATCCTTAAATACTGCCAAGAGGTATACCCAGACCTC

CAGATCACTAATGTAGTGGAGGCCAACCAGCCTGTCAGCATCCAGAACTGGTGCAAAATGG

GTCGCCGCCAGTGCCGCAGTCACACGCACATTGTTGTTCCCTACCGTTGCCTGGTTGGGGAG

TTTGTCAGCGATGCCCTCCTCGTCCCAGATAAGTGCAAGTTCTTGCACCAGGAGCGAATGGA

CATGTGCGAGAGTCATCTACACTGGCATACAGTGGCCAAAGAGTCCTGTGGTGATCGCTCCA

TGAATCTGCATGATTATGGTATGCTGTTGCCGTGTGGAATCGACCGTTTCCGGGGCGTGGAG

TTTGTGTGCTGCCCAATGGAGGAGCAGAAAGACTTGGACAGTGAGGAGCAGGAGGAGGCTA

ACTCTGACGTGTGGTGGGGCGGTGCTGAGACTGAGTACACTGACGCCAGCGTGCTGAAAGA

ACAGGTCACAGCCAAGCCTGATCCTGCAGTGACTGAGGATGATGAGGATCTCAACAATGAG

GAAGAGGAAGTCTGGGACAACGATGAAGACGGTGACGGTGAAGATGATGAAGATGAGGAG

GACGATGATGAAGATATAATCGATGAGCAAGACACCAGTGAACAGACCTCCAACATTGCAA

TGACGACCACCACCACAACCACAACAGAGTCCATAGAGGAGGTTGTGCGAGTGCCAACCAT

GGCCCCGAGCCCTGCTGATGCGGTGGATCGTTACCTGGAAGCTCCAGGAGATATGAACGAA

CACATGCGCTTCCAGAAGGCAAAGGAGAGCCTGGAGGCCAAACACCGAGAGAAAATGTCA

GAGGTGATGAGGGAATGGGAGGAGGCCGAGAGACAGGCCAAGAACCTTCCTCGTGCTGATA

AGAAGACCATAATTCAGCGCTTCCAGGAGAAGGTGGAGTCGCTGGAGAAGGAAGCGGCTGG

AGAGAGACAGCAGCTGGTGGAAACGCACATGGCTCGAGTGGAGGCCCTACTGAATGACCGC

CGCCGTCAGGCTCTGGAAAGCTACCTGAGCTCCCTTCAGTCTGACCAGCCTCGGCCTCGGCA

GGTGCTGAATCTGTTGAAGAAGTATATCCGTGCGGAGCAGAAGGACCGTCAGCACACTCTC

AAACACTTTGAACACGTGCGAGAGGTCGATCCCAAGAAGGCTTCACAGATTCGGCCATTTGT

GATGACCCACCTACGTGTGATTGAGGAACGCATGAACCAGTCTTTGGGTTATCTCTATAAAG

TGCCTCAAGTGGCTAATGATATCCAGGATCAAGTGGCGGTGCTGGTTCAGCGTGATCAGGCT

GAGGTGACGCAGCAGCTGTCGTCTCTTCAGAGTAAGATGAGGGTCAGTTATGGGAATGATG

CCCTGATGCCGGATCTGCCCGACAGCACCACACCACTGGACAACCTTCCTCCAGAGCAGGAC

GGCCTGGGCTTCATCCATCCCGAGAGCTTTAACCAGGCCAACACTGACAACCACGTTGAACC

TGTAGATGCCCGTCCAATTCCCGAAAGGGGTTTGCCTACGAGACCCGAGATCCCAAAGGTTC

GGCTGGACATTGAGGAAAGGCACAACGCTGGCTATGATGTTCGTGACAAGAGACTGATGTT

CCTCGCG<mark>G[A]G</mark>GACATGGGCTCTAATAAGGGTGCGATCATTGGGCTGATGGTGGGGGGCGT

GGTCATCGCTACTGTCATCGTAATCACTCTGGTTATGCTCAGGAAGAAGCAGTACACCTCTA

TTCATCATGGAGTTATTGAGGTGGATGCGGCAGTGACTCCTGAAGAACGTCATCTGGCAAAG

ATGCAGCAGAATGGCTATGAAAACCCCACCTACAAGTTCTTTGAGCAAATGCAGAACTAA   -

3'

(B) The ~2000bp of the zebrafish *appb* sequence that was used to identify sgRNA sequences to target the fAD "Austrian" mutation is included below:

- Yellow highlight and brackets indicate the specific residue we wished to mutate

5'-

ATGGGTATAGACCGCACGGTATTCCTGCTGTTAATGCTGACGACTTTGTCCCTCGCCATCGA

GGTGCCGTCGGATGACTCGGTGGGCTTGTTGGCGGAGCCTCAGGTGGCCATGTTCTGTGGGA

AACTCAACATGCATATCAACGTCCAGAGTGGCAAGTGGGAGCCTGATCCAACTGGCACCAA

GAGCTGCATCAGCACCAAAGAGGGCATCCTTAAATACTGCCAAGAGGTATACCCAGACCTC

CAGATCACTAATGTAGTGGAGGCCAACCAGCCTGTCAGCATCCAGAACTGGTGCAAAATGG

GTCGCCGCCAGTGCCGCAGTCACACGCACATTGTTGTTCCCTACCGTTGCCTGGTTGGGGAG

TTTGTCAGCGATGCCCTCCTCGTCCCAGATAAGTGCAAGTTCTTGCACCAGGAGCGAATGGA

CATGTGCGAGAGTCATCTACACTGGCATACAGTGGCCAAAGAGTCCTGTGGTGATCGCTCCA

TGAATCTGCATGATTATGGTATGCTGTTGCCGTGTGGAATCGACCGTTTCCGGGGCGTGGAG

TTTGTGTGCTGCCCAATGGAGGAGCAGAAAGACTTGGACAGTGAGGAGCAGGAGGAGGCTA

ACTCTGACGTGTGGTGGGGCGGTGCTGAGACTGAGTACACTGACGCCAGCGTGCTGAAAGA

ACAGGTCACAGCCAAGCCTGATCCTGCAGTGACTGAGGATGATGAGGATCTCAACAATGAG

GAAGAGGAAGTCTGGGACAACGATGAAGACGGTGACGGTGAAGATGATGAAGATGAGGAG

GACGATGATGAAGATATAATCGATGAGCAAGACACCAGTGAACAGACCTCCAACATTGCAA

TGACGACCACCACCACAACCACAACAGAGTCCATAGAGGAGGTTGTGCGAGTGCCAACCAT

GGCCCCGAGCCCTGCTGATGCGGTGGATCGTTACCTGGAAGCTCCAGGAGATATGAACGAA

CACATGCGCTTCCAGAAGGCAAAGGAGAGCCTGGAGGCCAAACACCGAGAGAAAATGTCA

GAGGTGATGAGGGAATGGGAGGAGGCCGAGAGACAGGCCAAGAACCTTCCTCGTGCTGATA

AGAAGACCATAATTCAGCGCTTCCAGGAGAAGGTGGAGTCGCTGGAGAAGGAAGCGGCTGG

AGAGAGACAGCAGCTGGTGGAAACGCACATGGCTCGAGTGGAGGCCCTACTGAATGACCGC

CGCCGTCAGGCTCTGGAAAGCTACCTGAGCTCCCTTCAGTCTGACCAGCCTCGGCCTCGGCA

GGTGCTGAATCTGTTGAAGAAGTATATCCGTGCGGAGCAGAAGGACCGTCAGCACACTCTC

AAACACTTTGAACACGTGCGAGAGGTCGATCCCAAGAAGGCTTCACAGATTCGGCCATTTGT

GATGACCCACCTACGTGTGATTGAGGAACGCATGAACCAGTCTTTGGGTTATCTCTATAAAG

TGCCTCAAGTGGCTAATGATATCCAGGATCAAGTGGCGGTGCTGGTTCAGCGTGATCAGGCT

GAGGTGACGCAGCAGCTGTCGTCTCTTCAGAGTAAGATGAGGGTCAGTTATGGGAATGATG

CCCTGATGCCGGATCTGCCCGACAGCACCACACCACTGGACAACCTTCCTCCAGAGCAGGAC

GGCCTGGGCTTCATCCATCCCGAGAGCTTTAACCAGGCCAACACTGACAACCACGTTGAACC

TGTAGATGCCCGTCCAATTCCCGAAAGGGGTTTGCCTACGAGACCCGAGATCCCAAAGGTTC

GGCTGGACATTGAGGAAAGGCACAACGCTGGCTATGATGTTCGTGACAAGAGACTGATGTT

CCTCGCGGAGGACATGGGCTCTAATAAGGGTGCGATCATTGGGCTGATGGTGGGGGGCGTG

GTCATCGCT<mark>A[C]T</mark>GTCATCGTAATCACTCTGGTTATGCTCAGGAAGAAGCAGTACACCTCTA

TTCATCATGGAGTTATTGAGGTGGATGCGGCAGTGACTCCTGAAGAACGTCATCTGGCAAAG

ATGCAGCAGAATGGCTATGAAAACCCCACCTACAAGTTCTTTGAGCAAATGCAGAACTAA   -

3’

(C) The ~2000bp of the zebrafish *appa* sequence that was used to identify sgRNA sequences to target the fAD "Leuven" mutation is included below:

- Yellow highlight and brackets indicate the specific residue we wished to mutate

5'-

ATGCGGTCGAGGGAGCTCTTCATATTACTGATGGCCGTCGCGTCGACGCTCGCGGTGGAGGT

GCCGTCAGACTCCGGGACGGGGCTTCTGGCCGAGCCACAGATCGCCATGTTCTGCGGGAAA

CTCAACATGCACATCAACATCCAGAGCGGGAAATGGGAGCCGGATCCGTCCGGAAGCAAGA

GCTGCATCGGGAATAAAGAGGGAATCCTGCAGTACTGCCAGGAGGTGTATCCGGAGCTGCA

GATCACCAATGTGGTGGAGGCCAATCAGCCGGTCAGCATCTGGGACTGGTGTAAGAAGAGC

CGCAAGCAGTGCCGCAGTCACATGCACATCGTAGTGCCGTACCGCTGCCTGGTTGGGGAGTT

TGTCAGCGATGCTCTGCTGGTTCCTGATAAGTGTAAGTTCCTGCATCAGGAGCGCATGGACA

TGTGTGAGAGTCACCTGCACTGGCACACCGTCGCCAAAGAGTCGTGTGGTGACCGCAGCATG

AATCTGCATGATTATGGGATGCTGTTGCCGTGTGGAATTGACCGTTTCCGGGGTGTGGAGTT

CGTGTGCTGTCCTGCAGACGCGGGTAAAGAGTCTGAGAGCGCCGCTGTGGAGGAGGACGAT

TCAGACGTGTGGTGGGGCGGAGCGGAGGCTGATTACACCGAGAACAGCATGACTCGTGATG

CCGCAGCGGAGCCGGCGGTGCTGGAGGATGATGAGGATGCGGACGAGGAAGAGGACGAGG

ATCAGGATGGAGATGGAGATCGGGACGAGAAGATAGAGGAAGAGGAGGAGGAGGAGGAG

CGCACCCAGAGCACCAGCGCAGCCCTGACCTCCACCACCACCACCACCACTGAGTCTGTAGA

GGAGGTGGTGCGAGAGGTGTGTTTTGCGAGTGCAGAGACGGGTCCCTGTAGGGCCATGTTGT

CCCGCTGGTATTATGTGCGTGAGGAGCGCCGCTGTGCGCCCTTCATCTACGGCGGCTGCGGA

GGAAACCGTAATAACTTTGAGTCGGAGGAATACTGTCTGTCCGTCTGCAGCGGTGTGTTGCC

GACTCCATCCTCCAGCCCTCCGGATGCAGTGGACCGATATCTGGAGACGCCGGCGGACGAG

AACGAACACGCTCACTTTCTACAGGCCAAAGAAAGTTTGGAGACCAAACACCGCGAGCGCA

TGTCTCAGGTGATGAGGGAATGGGAAGAGGCCGAGAGACAGGCCAAGAGTTTACCACGCAA

CGACAAGAAGGCCGTGATCCAGCACTTCCAGGAGAAGGTGGAGGCTCTGGAGCAGGAGTCG

GCCAGCGAGCGTCAGCAGCTGGTCGAGACACACATGGCACGGGTGGAGGCTCTGCTGAACG

ACCGCCGCCGCCTCGCCCTCGAGAGCTACCTGTCTGCCCTGCAGGCCGATCCACCACGGCCT

CGTCATGTGTTCAGCCTGCTGAAGAAGTATGTACGGGCCGAGCAGAAGGACCGTCAACACA

CACTCAAACACTTCGAACACGTGCGCATGGTGGACCCCAAGAAAGCAGCACAGATACGGCC

TCAGGTGTTGACGCATCTGCGTGTGATCGAGGAGCGCATGAACCAGTCTCTCGGCCTGCTGT

ACAAAGTGCCCGGCGTTGCTGACGACATCCAGGACCAAGTCGAGCTGCTGCAGCGTGAGCA

GCAGGAGATGTCCGCTCAGTTGGCGAATCTGCAGAGTGACGCGCGTGTGAGTTATGGGAAC

GATGCGCTGATGCCCGACAGCACCGCCGGCCTGGAGCTGCTGCCCGCCGAGGACACACAGG

GGTTCGGCTTCATACACCCCGAGAGCTTCAACCAGCCCAACACACACAACCAGGTTGAGCCT

GTCGATGCCCGACCTGTTCCAGACCTAGACCTGGCAACCCGACCAGTGTCTGGACTGAAGCC

TGATGACATTCCTGAGCTGCGGATGGAAGCTGAAGAGAGACACAGT<mark>G[A]A</mark>GTCTACCACCA

GAAGCTGGTTTTCTTTGCGGAGGACGTGAGCTCCAATAAAGGAGCTATTATTGGCCTGATGG

TCGGAGGCGTCGTCATAGCAACCATCATCGTCATCACGCTGGTGATGCTGAGGAAGAAGCA

GTACACGTCCATCCACCACGGCATCATCGAGGTGGACGCGGCCGTGACTCCAGAGGAGCGT

CATCTGTCTAAGATGCAGCAGAACGGCTATGAAAACCCCACCTACAAGTTCTTTGAGCAGAT

GCACAACTGA-3'

(D) The ~1300bp of the zebrafish *appa* sequence that was used to identify sgRNA sequences to target the start codon is included below:

- Yellow highlight and brackets indicate the specific residue we wished to mutate

5'-

ACTAAATAATTTATATATTATACCTTATAACGAACACATCCCTGACTGAAAAGTTAAATATT

AATTAATTAACATTTAATAGATAATTAAAATGTATACTATTTCTTTTTTATTATAACATCAAT

TAGATCATTATTGTGTTTAACTATTACATTTCTATAAAATAGATAACTTATAAATGATATTTT

AAAAAGTAGATGCCTGACTATAGACTTAAATATTTGATTGTTTTTCTCAATATAAAATATATA

CGTTTTCTTTTTTTATTATTATATAAAATATGTATTCGATCACAATTATATTTGTTTGTACTTTT

TGTCAAATACGTTTTTAAAAATATAAAAAACAGATGCTTGTATCAGATCATAGTTGATTTATT

TATTATATTAAATAATTTATATATTATAAAATATAATAAACACATCCCTGCCCGAAAAAAAT

ATATCTTATTTTTCTAAAATTATGATATATACCTTTCATTTTATTTTTATTATTAATTTATTTTT

TATTTCTCACAGCACATTATGTAAATAAGCTCCTTCCTTGATCATTGATTGGCGTGTTCTCCG

TTTGCTCCTCCCCCTACCGTGACGTACGGCTCTGATTCTCACCGCAGGCGCATCGCGTTCTTC

ACAGAGCCACAAACACATCGACATTCATTCACAGCGAACGGATAAATCTAAATACAACCAA

AAACACAGAAGGAAAGGCGAAATAAAGAGGAGCAGAAGC[ATG]CGGTCGAGGGAGCTCTT

CATATTACTGATGGCCGTCGCGTCGACGCTCGCGGTGGAGGTAAGAAAACGACAAAGATA

ACGTTACGTTTGATCTCTGAGGTAAAACATATTCACAACTGACGATTATTCCTATAGAGTCCT

ATTGAGCCTTGATATTTAATGTTTTGAGGCGTGTAAATGTTAGATGTGTTATCATTTCATCTG

CTGGTGTGGATTATTCGGTGAAGATGTTCGCTGGACGGGCGGAATGTGTAGCGGGATCAGCC

GCGTTACATCAGCGTTACATTTAAAACATCGATGATTAATTCATACTGTACACTGTAACGCA

TCTGCTTATTAATATTAATAAACAAACTGACGCGCTAAAAATACCGCGAGTGTCGGTGGTGG

TAAACGCGTTTATAACATTCACTGCGTTACATGTTCACATTCTGTTGATTTGTATTGATTCAA

TGGACAAATGTTGTGTTTATAGTGGTTTTTAGTCGATGTTTAACCTTAGACGATGGTTTAATA

TCGAATAATAAGGCTTCCACGAATAATATGGCTCCCAGTGTTGTGGTTGTGAAATATGATTT

AATATTCCCTCGAGCTATGCTGGTGTCAGTGGGTGTGTAGGGAGGGAACTTGATTTTATGAC

GCTC-3'

(E) The ~450bp of the zebrafish *appb* sequence that was used to identify sgRNA sequences to target the start codon is included below:

- Yellow highlight and brackets indicate the specific residue we wished to mutate

5'-

CACACGATAGGGGCACACGGAGCAGAAAATCGCGACAGAAAAACCCTGATCCGCTCAGGAT

ATATATTCACCAGGACGTGCTGCGCTTGGGAACACAGCC[ATG]GGTATGGACCGCACGGTAT

TCCTGCTGTTAATGCTGACGACTTTGTCCCTCGCCATCGAGGTAAGAATGATTGTGTAATGG

AGAAGGAGCTTGGTTCTCCTCCATACTTTAAAGGGCGGCCTGGGAGTGAAGGGAAAACGCA

TGACACGGATGCAGACAGACATTTTGGGCGCTGCATGATTGGCATTTGCAAGATGATCGTTT

TTTATAATGCAATATATATGTTTATATTTGGGACGTGCGAGTTTTAAATCGGTTAGTGTTGGG

TGGAGGCTTCTGGTGCAGGCTGCTTTACACTAATAGTATCAGGCGTAGTGTAGCGCGTGAAA

GTGATACACATCACTGAA-3'

144

File S2. Supporting tables.

Table S1. sgRNAs designed for the ADDGene method using Zfit targeter, with respective CRISPRscan scores.

- N/A indicates that an sgRNA was not identified via CRISPRscan software.

- All sequences are 5'-3'.

Gene-mutation	Target site	Oligo 1	Oligo 2	CRISPRscan score
Appb-Arctic	GGAGGACATGG GCTCTAATA	TAGGAGGACATG GGCTCTAATA	AAACTATTAGAGC CCATGTCCT	61
Appb-Austrian	GGTACTGTCATC GTAATCACTC	TAGGTACTGTCAT CGTAATCACTC	AAACGAGTGATTA CGATGACAGTA	N/A
Appa-Leuven	GGTGAAGTCTAC CACCAGAAGC	TAGGTGAAGTCTA CCACCAGAAGC	AAACGCTTCTGGT GGTAGACTTCA	N/A
Appa- start codon	GGAGCAGAAGC ATGCGGTCG	TAGGAGCAGAAG CATGCGGTCG	AAACCGACCGCAT GCTTCTGCT	70
Appb- start codon	GGCCATGGGTAT GGACCGCA	TAGGCCATGGGTA TGGACCGCA	AAACTGCGGTCCA TACCCATGG	66

Table S2. sgRNA primers designed for the PCR method using CRISPRscan software.

- Red text highlights guanine nucleotides artificially inserted to fit CRISPR/Cas9 requirements.

- Blue text indicates the identified sgRNA target sites

Gene-mutation	Backbone + sgRNA design (5'-3')	CRISPR scan score
Appa- start codon	GAAATTAATACGACTCACTATAGG[G]GCGTCGACGCTCGCG GTGGGTTTTAGAGCTAGAAATAGC	81
Appa- Austrian	GAAATTAATACGACTCACTATAGG[G]GGGCTGATGGTGGGG GGCGGTTTTAGAGCTAGAAATAGC	66
Appb- Austrian	GAAATTAATACGACTCACTATAGGCACCAGCGTGATGACGA TGAGTTTTAGAGCTAGAAATAGC	79

Table S3. Primers used for *appa* 5'UTR amplification and sequencing

- F = forward, R = reverse

Primer name	Use	Direction	Primer Sequence (5'-3')
appa 5'UTR	Amplification	F	ATAACGAACACATCCCTGACTGA
appa 5'UTR	Amplification	R	CCCACTGACACCAGCATAGC
M13 F	Sequencing	F	GTAAAACGACGGCCAGT
M13 R	Sequencing	R	AACAGCTATGACCATG

Table S4. Single-stranded donor DNA oligonucleotide templates for homology-directed repair

- The codon incorporating the mutation is underlined.

- Bold lettering indicates nucleotides that were substituted from the original genomic DNA sequence to give the desired mutation upon incorporation via HDR.

- Bold red text indicates silent substitutions that were made to aid specific binding of the primer.

Gene-mutation name	Oligonucleotide sequence (5'-3')
appb-Arctic	TGTTGAATCAGATGTTCCTCGCGG**GA**GACATGGGCTCTAATAAGGGTGCG
appa-Austrian	ATGGTCGGAGGCGTCGTCATAGC**G**ATAATCATCGTCATCAC**CT**TGGTGAT

Table S5. Primers used to identify double-stranded breaks induced by TALEN/CRISPR and to detect successful incorporation of ssONs carrying the mutation of interest.

- F = forward, R = reverse

Primer name	Use	Direction	Primer sequence (5'-3')
Arctic and Austrian *appb* F	T7 assays	F	GAACCTGTAGATGCCCGTCCAA
Arctic and Austrian *appb* R	T7 assays	R	CGCTGCTCAGTAGGTATTCACTTGC

Arctic *appa*	Sequencing	F	CCGTCCAATTCCCGAAAG
London *appb* F	T7 assays	F	CCTGTAGATGCCCGTCCAA
London *appb* R	T7 assays	R	GCTGCTCAGTAGGTATTCACTTGC
Leuven *appa* F	T7 assays	F	CAGTGTCTGGACTGAAGCCTG
Leuven *appa* R	T7 assays	R	TTTCCCTCACACAACCACAG
appa start codon F	T7 assays	F	GGCTCTGATTCTCACCGCA
appa start codon R	T7 assays	R	CTTCACCGAATAATCCACACCAG
appb start codon F	T7 assays	F	CTCCAACCTCCGAGTCCATTCT
appb start codon R	T7 assays	R	TGCGTTTTCCCTTCACTCCC
Arctic *appb* F	ssON detection	F	AATCAGATGTTCCTCGCGGGAG
Arctic *appb* R	ssON **and** WT detection	R	CGCTGCTCAGTAGGTATTCACTTGC
Arctic *appb* WT F	WT detection (ssON expt)	F	AATCAGATGTTCCTCGCGGAGG
Austrian *appa* F	ssON detection	F	GATAATCATCGTCATCACCTTGG
Austrian *appa* R	ssON detection	R	CATTTCTGACCGTGATCTGGC
Austrian *appa* WT F	WT detection (ssON expt)	F	GCAACCATCATCGTCATCAC
Austrian *appa* WT R	WT detection (ssON expt)	R	CTCCCCGAACCCTCCTCA

File S3. Sequence containing the region of genomic DNA just upstream of the start codon of *appa* generated by Sanger sequencing

- Blue text indicates exon 1 sequence.

- Start codon is indicated in bold blue text.

5'-

ATAACGAACACATCCCTGACTGAAGTTAAATATTAAATAATTAACATTTAGTAGATAATTAA

AATGTATACTATTTATTTTTTATTATAACATCAATTAGATCATTATTGTGTTTAACTATTACAT

TTCTATAAAATAGATAACTTATAAATGATATTTTAAAAAGTAGATGCCTGACTATAGACTTA

AATATTTGATTGTTTTTCTCAATATAAAATATATACGTTTTCTTTTTTTATTATTATATAAAAT

ATGTATTCGATCACAATTATATTTGTTTGTACTTTTTGTCAAATACGTTTTTAAAAATATAAA

AAAACAGATGCTTGTATCAGATCATAGTTGATTTATTATATTAAATAATTTATATATTATAAA

ATATAATAAACATATCCCTGCCCGAAAAAATATATCTTATTTTTCTAAAATTATGATATATAC

CTTTCATTTTATTTTTATTATTAATTTATTTTTTATTTCTCACAGCACATTATGTAAATAAGCT

CCATTCCTTGATCATTGATTGGCGTGTTCTCCGTTTGCTCCTCCCCCTACCGTGACGCGCGGC

TCTGATTCTCACCGCAGGCGCATCGCGTTCTTCACAGAGCCACAAACACATCGACATTCATT

CACAGCGAACGGATAAATCTAAATATAACCAAAAACACAGAAGGAAAGGCGAAATAAAGA

GGAGCAGAAGC**ATG**CGGTCGAGGGAGCTCTTCATATTACTGATGGCCGTCGCGTCGACGCT

CGCGGTGGAGGTAAGAAAACGACAAAAGATAACGTTACGTTTGATCTCTGAGGTAAAACAT

ATTCACAACTGACGATTATTCCTATAGAGTCGCATTGAGCCTTGATATTTAATGTTTTGAGGC

GTGTAAATGTTAGATGTGTTATCATTTCATCTGCTGGTGTGGATTATTCGGTGAAGATGTTCG

CTGGACGGGCGGAATGTGTAGCGGGATCAGCCGCGTTACATCAGCGTTACATTTAAAACATC

GATGATTAATTCATACTGTACACTGTAACGCATCTGCTTATTAATATTAATAAACAAACTGA

CGCGCTAAAAATACCGCGAGTGTCGGTGGTGGTAAACGCGTATATAACATTCACTGCGTTAC

ATGTTCACATTCTGTTGATTTGTATTGATTCAATGGACAAATGTTGTGTTTATGGTGGTTTTTA

149

GTCGATGTTTAACCTTAGACGATGGTTTAATATTGAATAATAAGGCTTCCACGAATAATATG

GCTCCCAGTGTTGTGGTTGTGAAATATGATTTAATATTCCCTCGAGCTATGCTGGTGTCAGTG

GG-3'

File S4. Expression levels of *appa* and *appb* during early development
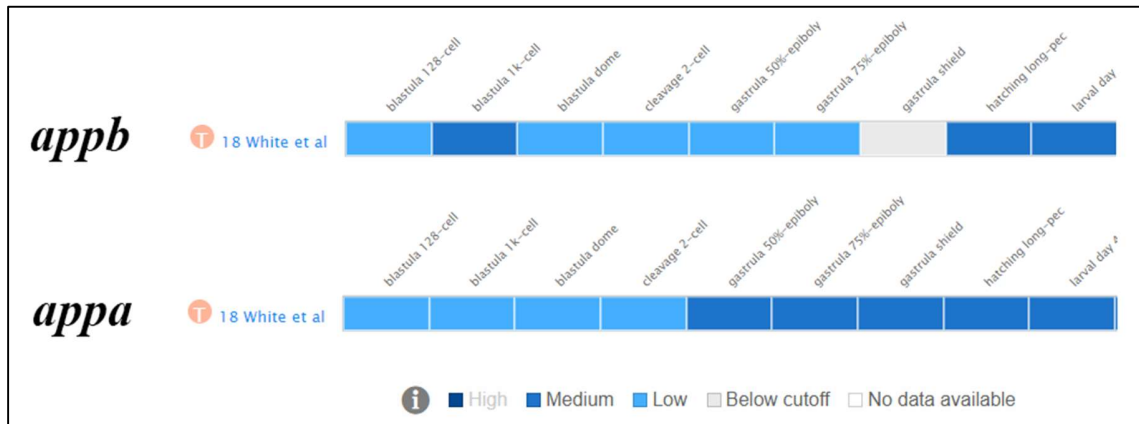
Figure S1. Screen shots from the EMBL-EBI "Expression atlas" showing expression levels of **appa** *and* **appb** *during early developmental stages. Legend at the bottom of the figure indicates expression level by colour.*

Chapter 4 – Preliminary analysis of a putative null mutation in zebrafish *appb*

4.1 Introduction

γ-secretase is an important protease complex due to its cleavage of over 100 substrates within their transmembrane domains [1, 2]. γ-secretase plays a role in Alzheimer's disease (AD) through its processing of APP to produce the aggregation-prone Aβ peptide [3]. Sequential processing of APP that produces the Aβ peptide is thought to be a major mechanism affected in AD disease progression, as described by the amyloid cascade hypothesis [4]. In this hypothesis, when APP is processed by β-secretase (BACE1) followed by γ-secretase, the Aβ peptide is released. This is in contrast to the processing of APP by α-secretase followed by γ-secretase, which releases the non-amyloidogenic P3 peptide [5]. However, as mentioned previously, although Aβ accumulation into plaques is evident in the disease, there is no quantitative relationship between this and other factors of the disease [6]. There are many alternative hypotheses that attempt to explain this disparate relationship between Aβ plaques and other factors. One such hypothesis is the Hypoxia/Ischemia hypothesis.

Local hypoxia refers to a state in which part of the body lacks adequate oxygen supply to tissue. Ischemia is one of the more common causes of hypoxia and refers to times where there is insufficient blood flow to a tissue. Interestingly, ischemia has been observed in the brains of individuals experiencing Mild Cognitive Impairment (MCI), which is a condition considered a precursor to dementia [7]. APP, β-secretase (BACE1) and γ-secretase expression have all previously been found to be increased during hypoxia while α-secretase expression is decreased [7, 8]. It is presumably this shift that results in an observed increase in Aβ production under

hypoxic stress. Some studies have also suggested that oxidative stress, induced by acute hypoxia, may increase APP cleavage by γ-secretase [8]. Furthermore, stimulation of oxidative stress due to accumulation of Aβ peptides has also been found to occur, which would create a positive feedback loop potentially accelerating brain aging [8]. Hypoxia-inducible factor 1 (HIF-1) is considered the master transcriptional regulator of the cell's responses to hypoxia. HIF-1 is a transcription factor consisting of both an oxygen-regulated α subunit and a constitutively expressed β subunit. The oxygen-regulated HIF1-α subunit is degraded under normal oxygen conditions (normoxia), whereas, under hypoxia this subunit is stabilised and translocated to the nucleus. Once in the nucleus, HIF1-α and HIF1-β dimerise and can induce expression of hypoxia-regulated genes [reviewed in 9]. During hypoxia, HIF-1 induces transcription of over 70 HIF-1 responsive genes (HRGs) including *IGFBP3* and *EDN1* [10]. These HRGs can be used to measure the hypoxic response in genetically modified organisms.

Zebrafish are widely used to study genetic influences on diseases as their genome is approximately 70% conserved with humans [11]. A study by Moussavi Nik *et al* (2012) found that transcription of zebrafish genes encoding Bace1 and Appa and Appb, along with Psen1 and Psen2 (the active components of γ-secretase), are also induced by hypoxia [12]. This finding suggests that regulation of these genes by hypoxia is conserved in the zebrafish, making it a good model for studying the mechanisms by which hypoxia/ischemia might lead to AD. Recently, observations by Newman *et al* (2019) revealed an interesting response to hypoxia in the brains of zebrafish carrying fAD-like mutations in the *psen1* gene [13]. When measuring HRGs, they saw an acceleration of normal age-dependent changes in the brains of their mutant zebrafish. This supports the idea that age is one of the major risk factors for AD. Further investigation of the hypoxic response of fAD-like mutations

in other fAD mutant genes (*psen2*, *appa*, *appb* and *sorl1*), may provide further support to the notion that, not only is age a risk factor for AD, but AD is actually an inevitable consequence of aging. In Chapter 3, it was established that sgRNAs designed to target the Arctic mutation in *appb*, the Austrian mutation in *appa* and the start site of *appb* were successful in generating double-stranded breaks (DSBs). The purpose of targeting the start site of *appb* was to generate a null mutation. One original aim for this project was to generate zebrafish null models both for *appa* and *appb,* so that the effects of complete loss of these proteins could be compared to the effects of fAD-like mutations in zebrafish *psen1*, *psen2*, *appa*, *appb* and *sorl1* using transcriptomic analyses. A long-term goal of the laboratory is to perform these analyses in the hopes of identifying common molecular changes between fAD-like mutations, to assist in our understanding of the molecular progression into AD.

In this study, the identification of a null mutation of *appb* will be outlined, along with some preliminary experiments conducted to investigate whether loss of *appb* expression, and therefore a loss of Appb protein, affects the response to acute hypoxia in zebrafish brain. Furthermore, successful breeding of the putative *appb* null to produce homozygous mutant fish was achieved. These fish can now be utilised in future transcriptomic analyses to contribute to the wider laboratory program, to identify the molecular "signature" of AD.

4.2 Results and Discussion

Screening F0 founders for potential null mutations

The *appb* start site sgRNA, that was shown to induce DSBs in Chapter 3, was co-injected with Cas9 protein into fertilised one cell stage embryos. Embryos injected with *appb* sgRNA were

raised to adulthood and will henceforth be referred to as "F0" or "founder" fish. At ~3 months of age, founder fish were screened across the 5'UTR/Exon 1 region for potential mutations that would prevent translation of Appb protein. Screening was performed using primers amplifying a 345bp fragment spanning the 5'UTR/Exon 1 region of the *appb* gene, ~216bp upstream and ~129bp downstream of the PAM site (Figure 4.1).

```
...tttactcattgaccacctccaacctccgagtccattctcctccacttctagacgtgtggtcag
ctgactttccctggagcagctgtgtgcagtcgcgcattctcacaaagcacaaccacacgatagggg
cacacggagcagaaaatcgcgacagaaaaaccctgatccgctcaggatatatattcaccaggacgt
gctgcgcttgggaacacagccATGGGAATGGATCGCACGGTATTCCTGCTGTTAATGCTGACGACT
TTGTCCCTCGCCATCGAGGTAAGAATGATTGTGTAATGGAGAAGGAGCTTGGTTCTCCTCCATACT
TTAAAGGGCGGCCTGGGAGTGAAGGGAAAACGCATGACACGGATGC...
```

Figure 4.1 appb *start site, sgRNA binding site, and primer design. – a) region of 5'UTR and exon 1 sequence within which primers were designed to amplify genomic DNA for T7 cleavage assays. The* appb *start site sgRNA binding site is represented by the black bar, the PAM by the red bar. The green box highlights the start codon in exon 1 of* appb. *The forward primer sequence is highlighted in blue and the reverse primer binding site is highlighted in yellow. Lowercase letters represent 5'UTR sequence, while uppercase represent sequence from exon 1. The sequence in this figure reads 5'-3'.*

Of the 15 adult fish that were biopsied by tail clipping in the initial founder screen, T7 endonuclease 1 assays identified 11 individuals that were positive for what could potentially be an *appb* null mutation (Figure 4.2 B). As CRISPR/Cas9 can generate DSBs continuously as the embryo develops, zebrafish founders injected with the CRISPR/Cas9 system are generally mosaic for mutations [14]. Therefore, founders that were positive for mutations were outcrossed to wild type, Tübingen (TU), fish and the progeny of these crosses (F1) were screened for germline-transmitted mutations.

3 of the 11 founder fish were selected to be outcrossed to wildtype for mutation screening. Fish F0-1, F0-3, F0-6, F0-7, F0-10 and F0-13 displayed a banding pattern corresponding to the cleavage fragment sizes predicted from the DSB test (figure 4.2 B). Fish F0-3 and F0-6 from this cohort were initially selected for further breeding and analysis. Interestingly, our screen also revealed a single band above the larger of the predicted (216bp) cleavage fragments in fish F0-5, F0-12 and F0-13 (figure 4.2 B). It was suspected that this discrepancy may be indicative of a larger deletion/insertion introduced through NHEJ. Fish F0-5, displaying this unexpected banding pattern was therefore selected for further breeding and analysis. Fish F0-11, F0-14 and F0-15 also displayed aberrant banding patterns to those described above, that were not present in the wildtype and did not correspond to the predicted fragments but were not selected for further analysis.
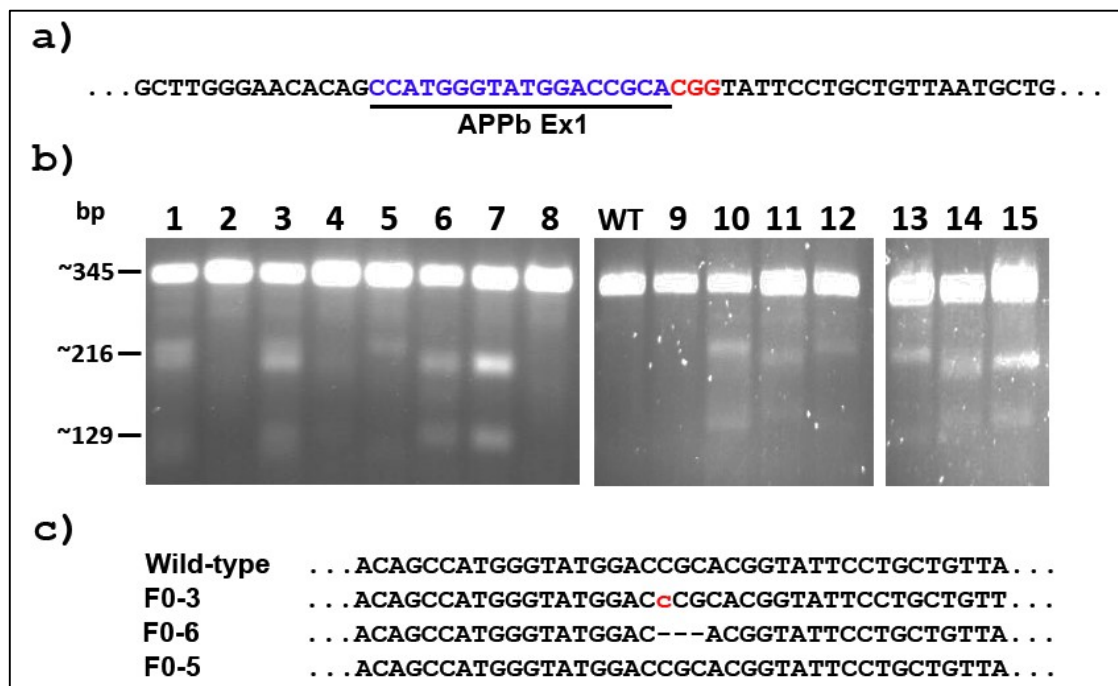
Figure. 4.2. Analysis of F0 founder fish generated from **appb** *start site sgRNA injections – a)* appb *start site sgRNA binding site at the boundary of 5'UTR and exon 1 (Ex1). The sgRNA sequence is indicated in blue text and is underlined. The 3bp PAM sequence is indicated in red text; b) T7 endonuclease I assay of*

F0 founders. Each label number above the gel image represents an arbitrarily numbered adult zebrafish and the wildtype control is labelled WT; c) the sequence of zebrafish appb *around the sgRNA binding site for founder fish F0-3, F0-5 and F0-6 (from the gel images in b), above). Dashes indicate deleted nucleotides; red and lowercase text indicates an inserted nucleotide.*

Attempting to identify appb *mutations in F0-5 and F0-6 founder fish*

In order to identify potential mutations in our selected founder fish, a random pool of 10 embryos from the separate outcrossing of wildtype to F0-5 or F0-6 were first screened to confirm germline transmission (supplementary data file S1, Figure S2). A further screen of 6 adult F1 progeny from fish F0-5 revealed that approximately 67% (4/6) were positive for the same banding pattern that was observed in Figure 4.2 B, indicating that these fish are transmitting a mutation (supplementary data file S1, Figure S3). These positive F1 progeny (F1-5.2, F1-5.3, F1-5.4 and F1-5.6) of F0-5 outcrossed to wildtype were sequenced using the Sanger method to identify any mutations present near the start codon. Unfortunately, Sanger sequencing did not reveal any mutations at or around the start codon (supplementary data file S2, Figure S6). The absence of a mutation at the start codon in these F1 fish suggests that potential mutations observed from tail clippings of F0-5 fish do not exist in both tail fin cells and the germline. However, it is also possible that only a few of the embryos in the original random pool that was used for screening carried the detected mutant genotype. Therefore, as individual adult F1 generation zebrafish were selected for sequencing randomly, it could be purely chance that this genotype was not represented in the sequencing. This result could be due to the mutation not being carried through to all cells in the germline of the F0 fish. Another possibility to explain this result may be the presence of a natural polymorphism within the 5'UTR and exon 1 region that was used for screening.

Screening of 39 adult F1 progeny from fish F0-6 crossed with wildtype, revealed that approximately 8% (3/39) carry a banding pattern identical to that of the founder fish, which aligns with the predicted fragments expected from DSB induction by this sgRNA (Figure 4.2 B). Interestingly, some fish showed faint bands that do not align with the original cleavage pattern as seen in Figure 4.2 B (supplementary data file S1, Figure S4). Sanger sequencing of F1 fish 6-13, that carries the same banding pattern as F0-6, revealed a 3bp deletion within the 5'UTR/Ex1 region (Figure 4.2 C, supplementary data file S2, Figure S7). Unfortunately, the 3bp deletion isolated in this mutant fish would not result in the *appb* transcript being read out of frame hence an (almost) full length Appb protein would be generated (supplementary data file S2, figure S9). Since neither of the F0-5 or F0-6 lines carried a mutation that would result in an Appb null protein they were not included in further analyses.

Identification of a putative appb *null mutation in founder fish F0-3*

A pool of 10 randomly selected embryos from the outcrossing of F0-3 to a wildtype fish was also screened, confirming germline transmission (supplementary data file S1, Figure S5 a). Further screening of 20 adult F1 fish from this mating by T7 endonuclease 1 assay, showed that 10% (2/20) of these F1 progeny carry the same banding pattern as F0-3 as shown in Figure 4.2 B (supplementary data file S1, Figure S5 b-e). Sanger sequencing performed on F1 fish F1-3.8 of this mating identified a single C-insertion mutation at position 13 in the 5th codon (supplementary data file S2, Figure S8). In silico translation of cDNA containing the C-insertion using the ExPASy software [15] revealed a frameshift, resulting in replacement of the arginine for proline at position 5 in the amino acid sequence, followed by replacements of all amino acids up until the premature

stop codon (PTC) at position 26 in the amino acid sequence. Therefore, the *appb* C-insertion mutant will henceforth be referred to as *appb^R5PfsTer26^* (Figure 4.3).

Figure 4.3. Outline of* appb^R5PfsTer26^ *mutation. *Red star indicates position of the single C insertion that shifts the protein out of frame, resulting in an early stop codon, indicated by black box and star.*

Prediction of alternative start site priming for this mutant using the NetStart 1.0 prediction software [16], suggests that the start site that gives the truncated protein is preferred (supplementary data file S3, Figure S10). An additional search for alternative methionine (start codons) within the first and second exons of zebrafish *appb* that could re-initiate translation, as displayed on the Ensembl genome browser (https://asia.ensembl.org/index.html), revealed that there are no alternative start codons in either of these exons. Also, translation of *appb^R5PfsTer26^* obtained from the ExPASy software (supplementary data file S3, Figure S11) showed that there are no alternative start codons for the correct reading frame downstream of the PTC.

Breeding the appb *null mutant to generate families of WT, heterozygous and homozygous fish for analysis of mutational effects*

103 F1 fish from the mating of F0-3 x TU were screened for the *appb^R5PfsTer26^* mutation. The initial 20 fish were screened by T7 endonuclease 1 assay, while the remainder were screened using mutation-detecting primers. Designing primers to detect the single base pair change of the

appb^{R5PfsTer26} mutation was challenging. The *appb^{R5PfsTer26}* mutation consists of three consecutive cytosines, where previously there were only two. Therefore, a mutation-detecting primer was designed that placed the triple cytosine as close to the 3' end of the primer as possible, as DNA polymerase can only synthesis in the 5'-3'direction and thus will first bind strongly to the 3' end to commence replication. Approximately 13% of the 103 F1 fish screened positive for germline transmission of the mutation (13/103, as shown in supplementary data file S4, Table S2). Fish F1-3.30, that was genotyped as heterozygous for the null mutation using mutation detecting primers, was confirmed by Sanger sequencing to carry *appb^{R5PfsTer26}*. In order to minimise off target mutations carried through to further experiments, fish F1-3.08 (that had been confirmed by Sanger sequencing to carry *appb^{R5PfsTer26}*) was outcrossed to a wildtype fish from a different family to generate an F2 family. Approximately 36% of the 33 (12/33) F2 fish that were tested, screened positive for a heterozygous mutation.

In zebrafish, *appb* has previously been observed to play a role in convergent extension movements, development of body length and neural development [17, 18]. Involvement in these processes suggests a key role for *appb* in zebrafish development. Interestingly, we observed no obvious developmental abnormalities in zebrafish heterozygous for the *appb* null mutation. The lack of an observable phenotype in individuals heterozygous for this putative null mutation (as had previously been observed with morpholino knockdown) is perhaps not surprising, as previous studies have shown that phenotypes caused by morpholino knockdown are most often different to those observed from genetic mutations that have produced null alleles in the same gene [19].

Generating homozygous appb *nulls for phenotypic analysis and future transcriptomic analyses*

Further breeding of *appb^{R5PfsTer26}* mutant zebrafish was conducted in order to; 1) investigate the potential of generating zebrafish that were homozygous viable for *appb^{R5PfsTer26}*, and 2) generate families containing a combination of heterozygous, homozygous and wildtype genotypes for multiple three-way comparisons (i.e. pairwise comparisons between heterozygous/homozygous, heterozygous/wildtype and homozygous/wildtype for later transcriptomic and other analyses). Therefore, a pair of heterozygous *appb^{R5PfsTer26}* fish were mated to produce an F2 generation family that should contain a combination of heterozygous, homozygous and wildtype in a 2:1:1 ratio. The progeny of this original mating were raised to 5dpf, at which point we observed their phenotype to look for any changes that might be associated with homozygous siblings.

Since we previously observed no phenotype in heterozygous individuals, we were surprised to observe phenotypic changes from this mating. We classified larvae from this mating into 3 phenotypic groups (supplementary data file S5, Figures S12, S13 and S14): severe (with phenotypic traits ranging from: tail curved down, edema, short tail and up curved fish [20]), mild/moderate (up curved tail, down curved fish (mild) [20] and under-developed swim bladder) and wildtype (normal). These larvae were then harvested for genomic DNA and genotyped to determine whether genotypes correlated with the observed phenotype groups. Unexpectedly, genotypes did not align with our predictions from the observed phenotypes. We expected larvae displaying severe phenotypes to be homozygous for *appb^{R5PfsTer26}*, whereas mild/moderate phenotypes might be homozygous or heterozygous and normal phenotypes might be heterozygous or wildtype. However, we observed no correlation between genotype and phenotype. Of the 21 larvae genotyped, ~33% (7/21) were heterozygous, ~38% (8/21) were homozygous mutant and

~29% (6/21) were wildtype. Interestingly, no obvious phenotype was observed in the F2 offspring from additional mating's of two alternative zebrafish heterozygous for the *appb* null mutation. Since the progeny of only a single pair mating was observed in this analysis, the phenotypical differences observed in the F2 family may have been due to individual fish variability or genetic impacts at genome editing other than generation of *appb*^{R5PfsTer26}. For example, an off-target mutation may unintentionally have been created during the process of generating this mutant fish line. Alternatively, our laboratory has previously observed poor quality embryos produced from the pairing of two non-mutant TU fish. This may be due to homozygosity for deleterious recessive mutations present in this inbred fish line. The observation of the F2 offspring from additional mating's of two alternative zebrafish heterozygous for the *appb* null mutation having no observable phenotype better fits what we would expect from the genetic compensation response.

Investigating the hypoxic response of young appb^{R5PfsTer26} *mutant zebrafish by examining HRGs*

Zebrafish are a versatile model for the genetic analysis of the response to acute hypoxia [12]. A previous study in our laboratory used digital real-time polymerase chain reactions (dqPCRs) to compare transcript levels of five HIF-1 responsive genes (HRGs) in 6-month-old zebrafish brains, carrying two different fAD-like mutations in *psen1*, after treatment with normoxia or acute hypoxia [13]. In the study, an expected increase in expression of HRGs was observed in wildtype zebrafish after treatment with acute hypoxia. Interestingly, the expression of HRGs in *psen1* fAD-like mutant zebrafish under normoxia was raised to levels similar to that of the wildtype hypoxia-treated zebrafish, suggesting hypoxic-like stress in the young mutant brain [13]. Treatment of 6-month-old *psen1* fAD-like mutant zebrafish with acute hypoxia further exacerbated the observed increase in HRGs expression [13].

A study using mouse embryonic fibroblasts (MEFs) suggested that the APP intracellular domain (AICD), generated by cleavage of the APP protein by PSEN1/2, can induce *Hif1a* gene expression and HIF-1α protein stability [21]. Furthermore, it has also been demonstrated that hypoxia upregulates both APP mRNA and protein expression, resulting in accumulation of Aβ [reviewed in 22]. If APP plays an integral role in the hypoxic response through generation of its AICD, one would expect that knocking out APP would reduce the brain's ability to respond to hypoxia. Our study also utilised dqPCR to establish the effects that our *appb* putative null allele asserts on the HIF-1 hypoxic response, by comparing transcript levels from two HRGs: *edn1* and *igfbp3* [23], in 6-month-old zebrafish brains that had been exposed to either normoxia or acute hypoxia. Unlike the study with *psen1* fAD-like mutant zebrafish, we did not observe increased levels of HRGs in the normoxia treated heterozygous *appb* putative null mutation fish, nor did we observe the decrease in hypoxic response we would expect from loss of the Appb AICD. However, expression of both HRGs was raised under acute hypoxia in the 6-month-old brains of both wildtype and *appb* putative null zebrafish (although there was an upwards trend observed, *edn1* did not reach significance in the wildtype) (Figure 4.4).

The human APP protein has been identified in three different isoforms that occur due to alternative splicing of exons 7 and 8 of *APP* transcripts [24], the full-length isoform APP770, isoform APP751 lacking exon 8, and isoform APP695 which lacks both exon 7 and exon 8. In zebrafish there are two co-orthologues of the human APP gene, *appa* and *appb* [25]. The *appa* orthologue shares greater identity to APP770 and APP751, while the *appb* orthologue is more closely related in structure to APP695. The human APP695 isoform is nearly exclusively expressed in the CNS (the site of increased hypoxia in AD), while APP770 and APP751 are expressed elsewhere [26].

Although, structurally, *appb* resembles more closely the human APP695 isoform, while *appa* is closer to the APP751 and APP770 isoforms, physiologically these zebrafish genes are both expressed widely throughout the developing embryo. A study by Kaiser *et al* (2012) found that *appa* and *appb* were mostly redundant during zebrafish development, where messenger RNA (mRNA) of *appa* rescued the phenotype induced by Appb morpholino knockdown, and *vice versa* [27]. Hence, the observed lack of decrease of the hypoxic response in our heterozygous *appb* putative null mutants is perhaps not unexpected, as in zebrafish two healthy copies of the *appa* allele may be able to compensate for loss of the Appb AICD. Another possibility is that a single wildtype copy of the *appb* allele may be enough to sufficiently respond to hypoxia. Zebrafish homozygous for the *appb* putative null mutation should therefore be tested to investigate whether knocking out both copies of *appb* is sufficient to decrease the hypoxic response. Furthermore, the lack of statistical significance in induction of the hypoxic response observed when measuring *edn1* in our wildtype zebrafish brains is consistent with what has previously been observed in our laboratory, where recent experiments have shown that *edn1* expression is only increased under very severe acute hypoxia (unpublished data). This variability suggests that *edn1* is perhaps not the best marker to measure the hypoxic response in our zebrafish.
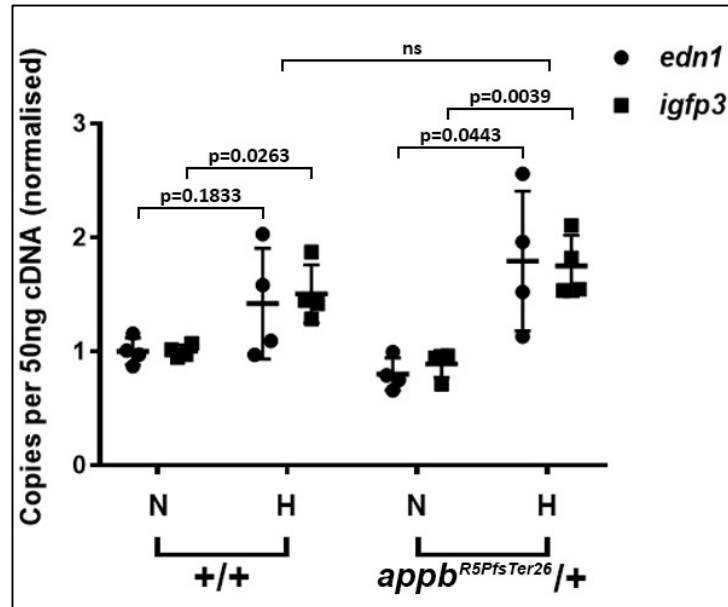
Figure 4.4. Hypoxia response gene expression in 6-month-old zebrafish brains under normoxia and hypoxia. *Data points on the graph indicate relative transcript levels of edn1 (●) and igfbp3 (■) in 50ng of cDNA generated from a single zebrafish brain RNA sample. The genotype of each sample is indicated at the bottom of the graph. Normoxia treated samples are indicated with "N", while hypoxia treated samples are indicated with "H" on the x-axis.*

4.3 Conclusions and Future Directions

In this study, we utilised the *appb* start site sgRNA that was demonstrated to generate DSBs in Chapter 3 in an attempt to isolate *appb* null mutations. Three zebrafish founders were screened for potential *appb* null mutations. Of these founders, F0-5 and F0-6 did not carry a null mutation. Therefore, neither of these fish were selected for further breeding and analysis. However, we successfully identified a founder carrying a putative null mutation in *appb* when sequencing F1 progeny from the F0-3 founder fish crossed with TU. This *appb^R5PfsTer26^* mutation was carried through to the germline and successfully bred to produce heterozygous and homozygous families. In future, western blot analyses could be performed on these homozygous fish, to confirm that our

putative *appb* null leads to loss of Appb protein. A study by Kaiser *et al* (2012) previously used the 22C11 APP antibody to detect changes in the amount of zebrafish Appa or Appb protein after morpholino knockdown [27]. They demonstrated that the 22C11 antibody binds both Appa and Appb in zebrafish. Unfortunately, preliminary western blot analyses we performed using this antibody to detect protein changes in response to our putative *appb* null mutation were unsuccessful (data not shown). In their paper, Kaiser *et al* (2012) did not provide details of their methods using the 22C11 antibody with zebrafish protein material, hence our preliminary tests utilised conditions that had previously been used in mouse studies. Further testing and adaptation of antibody conditions in zebrafish may lead to more successful identification of the effect of *appb^{R5PfsTer26}* on Appb protein levels.

Nonsense mediated decay (NMD) is a mechanism through which mRNAs harbouring PTCs are rapidly degraded [28]. The single cytosine insertion in the *appb^{R5PfsTer26}* mutant creates a PTC at amino acid position 26, therefore we expect that this transcript would be down-regulated by NMD. To measure the effect of NMD on the *appb^{R5PfsTer26}* transcript, dqPCR could be employed. Allele-specific primers that amplify either mutant or wildtype would be utilised to directly compare transcript expression levels in cDNA synthesised from total RNA extracted from wildtype, heterozygous and homozygous mutant fish brains. If *appb^{R5PfsTer26}* is a true null, we would expect to see decreased, or complete loss of, expression of this allele in heterozygous and homozygous mutants due to NMD of the mutant transcripts. Unfortunately, preliminary attempts to design allele-specific dqPCR primers that could differentiate between our mutant *appb* fish carrying a single cytosine insertion and wildtype were unsuccessful (data not shown). In future, further

primers may be designed and conditions tested in attempts to better differentiate between these two very similar sequences.

Previous experiments in zebrafish that used morpholinos to block *appb* expression resulted in several developmental phenotypes [17, 18]. Despite this, we observed no obvious phenotypical difference in our *appb*^{R5PfsTer26} putative null mutants (heterozygous or homozygous) compared to wildtype. However, the observed lack of phenotype in our mutant *versus* the developmental phenotypes observed for the previously studied morphants was not unexpected, as this phenomenon has now been observed for many genes. Rossi *et al* (2015) showed that phenotypic differences that are now commonly observed between morphants and mutants are due to the phenomenon of "genetic compensation" [19]. In their experiments, genetic compensation was only observed in zebrafish when a loss of protein function was achieved by mutation, not morpholino knockdown. Therefore, it is probable that the lack of phenotype observed in our *appb*^{R5PfsTer26} heterozygous and homozygous fish is due to genetic compensation, most likely by the *appb* co-orthologue *appa*, but also possibly by another member of the large APP family of genes such as *aplp1* or *aplp2* [29].

6-month-old wildtype and heterozygous siblings from an F2 family generated by outcrossing a zebrafish heterozygous for *appb*^{R5PfsTer26} to TU, were treated with either normoxia or hypoxia and their hypoxic response measured by dqPCRs detecting the levels of two HRGs (*edn1* and *igfbp3*). We observed an increase in the transcript levels of both HRGs under hypoxia to the same extent as observed for wildtype fish at the same age. However, we did not observe an increase in the transcript levels of HRGs in the normoxia treated *appb* null mutants, as had been observed

previously for two *psen1* mutants in our laboratory at the same age point [13]. In their study, Newman *et al* (2019) measured the hypoxic response in both wildtype and *psen1* fAD-like mutant zebrafish brains at 24-months-old (which we consider to be an aged brain in zebrafish as the fish are no longer fertile) with some interesting findings. At 6-months-old, their mutant fish showed increased HRG expression under normoxia like that of 24-month-old wildtypes, suggesting that this *psen1* fAD-like mutation accelerates brain aging. Under hypoxia, HRG transcript levels in wildtype fish are increased twofold relative to that of fish remaining under normoxia, while the HRGs of normoxic 24-month-old *psen1* fAD-like mutants are expressed at very similar levels to HRGs in the hypoxia-treated wildtype brains. If 24-month-old *psen1* fAD-like mutant brains are already under hypoxic stress prior to hypoxia treatment, it would be interesting to observe the effect our *appb* putative null on brain aging. As previously mentioned, the APP AICD can directly upregulate HIF-1α [21]. If this interaction is important for the cells ability to respond to hypoxia, we would expect our *appb* putative null zebrafish to have a decreased hypoxic response. Although we did not observe the expected decrease in our experiments presented in this study, the importance of this interaction cannot be discounted for a number of reasons. Firstly, as we measured the hypoxic response in zebrafish brains heterozygous for *appb*^{R5PfsTer26}, there may be enough AICD generated by the second wildtype *appb* allele to sufficiently respond to hypoxia in these brains. This could be further investigated by measuring the hypoxic response in zebrafish brains homozygous for *appb*^{R5PfsTer26}. Secondly, zebrafish have two co-orthologues of the APP gene, *appa* and *appb*. Therefore, it is possible that in our *appb*^{R5PfsTer26} zebrafish, the AICD of Appa can compensate for loss of Appb. Indeed, AICD is highly conserved between human, mouse and both zebrafish APP's, supporting this suggestion. Analysis of the hypoxic response in a double knockout model of zebrafish *appa* and *appb* might be able to resolve this question.

In a previous mouse study, Koike *et al* (2012) observed a higher mortality rate when APP knockout mice were subjected to hypoxia. They suggested that loss of APP, or one of its cleavage fragments, may have increased mortality in these animals by increasing their vulnerability to hypoxia [30]. In their case, complete knockout of APP would mean a complete loss of the AICD, hence observations from their study may provide support to the idea that the interaction between AICD and HIF1-α is critical for the cell to respond to hypoxia. Alternatively, BACE1 expression is also upregulated in response to hypoxia [ref], so it is also possible that one or more β-secretase-specific APP cleavage fragments (such as sAPPβ or even Aβ) play an important role in the hypoxic response. Expression of APP in the vasculature has been observed experimentally [reviewed in 31] and vascular dysfunction in the aging brain has been suggested to be a major risk factor for AD [32]. Although APPs role in the vasculature remains incompletely understood, hypoxia (in the form of ischemia) has been shown to upregulate both the mRNA and protein levels of APP, with subsequent Aβ accumulation [reviewed in 22]. It is therefore possible that APP can act as a stress response protein that protects the brain against hypoxia, although further investigation is needed to confirm this.

Finally, transcriptomic analyses should eventually be performed to identify changes that are specific to an *appb* null mutation and can be excluded from our developing understanding of transcriptomic changes caused by fAD mutations. This will be discussed further in Chapter 6.

4.4 Methods

Ethics

This work was conducted under the auspices of the Animal Ethics Committee of the University of Adelaide and in accordance with EC Directive 86/609/EEC for animal experiments and the Uniform Requirements for manuscripts submitted to Biomedical journals.

Genomic DNA extraction from zebrafish embryos and tail fin clips

A small section of the dorsal fin (~3 mm) of adult fish was removed with a sterile scalpel blade under anaesthesia (Tricaine solution, 1.68 µg/mL).

Embryos, or, tail fin tissue, were placed in 50µl of a 17mg/ml Proteinase K (Roche, Basel, Switzerland), 1XTE solution and then incubated at 55-65°C until dissolved (2-4 hours). The lysis buffer was then placed at 95°C for 5 mins to inactivate the Proteinase K. Either mixture was then centrifuged at 12,000 rpm for 2 mins to pellet cellular debris. The supernatant was transferred into a clean microfuge tube ready for subsequent PCR [33].

Mutation detection assays

Mutations generated by CRISPR or TALEN were detected by the T7 endonuclease 1 method adapted from Reyon *et al* (2012) [34]. In the T7 endonuclease 1 method, loci were amplified using primers presented in supplementary data file S4, Table S1, with GoTaq polymerase (New England Biolabs, Ipswich, Massachusetts, USA) and the resulting DNA was re-annealed under the following conditions; 95°C for 5 mins then ramped down from 95°C-85°C at a rate of -2°C/sec, followed by ramping down from 85°C-25°C at a rate of -0.1°C/sec. After reannealing, 1µl of T7

endonuclease I (New England Biolabs, Ipswich, Massachusetts, USA) was added and the entire mixture was incubated at 37°C for 15 mins.

Sanger sequencing

The genomic region surrounding the *appb* start site sgRNA target site was PCR amplified using *appb* Ex1 forward and reverse primers (supplementary data file S4, Table S1). The purified *appb* 5'UTR/Exon1 genomic DNA fragments were then used for Sanger sequencing with the *appb* Ex1 sequencing F primer (supplementary data file S4, Table S1) (Australian Genome Research Facility, SA, AUS).

Detecting appb^{R5PfsTer26} *mutations via PCR*

We identified a single base pair insertion that resulted in a frameshift at nucleotide position 14 in exon 1 of *appb* (Figure 4.3#). PCR primer pairs were designed to detect either the *appb*^{R5PfsTer26} allele or wildtype allele at this position (supplementary data file S4, Table S1). Loci were amplified with GoTaq polymerase (New England Biolabs, Ipswich, Massachusetts, USA) with PCR conditions varying between primer pairs: The PCR conditions for wildtype-specific (WTF2-R1) detection were 95°C, 2min; 30 cycles of [95°C, 30 s; 64°C, 30 s; and 72°C 40s]; 72°C, 5 min and the anticipated length of the PCR product was 326 bp. The PCR conditions for mutant-specific (F4R2) detection were 95°C, 2min; 30 cycles of [95°C, 30 s; 60°C, 30 s; and 72°C 40s]; 72°C, 5 min and the anticipated length of the PCR product was 618 bp.

Hypoxia treatment and whole brain removal

Oxygen-depleted water was first generated by bubbling nitrogen into recirculated tank water. 6-month-old zebrafish were treated in low oxygen levels (0.9 ± 0.5mg/L of oxygen) or normal oxygen levels (6.6 ±0.5 mg/L of oxygen) by placing them inside a 500ml bottle of the respective water conditions for approximately 3 hours. Directly following hypoxia or normoxia treatment, 6-month-old zebrafish were euthanised in an ice water slurry for ~30 seconds and the brain subsequently removed for immediate RNA extraction.

RNA extraction from whole adult zebrafish brain and cDNA synthesis

Total RNA was isolated from mutant and WT siblings using the *mir*Vana miRNA isolation kit (Thermo Fisher, Waltham, Massachusetts, USA). RNA isolation was performed according to the manufacturer's protocol. First the brain is lysed in a denaturing lysis solution. The lysate is then extracted once with acid-phenol:chloroform leaving a semi-pure RNA sample. A glass-fiber filter is utilised to further purify the sample and return total RNA.

cDNA was generated from 500ng of 6-month-old zebrafish brain RNA using random primers following the First-Strand cDNA Synthesis protocol (Invitrogen, Life Technologies, Carlsbad, CA, USA). After inactivation the cDNA was treated with RNase H (New England Biolabs, Ipswich, Massachusetts, USA) and incubated at 37°C for 30 minutes, the RNase H was then inactivated by heating at 65°C for 20 minutes.

Assaying the hypoxia response by digital PCR

Primers for dqPCR of known hypoxia response genes, including a forward and reverse primer (supplementary data file S4, Table S3) detecting the *igfbp3* gene (NM_205751), and a forward and

reverse primer (supplementary data file S4, Table S3) detecting the *edn1* gene (NM_131519), were previously designed for analysis of these genes in response to hypoxia in the presence of fAD-like mutations. HRG gene transcript levels were measured using the QuantStudio™ 3D Digital PCR System (Life Sciences, Waltham, MA, USA) with the QuantStudio™ 3D Digital PCR 20K Chip Kit v2 and Master Mix (Life Sciences, Waltham, MA, USA, A26317) and SYBR™ Green I Nucleic Acid Gel Stain (Life Sciences, Waltham, MA, USA, S7563). The dqPCR conditions for assays of *edn1* or *igfbp3* expression in response to hypoxia were 96°C, 10 min; 49 cycles of [62°C, 2 min; 98°C, 30 s]; 62°C, 2 min. 50ng of total cDNA from each sample was loaded into one chip for the dqPCR. Chips were read using QuantStudio™ 3D AnalysisSuite Cloud Software (Life Sciences, Waltham, MA, USA).

4.5 References

1. Newman, M., et al., *Robust homeostasis of Presenilin1 protein levels by transcript regulation.* Neurosci Lett, 2012. **519**(1): p. 14-9.
2. Hatchett, C.S., et al., *Familial Alzheimer's disease presenilin 1 mutation M146V increases gamma secretase cutting of p75NTR in vitro.* Brain Res, 2007. **1147**: p. 248-55.
3. Haapasalo, A. and D.M. Kovacs, *The many substrates of presenilin/gamma-secretase.* J Alzheimers Dis, 2011. **25**(1): p. 3-28.
4. Hardy, J.A. and G.A. Higgins, *Alzheimer's disease: the amyloid cascade hypothesis.* Science, 1992. **256**(5054): p. 184-5.
5. Wolfe, M.S., *APP, Notch, and presenilin: molecular pieces in the puzzle of Alzheimer's disease.* Int Immunopharmacol, 2002. **2**(13-14): p. 1919-29.
6. Drachman, D.A., *The amyloid hypothesis, time to move on: Amyloid is the downstream result, not cause, of Alzheimer's disease.* Alzheimers Dement, 2014. **10**(3): p. 372-80.
7. Salminen, A., A. Kauppinen, and K. Kaarniranta, *Hypoxia/ischemia activate processing of Amyloid Precursor Protein: impact of vascular dysfunction in the pathogenesis of Alzheimer's disease.* J Neurochem, 2017. **140**(4): p. 536-549.
8. Pluta, R., et al., *Brain ischemia activates beta- and gamma-secretase cleavage of amyloid precursor protein: significance in sporadic Alzheimer's disease.* Mol Neurobiol, 2013. **47**(1): p. 425-34.
9. Harris, R.A., L. Tindale, and R.C. Cumming, *Age-dependent metabolic dysregulation in cancer and Alzheimer's disease.* Biogerontology, 2014. **15**(6): p. 559-77.
10. Harris, A.L., *Hypoxia--a key regulatory factor in tumour growth.* Nat Rev Cancer, 2002. **2**(1): p. 38-47.
11. Howe, K., et al., *The zebrafish reference genome sequence and its relationship to the human genome.* Nature, 2013. **496**(7446): p. 498-503.

12.   Moussavi Nik, S.H., et al., *The BACE1-PSEN-AbetaPP regulatory axis has an ancient role in response to low oxygen/oxidative stress.* J Alzheimers Dis, 2012. **28**(3): p. 515-30.

13.   Newman, M., et al., *Accelerated loss of hypoxia response and biased allele expression in zebrafish with Alzheimer's disease-like mutations.* bioRxiv, 2019: p. 526277.

14.   Mehravar, M., et al., *Mosaicism in CRISPR/Cas9-mediated genome editing.* Dev Biol, 2019. **445**(2): p. 156-162.

15.   Gasteiger, E., et al., *ExPASy: The proteomics server for in-depth protein knowledge and analysis.* Nucleic Acids Res, 2003. **31**(13): p. 3784-8.

16.   Pedersen, A.G. and H. Nielsen, *Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis.* Proc Int Conf Intell Syst Mol Biol, 1997. **5**: p. 226-33.

17.   Joshi, P., et al., *Amyloid precursor protein is required for convergent-extension movements during Zebrafish development.* Dev Biol, 2009. **335**(1): p. 1-11.

18.   Abramsson, A., et al., *The zebrafish amyloid precursor protein-b is required for motor neuron guidance and synapse formation.* Dev Biol, 2013. **381**(2): p. 377-88.

19.   Rossi, A., et al., *Genetic compensation induced by deleterious mutations but not gene knockdowns.* Nature, 2015. **524**(7564): p. 230-3.

20.   Jeanray, N., et al., *Phenotype classification of zebrafish embryos by supervised learning.* PLoS One, 2015. **10**(1): p. e0116989.

21.   Kaufmann, M.R., et al., *Dysregulation of hypoxia-inducible factor by presenilin/gamma-secretase loss-of-function mutations.* J Neurosci, 2013. **33**(5): p. 1915-26.

22.   Ogunshola, O.O. and X. Antoniou, *Contribution of hypoxia to Alzheimer's disease: is HIF-1alpha a mediator of neurodegeneration?* Cell Mol Life Sci, 2009. **66**(22): p. 3555-63.

23.   Paolicchi, E., et al., *Targeting hypoxic response for cancer therapy.* Oncotarget, 2016. **7**(12): p. 13464-78.

24.   Yoshikai, S., et al., *Genomic organization of the human amyloid beta-protein precursor gene.* Gene, 1990. **87**(2): p. 257-63.

25.   Musa, A., H. Lehrach, and V.A. Russo, *Distinct expression patterns of two zebrafish homologues of the human APP gene during embryonic development.* Dev Genes Evol, 2001. **211**(11): p. 563-7.

26.   Newton, J.R., D. Parkinson, and M.R. Clench, *Strategies for examination of Alzheimer's disease amyloid precursor protein isoforms.* Anal Bioanal Chem, 2006. **385**(4): p. 692-9.

27.   Kaiser, D.M., et al., *Amyloid beta precursor protein and prion protein have a conserved interaction affecting cell adhesion and CNS development.* PLoS One, 2012. **7**(12): p. e51305.

28.   Brogna, S. and J. Wen, *Nonsense-mediated mRNA decay (NMD) mechanisms.* Nature Structural & Molecular Biology, 2009. **16**(2): p. 107-113.

29.   Nicolas, M. and B.A. Hassan, *Amyloid precursor protein and neural development.* Development, 2014. **141**(13): p. 2543-8.

30.   Koike, M.A., et al., *APP knockout mice experience acute mortality as the result of ischemia.* PLoS One, 2012. **7**(8): p. e42665.

31.   d'Uscio, L.V., T. He, and Z.S. Katusic, *Expression and Processing of Amyloid Precursor Protein in Vascular Endothelium.* Physiology (Bethesda), 2017. **32**(1): p. 20-32.

32.   de la Torre, J., *The Vascular Hypothesis of Alzheimer's Disease: A Key to Preclinical Prediction of Dementia Using Neuroimaging.* J Alzheimers Dis, 2018. **63**(1): p. 35-52.

33.   Hruscha, A. and B. Schmid, *Generation of zebrafish models by CRISPR /Cas9 genome editing.* Methods Mol Biol, 2015. **1254**: p. 341-50.

34.   Reyon, D., et al., *FLASH assembly of TALENs for high-throughput genome editing.* Nat Biotechnol, 2012. **30**(5): p. 460-5.

4.6 Supplementary Information

This section is included in the thesis as information supplementary to Chapter 4. It contains additional information not included in the main text of the manuscript.

File S1. Mutation screening
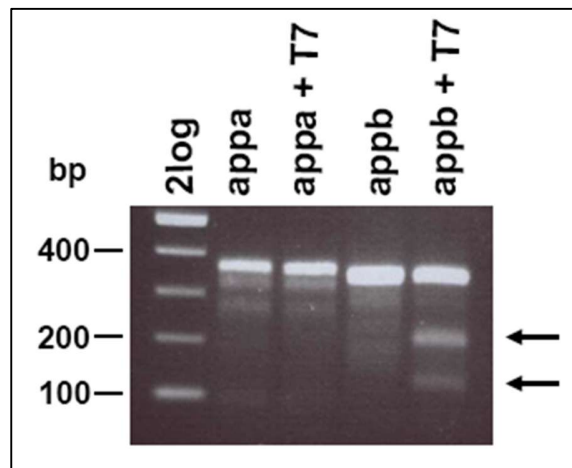
Figure S1. Confirmation of the ability of the* appb *start site sgRNA to generate mutations. *Image shows T7 endonuclease treated genomic DNA from* appb *sgRNA microinjected zebrafish. Primers were used to amplify a region of 345bp with the predicted cleavage site (PAM) positioned so that cleavage products of ~200bp and ~100bp would be generated that could be separated using gel electrophoresis. Fragments observed with sizes of approximately 200pb and 100 bps (black arrows) indicate that the sgRNA was successful in generating mutations at the site of interest.*
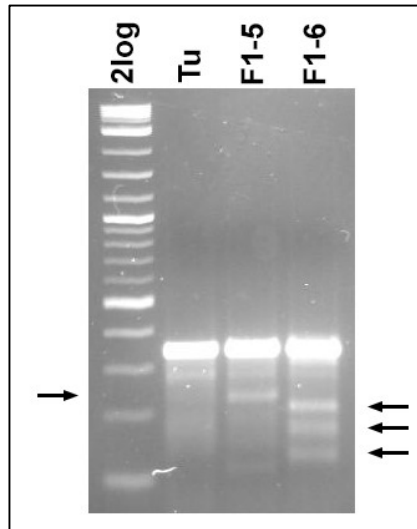
Figure S2. Screening for germline transmission of mutations in the embryos produced from of an outcross of fish F0-5 to wildtype and an outcross of fish F0-6 to wildtype. The 3 arrows on the right indicate cleavage products from a T7 endonuclease assay on DNA amplified across the cleavage site of the appb *sgRNA in embryos from F0-6 outcrossed to wildtype. The single arrow on the left indicates the cleavage product from a T7 assay on DNA amplified across the cleavage site of the* appb *sgRNA in embryos from F0-5 outcrossed to wildtype.*
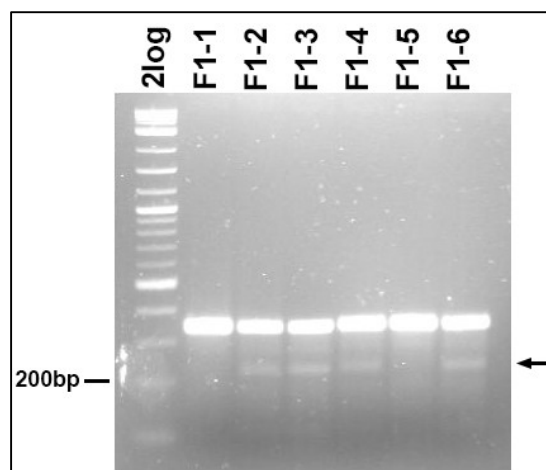
Figure S3. Screening for mutations in the adult progeny of an outcross of fish F0-5 to wildtype. Fish F1-5.2, F1-5.3, F1-5.4 and F1-5.6 carry a potential germline mutation indicated by arrow. Fish F1-5.1 and F1-5.5 are wildtype.
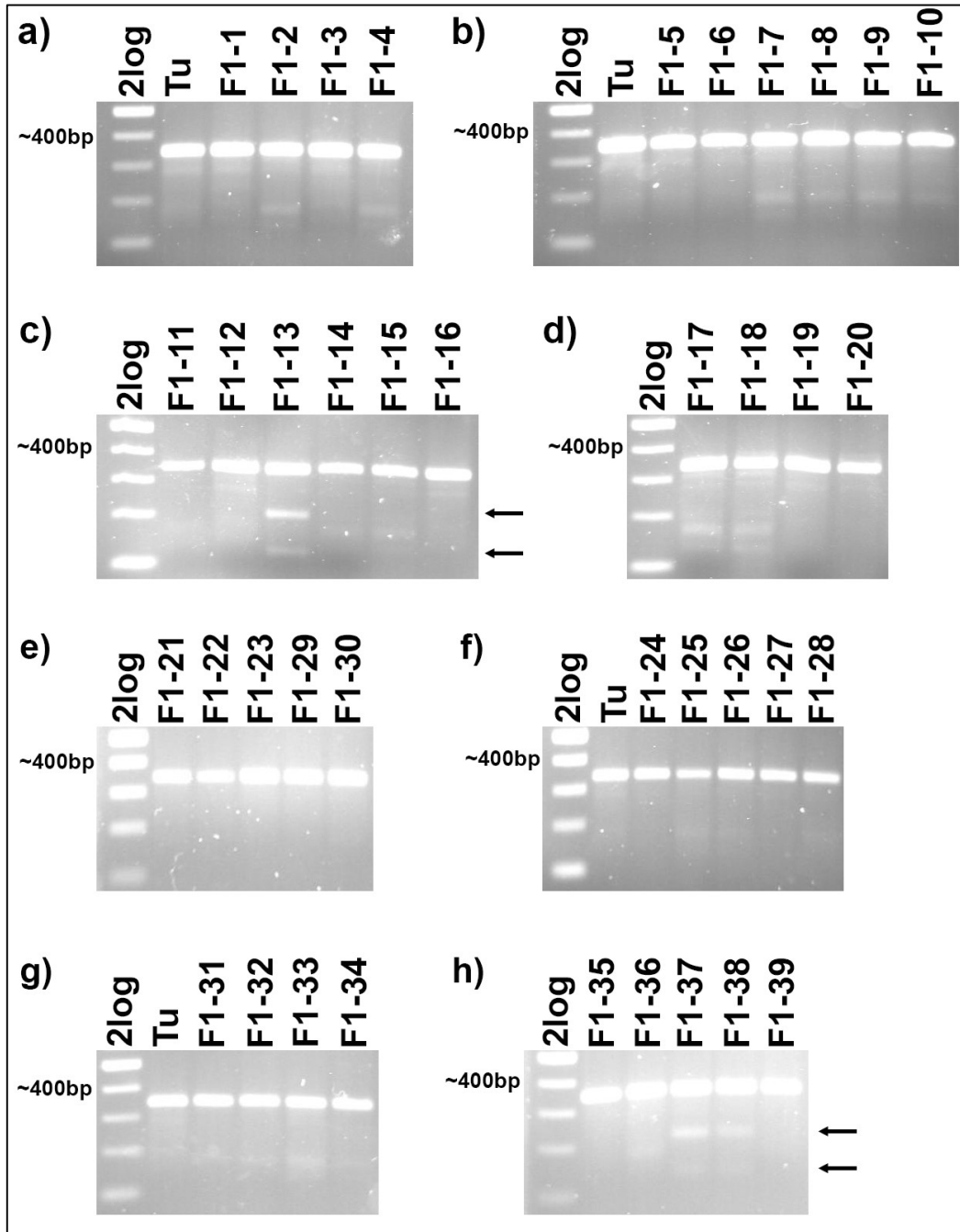
Figure S4. Screening for mutations in the adult progeny of an outcross of fish F0-6 to wildtype. *Fish F0-6.13 (a), F0-6.37 and F0-6.38 (h) carry a potential germline mutation indicated by arrows, while the remaining fish are wildtype.*
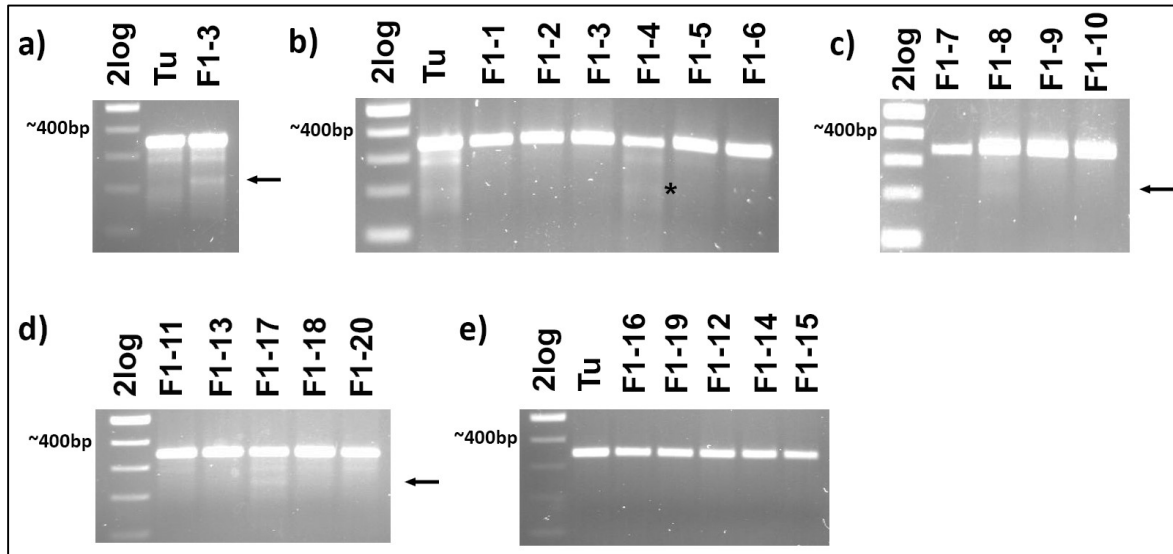
Figure S5. Screening for mutations in the adult progeny of an outcross of fish F0-3 to wildtype. *Fish F0-3.8 (c) and F0-3.17 (d) carry a potential germline mutation (indicated by black arrows), while the remaining fish are wildtype. Black star (fish F0-3.4, b) indicates a faint band that suggests the fish may carry a mutation, but this is inconclusive.*

File S2. Raw sequencing data files
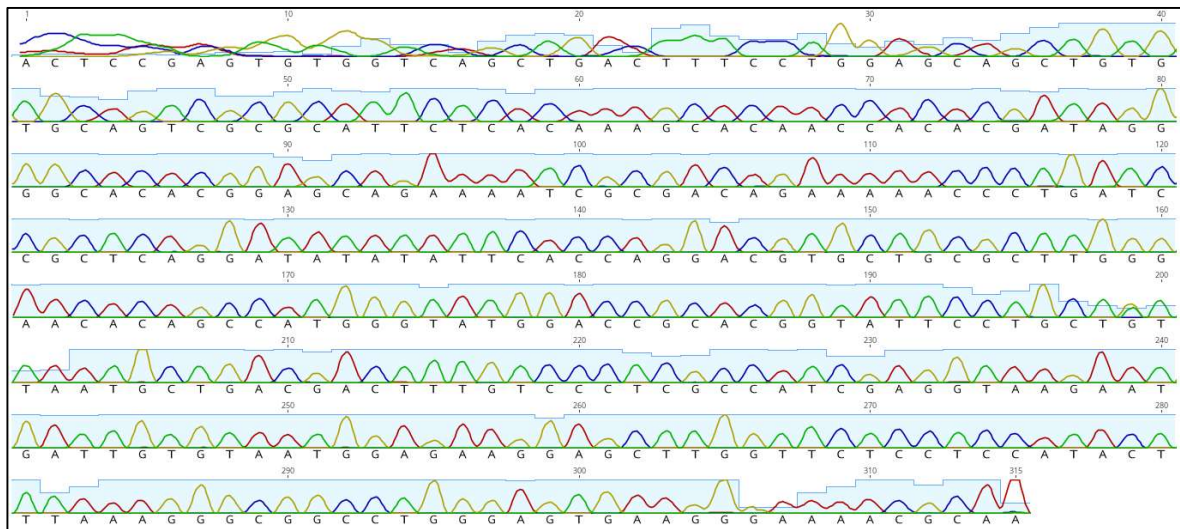
Figure S6. Sanger sequencing of a single adult from the progeny of an outcross of fish F0-5 to wildtype – Single peaks observed throughout the sequence suggest that both strands are wildtype. This same observation was made for sequencing of fish F0-5.2 (this image), F0-5.3, F0-5.4 and F0-5.6 from supplementary data file S1, Figure S3.
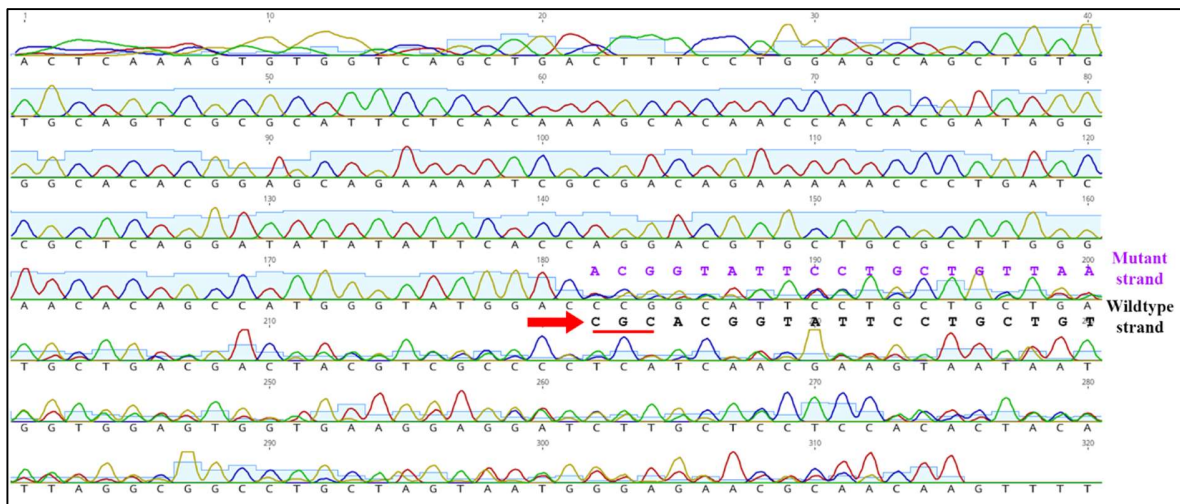
Figure S7. Sanger sequencing of a single adult from the progeny of an outcross of fish F0-6 to wildtype – Single peaks indicate wildtype sequence. Sequence with double peaks indicates a difference in the composition of the two strands (purple and black). Red arrow indicates mutation position. This same

179

observation was made for sequencing of fish F0-6.13 (this image), F0-6.37 and F0-6.38 from supplementary data file S1, Figure S4.
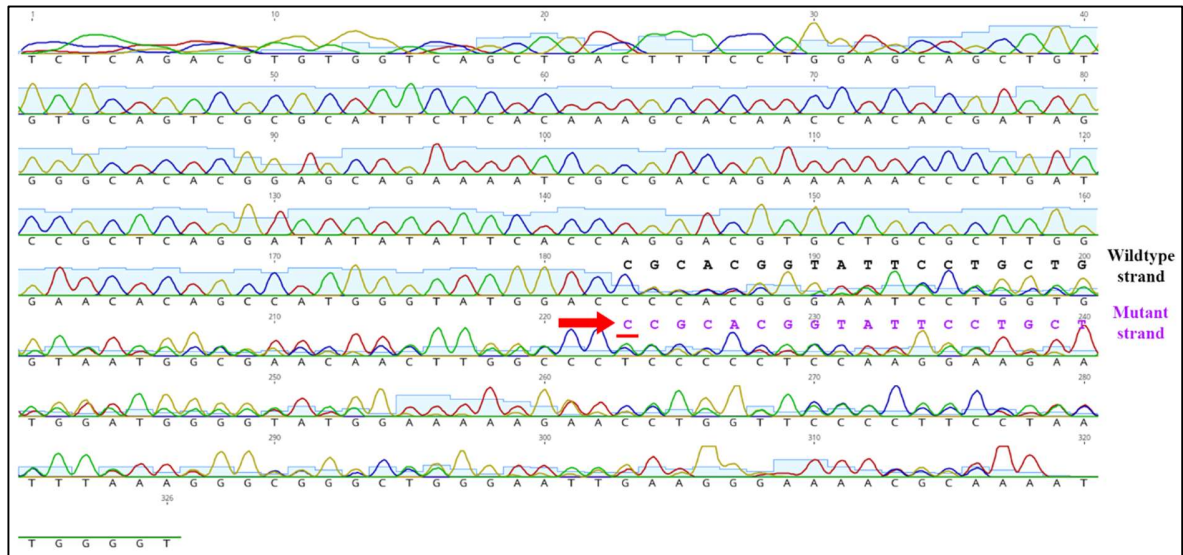
***Figure S8. Sanger sequencing of a single adult from the progeny of an outcross of fish F0-3 to wildtype** – Single peaks indicate wildtype sequence. Scrambled sequence with double peaks indicates a difference in the composition of the two strands (purple and black). Red arrow indicates mutation position. This same observation was made for sequencing of fish F0-3.8 (this image) and F0-3.17 from supplementary data file S1, Figure S5.*

***Figure S9. Outline of** **appb** **3bp deletion mutation identified in progeny of an outcross of fish F0-6 to wildtype.** 3 dashes indicate the 3 nucleotides that are deleted.*

File S3. Investigation of translation and start codons for *appb*^{R5PfsTer26}
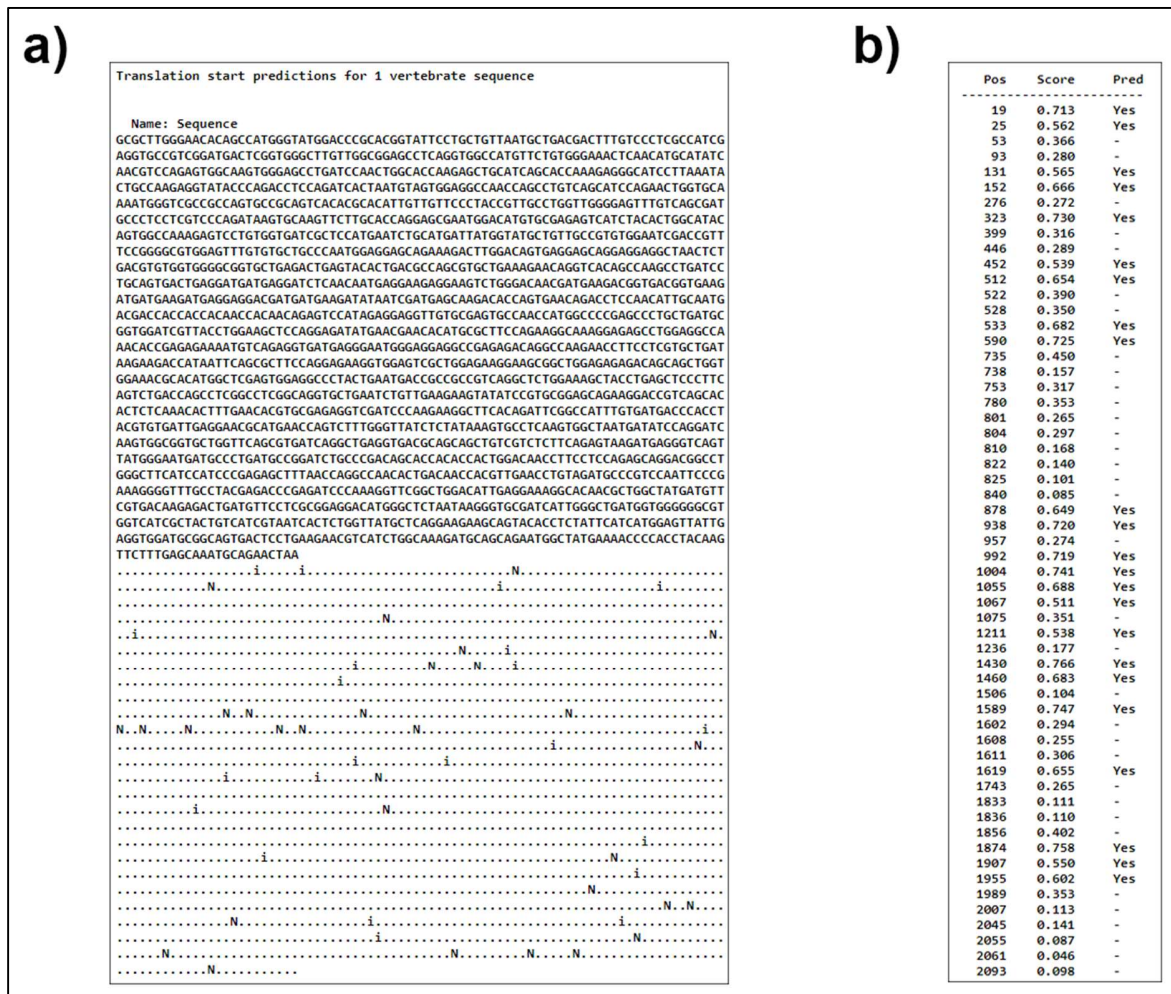
Figure S10. Netstart 1.0 prediction (screenshots) for alternative start codons in R5PfsTer26 mutant. a)
sequence run through software, output is below sequence with start codon (methionine) positions indicated
by "i" if initiation is possible and "N" if not. b) individual scores given to each start codon that could
potentially initiate translation.

a)

5'3' Frame 1

MGMDRTVFLLLMLTTLSLAIEVPSDDSVGLLAEPQVAMFCGKLNMHINVQSGKWEPDPTGTRSCISTKEGILKYCQEVYPDLQITNVVEANQPVSIQNWCKMGRRQCRSHTHIVVPYRCLVGEFVSDALLVPDKCKFLHQERMDMCESHLHW HTVAKESCGDRSMNLHDYGMLLPCGIDRFRGVEFVCCPMEEQKDLDSEEQEEANSDVWWGGAETEYTDASVLKEQVTAKPDPAVTEDDEDLNNEEEEVWDNDEDGDGEDDEDEEDDDEDIIDEQDTSEQTSNIAMTTTTTTTTESIEEVVRV PTMAPSPADAVDRYLEAPGDMNEHMRFQKAKESLEAKHREKMSEVMREWEEAERQAKNLPRADKKTIIQRFQEKVESLEKEAAGERQQLVETHMARVEALLNDRRRQALESYLSSLQSDQPRPRQVLNLLKKYIRAEQKDRQHTLKHFEHVR EVDPKKASQIRPFVMTHLRVIEERMNQSLGYLYKVPQVANDIQDQVAVLVQRDQAEVTQQLSSLQSKMRVSYGNDALMPDLPDSTTPLDNLPPEQDGLGFIHPESFNQANTDNHVEPVDARPIPERGLPTRPEIPKVRLDIEERHNAGYDVR DKRLMFLAEDMGSNKGAIIGLMVGGVVIATVIVITLVMLRKKQYTSIHHGVIEVDAAVTPEERHLAFMQQNGYENPTYKFFEQMQN-

b)

5'3' Frame 1

MGMDPHGIPAVNADDFVPRHRGAVG-LGGLVGGASGGHVLWETQHAYQRPEWQVGA-SNWHQELHQHQRGHP-ILPRGIPRPPDH-CSGGQPACQHPELVQNGSPPVPQSHAHCCSLPLPGWGVCQRCPPRPR-VQVLAPGANGHVRESSTL AYSGQRVLW-SLHESA-LWYAVAVWNRPFPGRGVCVLPNGGAERLGQ-GAGGG-L-RVVGRC-D-VH-RQRAERTGHSQA-SCSD-G--GSQQ-GRGSLGQR-RR-R-R--R-GGR--RYNR-ARHQ-TDLQHCNDDHHHNHNRVHRGGCAS ANHGPEPC-CGGSLPGSSRRYERTHALPEGKGEPGGQTPRENVRGDEGMGGGRETGQEPSSC--EDHNSALPGEGGVAGEGSGWRETAAGGNAHGSSGGPTE-PPPSGSGKLPELPSV-PASASAGAESVEEVYPCGAEGPSAHSQTL-TRA RGRSQEGFTDSAICDDPPTCD-GTHEPVFGLSL-SASSG--YPGSSGGAGSA-SG-GDAAAVVSSE-DEGQLWE-CPDAGSARQHHTTGQPSSRAGRPGLHPSREL-PGQH-QPR-TCRCPSNSRKGFAYETRDPKGSAGH-GKAQRWL-CS -QETDVPRGGHGL--GCDHWADGGGRGHRYCHRNHSGYAQEEAVHLYSSWSY-GGCGSDS-RTSSGKDAAEWL-KPHLQVL-ANAEL

Figure S11. ExPASy translations for primary reading frame of a) wildtype and b) R5PfsTer26 mutant respectively. Red highlight indicates open reading frame.

File S4. Supporting tables.

Table S1. Primers used in sequencing and mutation detection.

- F = forward, R= reverse

Primer name	Primer sequence (5'-3')
Appb Ex1 Sequencing F	CTCCAACCTCCGAGTC
Appb Ex1 F	CTCCAACCTCCGAGTCCATTCT
Appb Ex1 R	TGCGTTTTCCCTTCACTCCC
C insertion F4	CAGGAATACCGTGCGGG
C insertion R2	TATTTGGGTCTCAGGAGGATAA
Appb Ex1 WT F2	CATGGGTATGGACCGCA
C insertion R1 (used with WT F2)	CGCTACACTACGCCTGATACT

Table S2. Mutations detected in the progeny of an outcross of fish F0-3 to wildtype via genotyping PCR

- Fish were genotyped with T7E1 assay up to F1-3.20, then genotyped with the C insertion detecting primers up to F1-3.103.

- Orange shading indicates no mutation, Green shading indicates positive signal for a mutation and Yellow indicates faint bands that might be a signal

F1-3.1	F1-3.2	F1-3.3	F1-3.4	F1-3.5	F1-3.6	F1-3.7	F1-3.8	F1-3.9
F1-3.10	F1-3.11	F1-3.12	F1-3.13	F1-3.14	F1-3.15	F1-3.16	F1-3.17	F1-3.18
F1-3.19	F1-3.20	F1-3.21	F1-3.22	F1-3.23	F1-3.24	F1-3.25	F1-3.26	F1-3.27

F1-3.28	F1-3.29	F1-3.30	F1-3.31	F1-3.32	F1-3.33	F1-3.34	F1-3.35	F1-3.36
F1-3.37	F1-3.38	F1-3.39	F1-3.40	F1-3.41	F1-3.42	F1-3.43	F1-3.44	F1-3.45
F1-3.46	F1-3.47	F1-3.48	F1-3.49	F1-3.50	F1-3.51	F1-3.52	F1-3.53	F1-3.54
F1-3.55	F1-3.56	F1-3.57	F1-3.58	F1-3.59	F1-3.60	F1-3.61	F1-3.62	F1-3.63
F1-3.64	F1-3.65	F1-3.66	F1-3.67	F1-3.68	F1-3.69	F1-3.70	F1-3.71	F1-3.72
F1-3.73	F1-3.74	F1-3.75	F1-3.76	F1-3.77	F1-3.78	F1-3.79	F1-3.80	F1-3.81
F1-3.82	F1-3.83	F1-3.84	F1-3.85	F1-3.86	F1-3.87	F1-3.88	F1-3.89	F1-3.90
F1-3.91	F1-3.92	F1-3.93	F1-3.94	F1-3.95	F1-3.96	F1-3.97	F1-3.98	F1-3.99
F1-3.100	F1-3.101	F1-3.102	F1-3.103					

Table S3. Primers used for digital PCR

Gene	Accession		Forward primer 5'-3'	Reverse primer 5'-3'
edn1	NM_131519	*WT*	CGTTACAGTTTAAAGCAGCGTCA	TGTGTTTGCATTGCTTCCCAG
igfbp 3	NM_205751	*WT*	AGTGCAGTCCATCCATCCAAAGGC	GTCTCCATGTTATAGCAGTGGACCT

Table S4. Expression levels of *edn1* in 50ng total 6-month-old adult brain cDNA. (Assuming complete reverse transcription of total brain RNA).

edn1	
+/+ fish under normoxia (copies/µL)	APPb[R5PfsTer26]/+ fish under normoxia (copies/µL)
402.97	304.18
449.94	460.66
465.88	347.28
534.26	365.84
+/+ fish under hypoxia (copies/µL)	APPb[R5PfsTer26]/+ fish under hypoxia (copies/µL)
449.36	704.68
505.65	522.94
732.53	908.35
940.17	1185.3

Table S5. Expression levels of *igfbp3* in 50ng total 6-month-old adult brain cDNA. (Assuming complete reverse transcription of total brain RNA).

igfp3	
+/+ fish under normoxia (copies/µL)	APPb[R5PfsTer26]/+ fish under normoxia (copies/µL)
735.81	711.64

717.35	725.71
766.96	538.4
809.11	713.48
+/+ fish under hypoxia (copies/μL)	APPb^{R5PfsTer26}/+ fish under hypoxia (copies/μL)
1072.9	1592.8
971.37	1162.2
1094.8	1169.4
1417.2	1376.6

File S5. Classification of F2 generation larvae, from a pair mating of heterozygous F1 *appb^R5PfsTer26* fish, into 3 phenotypic groups (wildtype, mild/moderate and severe).
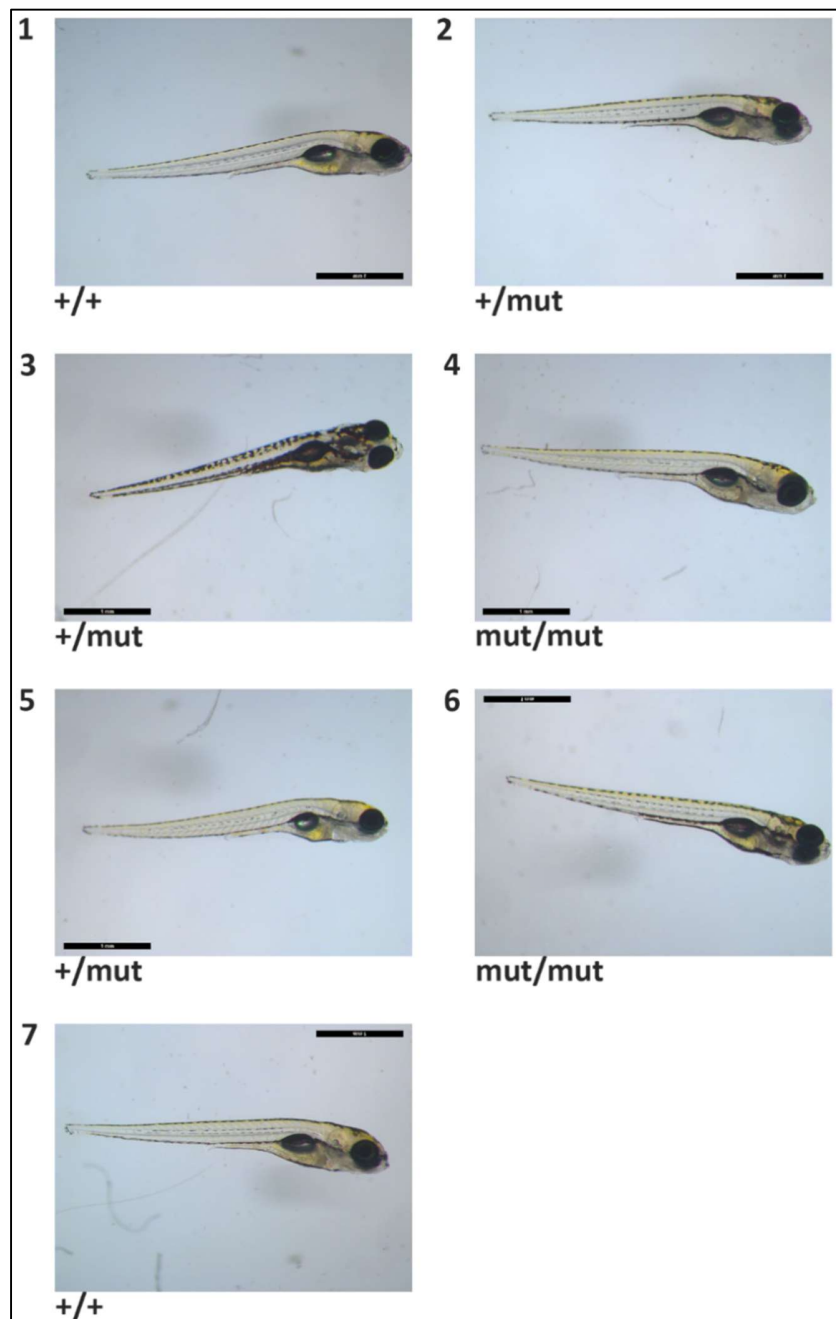
Figure S12. F2 progeny from a pair mating of two heterozygous F1 appb^R5PfsTer26 fish with the wildtype phenotype. *Larvae were subsequently genotyped, and their genotypes are included below each image as follows: +/+ wildtype, +/mut heterozygous and mut/mut homozygous for appb^R5PfsTer26.*
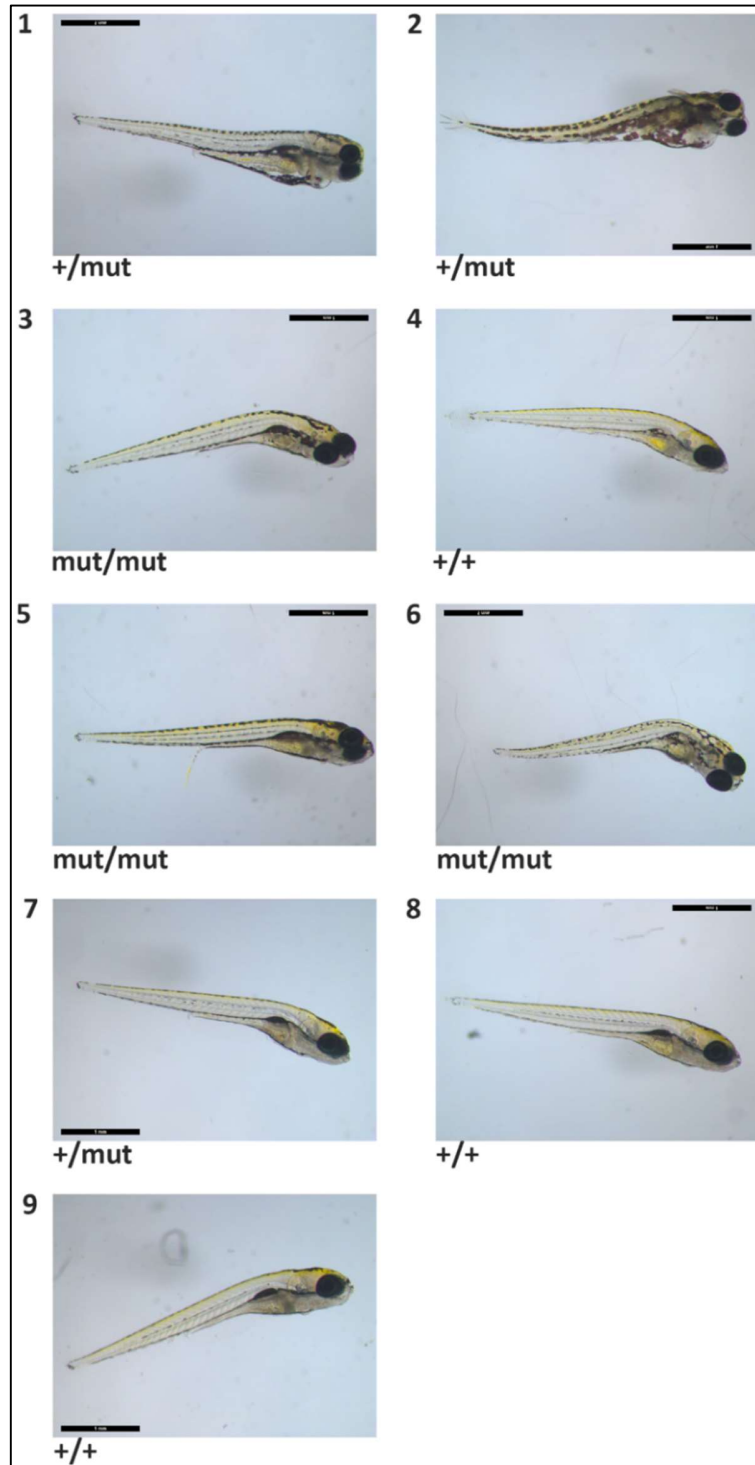
Figure S13. ***F2 progeny from a pair mating of two heterozygous F1 appb^{R5PfsTer26} fish with the mild/moderate phenotype.*** *Larvae were subsequently genotyped, and their genotypes are included below each image as follows: +/+ wildtype, +/mut heterozygous and mut/mut homozygous for appb^{R5PfsTer26}.*
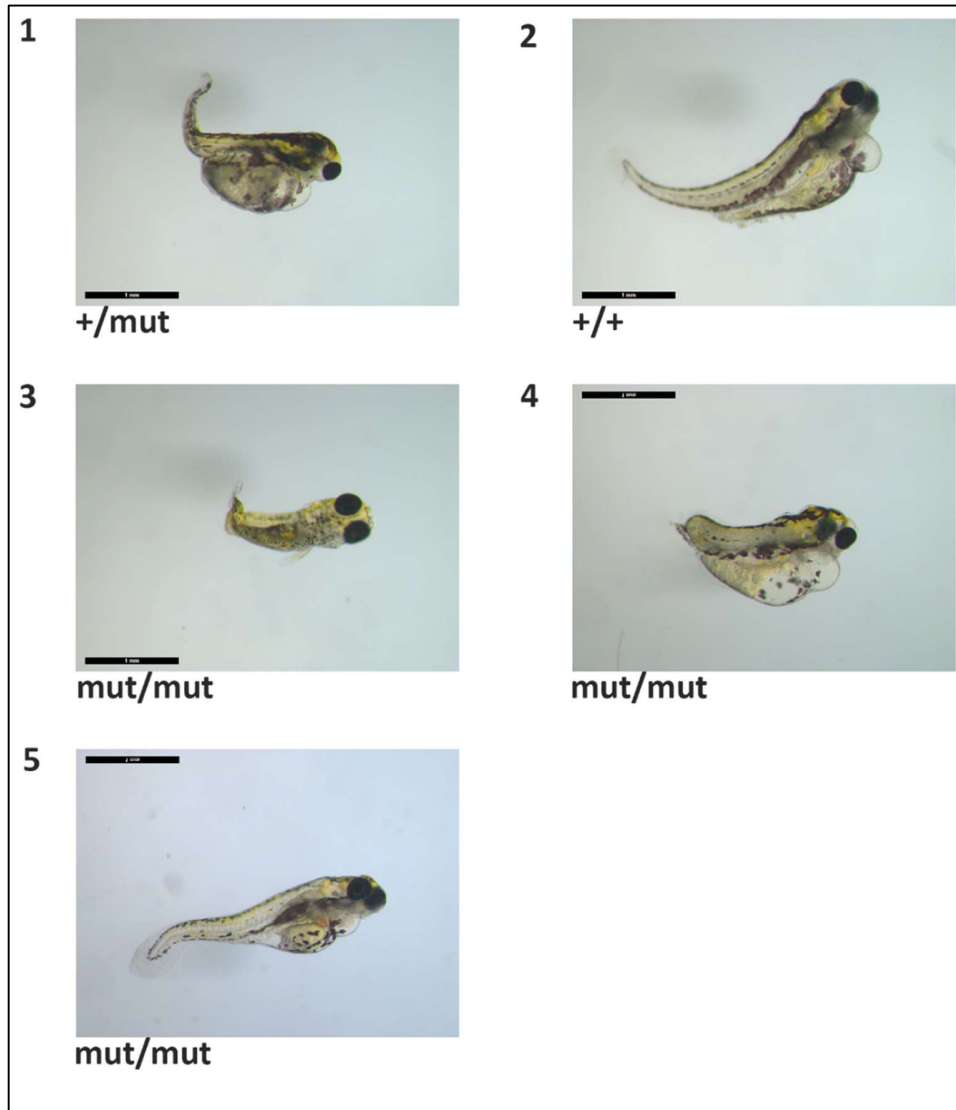
Figure S14. F2 progeny from a pair mating of two heterozygous F1 appb^{R5PfsTer26} fish with the severe

phenotype. *Larvae were subsequently genotyped, and their genotypes are included below each image as*

follows: +/+ wildtype, +/mut heterozygous and mut/mut homozygous for appb^{R5PfsTer26}.

Chapter 5 – Analysing the aggregation propensity of putative zebrafish APPa and APPb Aβ-42-like peptides

5.1 Introduction

The aggregation of proteins or peptides into β-sheets (i.e amyloid fibrils) occurs in several human diseases, such as: Alzheimer's disease (AD) and Parkinson's disease [1]. The amyloid precursor peptide (APP) has a diverse range of functions [2] but is perhaps most well-known for its predicted role in AD neuropathogenesis. APP is processed sequentially by the metalloprotease BACE (aka β-secretase) followed by the γ-secretase complex to release amyloid beta (Aβ) peptide [3]. The Aβ peptide varies in length between 38-43 amino acids, depending on the proteolysis by γ-secretase [4]. The two predominant Aβ species observed in the human brain are the soluble Aβ-40 peptide and the less soluble, more aggregation prone variant, Aβ-42, that is often observed in increased concentrations in AD brains. The Aβ-42 variant is concentrated in the senile/neuritic plaque deposits observed in AD patient's brain parenchyma [reviewed in 5]. Furthermore, mutations within the transmembrane domain of human APP that cause the rare, genetic form of AD, familial AD (fAD), have been demonstrated to increase the Aβ-42 to Aβ-40 ratio, which is thought to be a cause of neuronal cell death in the disease [4]. Both the Aβ-40 and Aβ-42 peptides have been shown to aggregate into amyloid fibrils [6] – which have also previously been defined as "fibrillar polypeptide aggregates that consist of a cross-β structure" [7].

In our laboratory, we have used the zebrafish as a model to study genetic and biochemical processes involved in fAD, such as: hypoxia, autophagy and APP processing. The zebrafish is an excellent model for studying these processes for several reasons. Firstly, genes that carry mutations in humans that are implicated in fAD pathogenesis are conserved in the zebrafish (human *PSEN1*

and *PSEN2* orthologs *psen1* and *psen2*, *APP* co-orthologs *appa* and *appb*, and *SORL1* ortholog *sorl1*) along with other genes involved in the Aβ cascade (other γ-secretase complex component genes *psenen*, *ncstn* and *aph1b*, and β-secretase orthologs, *bace1* and *bace2*) [reviewed in 8]. Zebrafish are also genetically manipulable through microinjection with a range of technologies, including: morpholinos, mRNA, transgenes and genome engineering tools, such as TALENs and CRISPR/Cas. Furthermore, zebrafish have neuroanatomical similarity with the human brain, fAD behavioural similarity with humans, and pathophysiological resemblance to fAD [reviewed in 9]. Previous research has investigated the Aβ peptide in zebrafish in a variety of ways. Donnini *et al* (2010) administered human-derived Aβ peptides directly to the water containing developing zebrafish embryos to examine toxicity [10]. Cameron *et al* (2012) treated zebrafish embryos with monomeric human Aβ and observed induced branching in cerebrovascular blood vessels in the zebrafish hind brain [11]. Newman *et al* (2010) fused the sequence of human Aβ-42 to the *mitfa* gene promoter and utilised this to generate a zebrafish transgenic toxicity model of Aβ [12]. Bhattarai *et al* (2016) generated an Aβ-42 toxicity model in zebrafish by injecting human Aβ-42 linked or unlinked to two peptides with the ability to penetrate cells, transportan and poly-arginine [13]. Sivaji *et al* (2019) injected human Aβ-42 peptide into the midbrain region of 24 hour old zebrafish, to study the relationship between Aβ and osteogenesis over Sox9a expression during early development [14].

To our knowledge there are no published studies that have investigated aggregation of endogenous Aβ peptides that would theoretically be produced from sequential processing of Appa or Appb in zebrafish. The Aβ peptide region, along with the secretases that produce Aβ from APP, are conserved in zebrafish. Also, it has previously been demonstrated that zebrafish *psen1* (γ-

secretase) is capable of processing human APP to secrete the Aβ-42 peptide, suggesting that this role of *psen1* (γ-secretase) is conserved in the zebrafish [15]. Despite this, there has not yet been identification of a zebrafish-specific Aβ peptide, nor has it been established whether the zebrafish equivalent β-secretase and γ-secretase enzymes process endogenous *appa* or *appb* as they do human APP [16]. One study performed sequence alignments of APP across all animal taxa and observed that the "normal" β-secretase cleavage site (EVKMDAE) is not completely conserved in zebrafish Appa or Appb [17]. However, this does not necessarily mean that zebrafish Appa and Appb will not be cleaved by β-secretase as; 1) the amino acid residues in zebrafish Appa and Appb that do not share conservation with the human APP sequence are still highly similar in chemical structure, and 2) it has previously been demonstrated that mutation of the β-secretase site in a Swedish fAD family (from EV**KM**DAE to EV**NL**DAE [18]) creates a peptide sequence that is actually preferentially cleaved by β-secretase [19]. One gap in the knowledge of presenilin function in zebrafish, in the context of generating fAD-like mutations in the equivalent endogenous zebrafish genes, is our lack of understanding of (or knowledge of the existence of) zebrafish-specific Aβ peptides. In order to utilise our existing fAD-like mutant zebrafish to study aggregation in response to these mutations, we first need to investigate a) whether zebrafish-specific Aβ peptides are released, and b) whether these peptides aggregate in a similar manner to human Aβ peptides.

In this study, we aimed to determine whether the putative Aβ-42 equivalent peptides from zebrafish Appa and Appb would aggregate with a similar propensity to that of human Aβ-42. We first predicted Aβ-42-like peptides for zebrafish Appa and Appb. We then analysed these Aβ-42-like peptides using three forms of different aggregation/fibril formation propensity predicting

software: TANGO, Zyggregator and Zipper DB. During these analyses, we observed two distinct regions of the human Aβ-42 peptide that are prone to fibril formation, one in the mid-peptide region and another at the C-terminal end of the peptide sequence. When comparing the predicted Aβ-42-like peptides of zebrafish Appa and Appb to the human Aβ-42 peptide, our analyses revealed similar aggregation propensities in the C-terminal ends across all peptides. However, we observed that the predicted Aβ-42 equivalent peptide of zebrafish Appb had distinctly lower, or even non-existent, aggregation propensity in the mid-peptide region. In contrast, the predicted Aβ-42 equivalent peptide of zebrafish Appa showed an aggregation propensity in this mid-peptide region similar to that of human Aβ-42. These findings suggest that both the Aβ-42-like peptides of zebrafish Appa and Appb have similar aggregation potentials to human Aβ-42. However, further investigation is required 1) to determine whether the lack of aggregation potential observed in the mid-peptide region of the predicted Aβ-42-like peptide of zebrafish Appb prevents it from forming amyloid fibrils, and 2) to demonstrate that the Aβ-42-like peptide of Appa does aggregate as predicted by the analyses in this chapter.

5.2 Results and Discussion

Determining Aβ-42-like peptide equivalents in zebrafish Appa and Appb

To investigate the amyloidogenic propensity of zebrafish-equivalent Aβ-42 peptides, we first identified Aβ-42 equivalents by performing a protein sequence alignment with human and mouse APP. Alignment of zebrafish Appa and Appb with human and mouse APP revealed the equivalent Aβ-42 peptide regions in these genes (Figure 5.1). While the Aβ-42-like peptide in zebrafish Appb maintained the same length as mouse and human Aβ-42, the zebrafish Appa peptide is only 39 amino acids in length due to three missing amino acids near the N-terminal end of the polypeptide

chain. The predicted amyloidogenic propensies of the identified zebrafish Appa and Appb Aβ-42-like peptides were then investigated using a variety of software tools, whose algorithms employ different methods to identify aggregation prone peptide regions, namely: TANGO, Zyggregator and ZipperDB.

Figure 5.1. Alignment of Aβ-42 peptides of human and mouse and Aβ-42-like peptides of zebrafish Appa and Appb used in aggregation analyses. Green boxes indicate regions of sequences that are identical across all peptides. The three amino acid residues that are different between mouse and human (and the equivalent amino acid residues in zebrafish Appa and Appb) are indicated by red boxes.

Investigating amyloidogenic propensity of zebrafish Aβ-40 and Aβ-42 using TANGO

The TANGO algorithm tests whether residues in a peptide sequence of interest occupy a β-aggregation conformation. One of the definitions of amyloid fibrils is that they consist of a cross-β structure [7], hence this method should give an indication of the propensity to form these fibrils. Only sections of five or more consecutive amino acid residues in the β-aggregation conformation are included in TANGO's analysis. TANGO has previously been used to predict the β-sheet aggregation propensity of Alzheimer's Aβ-40 and Aβ-42 variants [20]. TANGO identified two aggregation prone regions for both peptides: one region comprised of residues 17-21 in the middle of the peptide and a second region towards the C-terminus of the peptide containing residues 31-36 of Aβ-40 and residues 30-42 of Aβ-42. Interestingly, they found that Aβ-42 had a much higher

propensity to aggregate between residues 30-42 than Aβ-40 between residues 31-36 [20]. They also compared *in vivo* experimentally determined values of Aβ aggregation for four fAD mutations (Arctic, Dutch, Italian and Flemish) to predictions of β-sheet aggregation propensity using TANGO and found that three of the four predictions closely resembled the experimentally derived data thereby validating their algorithm [20].

Table 5.1. Combinations of physio-chemical conditions input into TANGO software				
Experiment	pH	Temperature (K)	Ionic Strength	Concentration
1	7.2	310.15	0.3	$9x10^{-12}$
2	7	298.15	0.3	$9x10^{-12}$
5	7.2	310.15	0.2	$9x10^{-12}$
6	7	298.15	0.2	$9x10^{-12}$
9	7.44	298.15	0.2	$9x10^{-12}$
10	7.44	298.15	0.3	$9x10^{-12}$
Notes: The experiments presented here are a subset of the experiments conducted. For a complete table and raw data tables from all experiments see supplementary data file S1, Table S1.				

In the zebrafish, APP exists as two co-orthologs, *appa* and *appb* [8], whereas in humans and mice there is only a single *APP* gene. We therefore analysed our predicted Aβ-42-like peptides for both zebrafish Appa and Appb, along with human and mouse Aβ-42, using the TANGO software under varying conditions. TANGO allows for input variations in pH, temperature, ionic strength and

concentration. Each of these parameters have been shown to affect protein folding or misfolding [20] and, therefore, their effect on these processes has been accounted for in TANGO's algorithm. The Aβ-42 peptides were each run through the TANGO software program using conditions laid out in Table 5.1. pH, ionic strength and concentration values for human and zebrafish were identified from the literature. For human pH conditions, we wished to use the pH of the brain as this is where the Aβ-42 peptide's function in fAD is predominantly studied. Orlowski *et al* (2011) stated the pH of human brain to be 7.2, hence we used this value in our analyses [21]. The *in vivo* pH for the brain ventricle of 5-day-old zebrafish larvae has previously been reported as ~7.0-7.44, hence we compared both the upper and lower limit of this range in our analysis [22]. For the temperature settings, we used values known to represent the standard *in vivo* temperature for humans (310.15K, 37°C) and zebrafish (298.15K, 25°C). The concentration value was derived from a study by Wang *et al* (1999), that measured Aβ-42 levels in pathological aging brain at 9 pmol/g wet tissue [23]. Animals maintain an intracellular ionic strength of 200-300 nM, therefore, we used the upper and lower limits of these values in our analyses [24]. The specific combinations of conditions presented in Table 5.1 were selected since we wanted to understand: a) whether each peptide would aggregate (and if these values would differ) with zebrafish or human-specific temperatures and pH, and b) whether the aggregation propensity would change with differing intracellular ionic strengths.
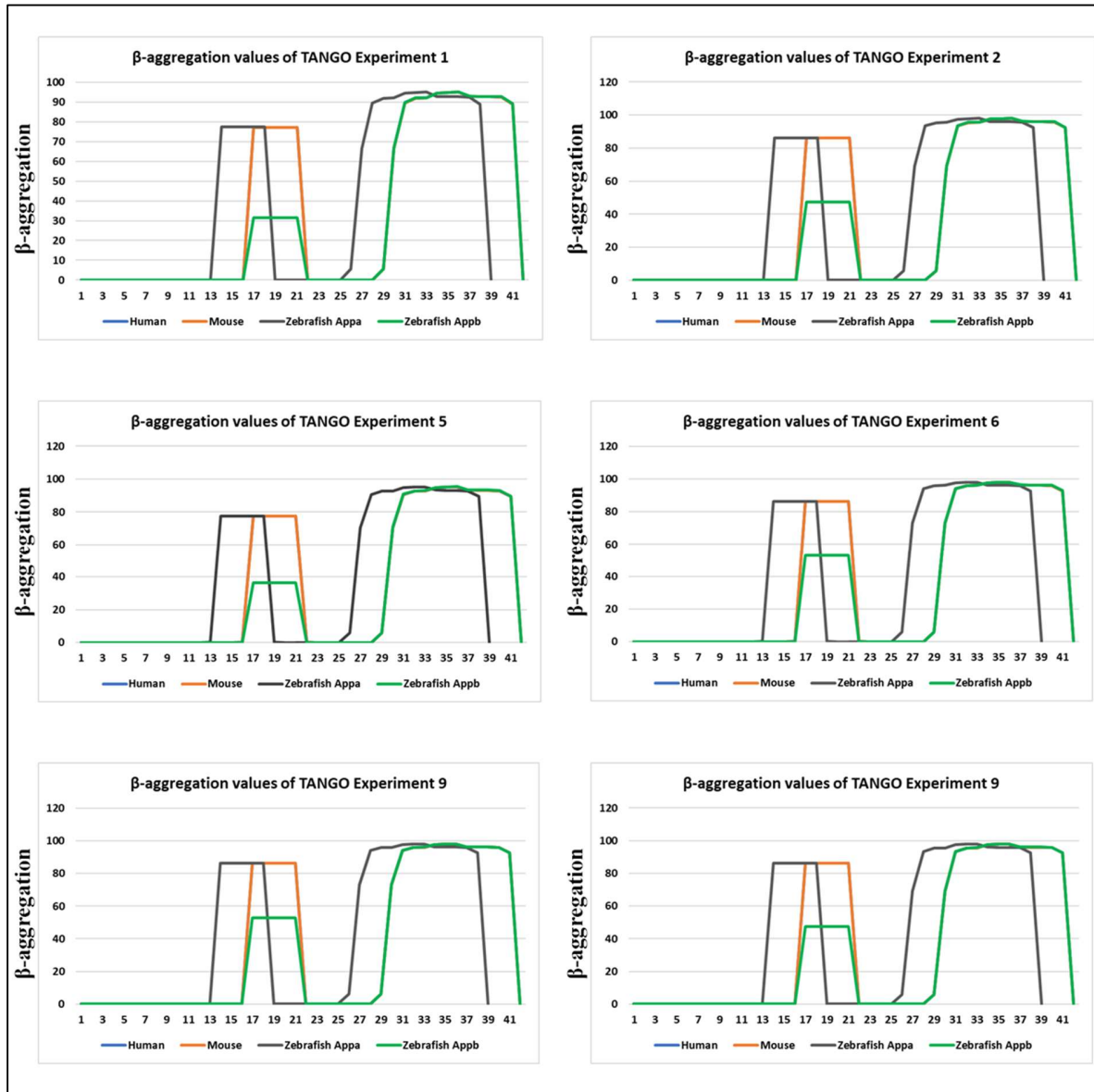
Figure 5.2. TANGO predictions of the aggregation propensity of Aβ-42 peptides for human, mouse, zebrafish Appa and zebrafish Appb under each set of experimental conditions (presented in Table 5.1). The blue line indicating aggregation propensity of human Aβ-42 is obscured by the mouse Aβ-42 line (orange) as they have highly similar values. Both human and mouse Aβ-42 lines are obscured in most sections by the zebrafish Appb line (green). Numbers corresponding to positions of the amino acid residues in each Aβ-42 peptide are on the x-axis.

During our analyses of the four Aβ-42 peptides using TANGO, we observed two regions of high β-aggregation propensity in each (Figure 5.2). As with the Aβ-40 and Aβ-42 study described above, the first region encompassed residues 17-21, and the second residues 30-41 for the Aβ-42 peptides of human APP, mouse APP and zebrafish Appb. For the shorter, zebrafish Appa Aβ-42-like peptide, the first region was comprised of residues 14-18, while the second region was composed of residues 27-38. Interestingly, while the Aβ-42 peptides of human, mouse and zebrafish Appa all returned highly similar β-aggregation scores in the mid-peptide region, the Aβ-42 peptide of zebrafish Appb returned distinctly lower scores in this region for all experimental conditions tested. Despite this, all four peptides returned highly similar scores in the C-terminal region across all conditions investigated. Finally, each of the peptides run through the varying conditions gave what the TANGO authors term an "Agg" score. Agg is a parameter indicating β-sheet aggregation propensity of the whole peptide being analysed, rather than at each section of 5 amino acids [20]. For each of the conditions tested, the Aβ-42-like peptide of zebrafish Appb had a lower Agg score than that of human, mouse or zebrafish Appa Aβ-42 equivalents (supplementary data file S1, Table S8). Taken together, the results of our TANGO analysis suggest that, while zebrafish Appa is three amino acids shorter than the human or mouse Aβ-42 peptide, it has a much more similar β-aggregation propensity than that of zebrafish Appb.

Investigating amyloidogenic propensity of zebrafish Aβ-40 and Aβ-42 using Zyggregator

To confirm the observations made using the TANGO software, we decided to utilise a similar aggregation predicting software, Zyggregator [1]. While TANGO predicts the β-aggregation propensity of peptides, Zyggregator calculates a Z_i^{agg} score at each position of the peptide sequence from each specific amino acid's physio-chemical properties. This can be otherwise described as a

prediction of the most important peptide sequence regions for promotion of aggregation and amyloid fibril formation [25]. The three physio-chemical properties Zyggregator calculates as a part of its algorithm are: hydrophobicity, charge, and the propensity to adopt α-helical or β-sheet structures [1]. Importantly, if a Z_i^{agg} score of 0 is recorded this suggests that the aggregation propensity at that position equals that of a random sequence (i.e. not likely to aggregate), whereas, if a score of 1 is recorded it is considered aggregation prone.

Tartaglia and Vendruscolo (2008) previously used their Zyggregator algorithm to predict the aggregation propensity of the human Aβ-42 peptide [1]. Similar to what was observed using the TANGO algorithm, Zyggregator revealed two regions with a high propensity to aggregate (identified by Z_i^{agg} values of >1, Figure 5.3). Also like the TANGO study, one region was identified in the middle of the peptide between residues 18-22 and another at the C-terminal end of the peptide between residues 32-42. Our analysis of human Aβ-42 using the Zyggregator algorithm complemented the Tartaglia and Vendruscolo study. We also analysed mouse Aβ-42 and the predicted Aβ-42-like peptides for zebrafish Appa and Appb. While the zebrafish Appa Aβ-42-like peptide is shorter than the others, the Z_i^{agg} profile of this peptide closely resembled that of human Aβ-42, with two regions identified: one at position 15-19 and the other at position 29-39 (Figure 5.3). Worthy of note is that the mouse Aβ-42 Z_i^{agg} profile is less similar to the human Aβ-42 profile than zebrafish Appa, with an additional aggregation prone residue close to the N-terminus (which is not represented in any other profile) and a slight shift in the positions of amino acid residues with high aggregation propensity in the mid peptide region (Figure 5.3). As was previously observed for the zebrafish Appb Aβ-42 peptide using the TANGO algorithm, its Z_i^{agg} profile showed only one region, at position 32-42, predicted to have high aggregation (Figure 5.3).
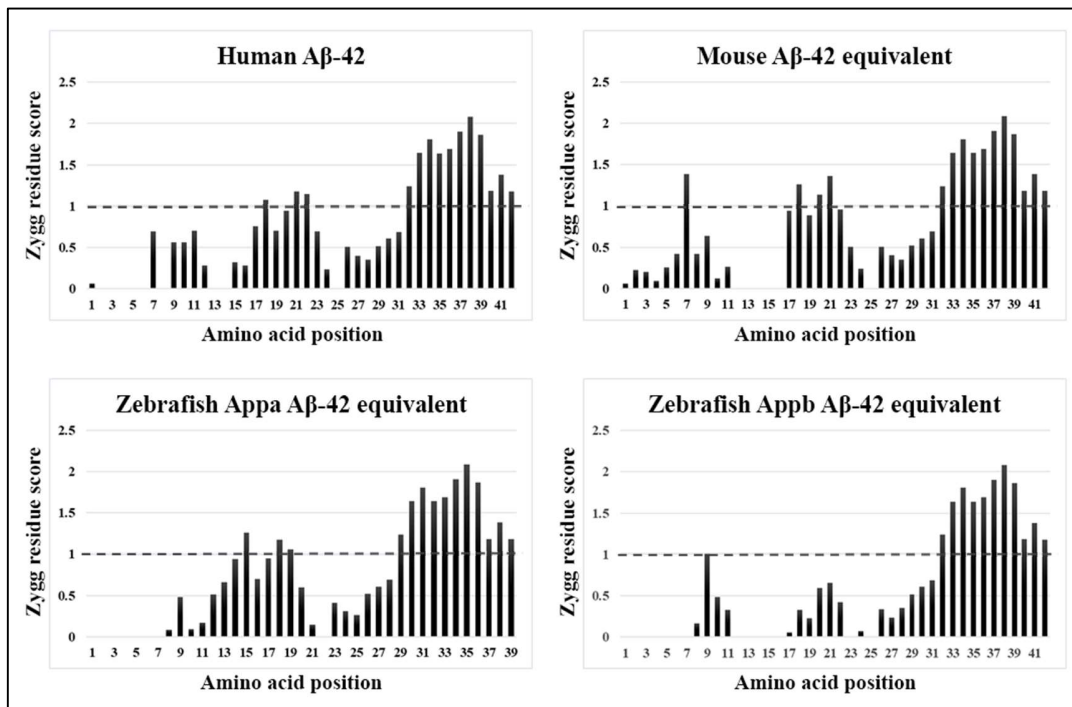
Figure 5.3. Aggregation propensity profiles of human, mouse, zebrafish Appa and zebrafish Appb Aβ-42 equivalent peptides. *Vertical lines show the intrinsic aggregation propensity profile, Z_i^{agg}. The horizontal dashed line indicates the threshold which the Z_i^{agg} score must reach for the amino acid residue at that position to be considered to have high aggregation propensity.*

Investigating amyloidogenic propensity of zebrafish Aβ-40 and Aβ-42 using Zipper DB

While the previously discussed algorithms, TANGO and Zyggregator, both rely purely on sequence-based information to calculate aggregation/fibril formation potential, Zipper DB utilises structural information to assess the possibility that a peptide will form fibrils. The Zipper DB algorithm focuses on the factors that permit a protein to convert to the amyloid state, by attempting to identify "steric zipper" forming segments [25]. In a study investigating their newly developed method, Goldschmidt *et al* (2009) validated their approach using pancreatic ribonuclease A (RNase A) as a model system [25]. The Zipper DB algorithm measures fibril formation potential

based on the RosettaDesign potential energy function, with low Rosetta energy numbers indicating high fibril forming propensity.

The authors established what they termed an "energetic threshold for high fibrillation propensity" (HP), where high fibrillisation potential is predicted for Rosetta energies of less than -23 kcal/mol. In our study, we used Zipper DB to assess the amyloid fibril forming propensity of human and mouse Aβ-42 along with our predicted Aβ-42-like peptides in zebrafish Appa and Appb. The Rosetta energies predicted for hexapeptides in each of these peptide sequences are presented in supplementary data file S3, Table S10.

Zipper DB analysis of the predicted zebrafish Appa and Appb Aβ-42 peptide equivalents revealed fibrillisation propensity profiles similar to the established aggregation prone human Aβ-42 peptide. Once again, Zipper DB identified two regions that appear to have high aggregation propensity, as was observed for TANGO and Zyggregator. For human, mouse and zebrafish Appa Aβ-42 sequences, the regions that are predicted to have high fibrillation propensity are mostly at the C-terminus of the peptide (hexapeptides 27-31, 33 and 35-37 for human and mouse, and hexapeptide 24-28, 30 and 32-34 for zebrafish Appa). Complementary to our findings from the analyses with TANGO and Zyggregator, the profile for the zebrafish Appa Aβ-42-like peptide most closely resembled that of human Aβ-42, even though its shorter length of 39 amino acids resulted in less hexapeptides analysed in this region. Also complementary to our previous two analyses, while the fibrillisation propensity profile of zebrafish Appb resembles human Aβ-42 at the 25th-37th hexapeptide positions, it lacks the aggregation prone hexapeptide that exists at position 15 in human and mouse profiles and position 12 in zebrafish Appa (Figure 5.4, Table S10). This

difference in its fibrillisation propensity profile may suggest that the zebrafish Appb Aβ-42-like peptide is less likely to aggregate. Interestingly, a separate analysis of the human APP Aβ-40 peptide with Zipper Db revealed a similar fibrillisation propensity to human, mouse and zebrafish Appa Aβ-42 in the mid-peptide (15th hexapeptide). However, due to it being two peptides shorter at the C-terminal end, it is missing the hexapeptides with high aggregation propensity at positions 36 and 37 (supplementary data file S4, Figure S1). Also worth noting is that, while there are several hexapeptides that are close to reaching the -23 kcal/mol threshold in the N-terminal region of both the human Aβ-42 peptide and zebrafish Appa Aβ-42-like peptide, the mouse Aβ-42 peptide hexapeptides have much higher energies in this region (and thus have less aggregation potential).
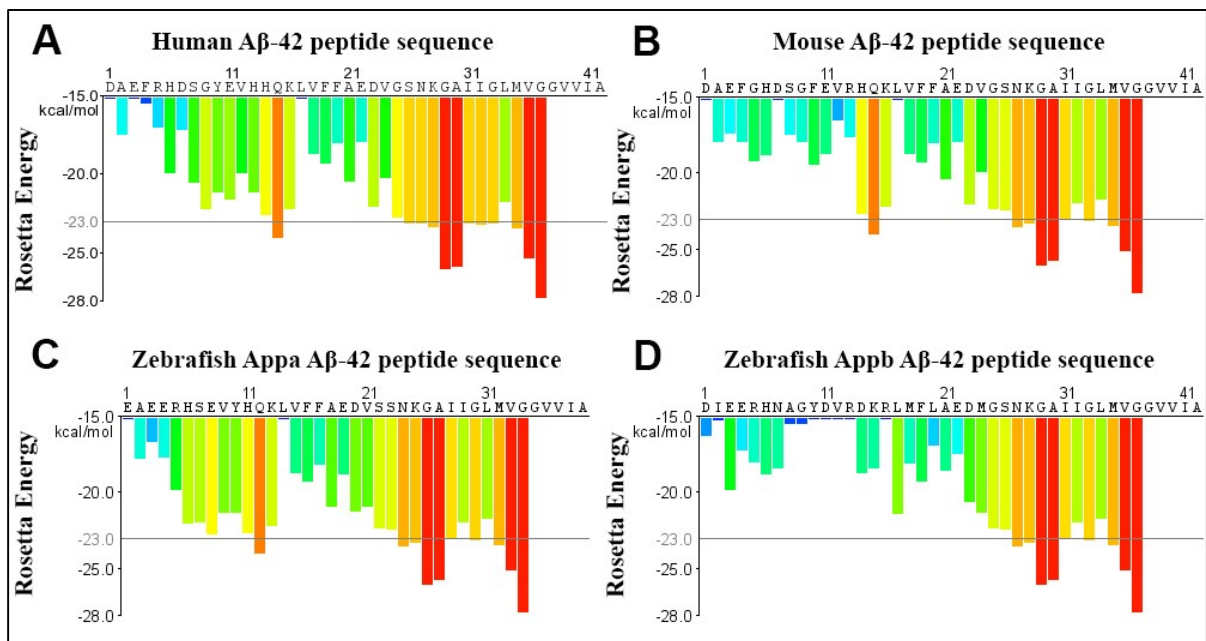
Figure 5.4. Zipper DB fibrillation propensity profiles of Aβ-42 peptide regions. *– A) human Aβ-42 peptide, B) mouse Aβ-42 peptide, C) zebrafish Appa predicted Aβ-42 peptide and D) zebrafish Appb predicted Aβ-42 peptide. Red histogram bars correspond to hexapeptides with energy < −23 kcal/mol and are predicted to form fibrils.*

5.3 Conclusions and future directions

In this study, we have predicted the Aβ-42-like peptides for the zebrafish co-orthologs, Appa and Appb. We performed aggregation analyses of these peptides using several software programs, namely; TANGO, Zyggregator and Zipper DB. Our analyses demonstrate that the zebrafish Appa Aβ-42 equivalent has a very similar aggregation propensity to that of the known aggregation-prone human Aβ-42 peptide. Interestingly, while the zebrafish Appb Aβ-42 equivalent peptide was observed to have comparable aggregation potential at its C-terminal end, each of the analyses suggested that the middle of this peptide did not have the same aggregation propensity as that of human and mouse Aβ-42, or the predicted Appa zebrafish Aβ-42-like peptide. In our TANGO analysis, this mid-peptide region had a much lower β-aggregation score, suggesting it was less likely to aggregate, whereas the analyses with Zyggregator and Zipper DB showed no aggregation propensity in this region. Although a region of high aggregation propensity appears to be missing in the zebrafish Appb Aβ-42-like peptide, it is still possible that it may form similar aggregates to the human Aβ-42 peptide. Further testing using *in vitro* analyses could confirm this. Another interesting observation from the Zipper DB analysis is that of the high fibrillisation propensity profile for human Aβ-40, similar to that of human Aβ-42, with the exception of the final two amino acid residues. Human Aβ-40 has long been thought to be less aggregation prone than the longer Aβ-42 variant, although more recent studies have demonstrated that Aβ-40 fibrils have common structural principles to Aβ-42 fibrils [6]. However, numerous other studies have demonstrated that the difference in the C-terminal regions of Aβ-40 or Aβ-42 results in variable aggregation mechanisms, suggesting that this C-terminal region is important in the dimer formation process of aggregation [26]. Considering this, we can speculate that perhaps the mid-peptide region, which is conserved between human Aβ-42, mouse Aβ-42 and Appa Aβ-42-like peptides, also has some

important (currently unknown) role to play in the aggregation process, making it possible that the Aβ-42-like peptide of Appb would not aggregate in the same way as the other Aβ-42 peptides. It is also worth noting that, while our TANGO analyses were performed using ionic strength values reflecting the zebrafish intracellular environment (200-300nM), peptide aggregation can also occur in the extracellular space where ionic strengths are typically in the mM range [27, 28]. In future, additional TANGO analyses could be conducted utilising ionic strength values representative of the extracellular environment. This should allow us to determine whether the aggregation propensities of these peptides in extracellular conditions remain consistent with the results of our analyses. The TANGO analysis could also potentially be improved by utilising a scrambled Aβ-42 peptide sequence to investigate whether random peptides with equivalent sequence composition are equally as aggregation prone.

Another interesting observation from our analysis was the differences between Mouse Aβ-42 profiles and human Aβ-42 profiles in the early-mid peptide region for both the Zyggregator and Zipper DB software analyses. Interestingly, zebrafish Appa Aβ-42-like peptide profiles were actually more similar to human Aβ-42 profiles than those of mouse Aβ-42. This is consistent with existing observations *in vitro* and *in vivo*, that mouse Aβ does not have as high an aggregation propensity as human Aβ, nor do mouse brains form Amyloid plaques [29, 30]. Interestingly, the altered aggregation propensity of mouse Aβ has previously been attributed to the three amino residues that differ in the mouse peptide (5, 10, and 13) (Figure 5.1) and the involvement of these residues in binding of metal ions [31]. When comparing these residues to those of the equivalent residues in the zebrafish Appa and Appb Aβ-42-like peptides, zebrafish Appa Aβ-42-like peptide shows no conservation with the equivalent human or mouse residues, while zebrafish Appb Aβ-

42-like peptide is conserved with human Aβ-42 at residues 5 and 10, but not 13 (it is instead conserved with mouse Aβ-42 at this position) (Figure 5.1). What does this mean regarding the aggregation propensities of these two putative zebrafish peptides? Further investigation of both zebrafish Aβ-42-like peptides is likely needed to address this observation. It is also worth noting that other studies have reported that residues 6 and 13 are most important for binding metal ions [32]. If these alternative studies are correct, and mouse Aβ-42 shares the same residue at position 6 as human Aβ-42, does this suggest that the residue at position 13 alone is critical for aggregation propensity? Further research is needed to address this question, although, if this is in fact the case then all three of mouse Aβ-42, zebrafish Appa Aβ-42-like and zebrafish Appb Aβ-42-like are not conserved at this residue, which may suggest that none of these peptides have the same aggregation propensity as the human Aβ-42 peptide.

Although human Aβ-40 also forms fibrils, these aggregates are considerably less toxic in AD. It follows that the toxicity of both zebrafish Aβ-42-like peptides should also be analysed. Regarding the observation by Moore *et al* (2014) of a lack of conservation in the putative β-secretase cleavage sites of zebrafish Appa and Appb and what this might mean for production of Aβ from these proteins, a study in our laboratory is currently investigating this using tagged forms of Appa and Appb. To confirm the observations made in this study, the Aβ-42 equivalent peptides that we have predicted for zebrafish Appa and Appb should be produced and used in *in vitro* aggregations studies. Future work should also include further investigation of these peptides *in vivo*. In order to confirm the existence of endogenous Aβ in the zebrafish brain, immunoprecipitation of Aβ peptide from the zebrafish brain should be attempted using existing Aβ antibodies. Zebrafish whole brain lysates could be incubated with the well documented 4G8 antibody. Aβ peptides could then be

sequestered from the lysates using protein A sepharose (or protein G sepharose) beads. The purified Aβ peptides would then be examined by western immunoblot. If these peptides could be isolated successfully from the zebrafish brain, we could investigate further the effects of our fAD-like mutation model zebrafish on Aβ aggregation.

5.4 Methods

Protein sequence alignments

Alignments were conducted using the Geneious software suite, version 5.6.7 (http://www.geneious.com, [29]) Alignments were performed with the following constraints; Cost matrix: Identity, Gap open penalty: 10, Gap extension penalty: 3, Alignment: Global.

Predicting β-aggregation propensity with TANGO

The β-aggregation propensies of human and mouse Aβ-42, along with our predicted Aβ-42 peptides for zebrafish Appa and Appb (Figure 5.1), were assessed with TANGO [20] using pH, temperature, ionic strength and concentration conditions presented in Table 5.1 and default parameters for Nterm and Cterm.

Predicting amyloidogenic regions with other software

The amino acid sequences of human and mouse Aβ-42, along with our predicted Aβ-42 peptides for zebrafish Appa and Appb were entered into the input windows of Zyggregator [1] and Zipper DB [25] to predict aggregation-prone regions of these peptides. When using Zyggregator, the pH was set to 7 as the input only allows for whole numbers and 7 is the closest to the pH of human brain (7.2) and zebrafish cytosol (7-7.44).

5.5 References

1. Tartaglia, G.G. and M. Vendruscolo, *The Zyggregator method for predicting protein aggregation propensities.* Chem Soc Rev, 2008. **37**(7): p. 1395-401.
2. Ulrike, C.M., D. Thomas, and K. Martin, *Not just amyloid: physiological functions of the amyloid precursor protein family.* Nature Reviews Neuroscience, 2017. **18**(5).
3. Hardy, J.A. and G.A. Higgins, *Alzheimer's disease: the amyloid cascade hypothesis.* Science, 1992. **256**(5054): p. 184-5.
4. Fernandez, M.A., et al., *Transmembrane Substrate Determinants for gamma-Secretase Processing of APP CTFbeta.* Biochemistry, 2016. **55**(40): p. 5675-5688.
5. Wolfe, M.S., *The gamma-secretase complex: membrane-embedded proteolytic ensemble.* Biochemistry, 2006. **45**(26): p. 7931-9.
6. Schmidt, M., et al., *Comparison of Alzheimer Abeta(1-40) and Abeta(1-42) amyloid fibrils reveals similar protofilament structures.* Proc Natl Acad Sci U S A, 2009. **106**(47): p. 19813-8.
7. Fändrich, M., *On the structural definition of amyloid fibrils and other polypeptide aggregates.* Cellular and Molecular Life Sciences, 2007. **64**(16): p. 2066-2078.
8. Newman, M., E. Ebrahimie, and M. Lardelli, *Using the zebrafish model for Alzheimer's disease research.* Front Genet, 2014. **5**: p. 189.
9. Saleem, S. and R.R. Kannan, *Zebrafish: an emerging real-time model system to study Alzheimer's disease and neurospecific drug discovery.* Cell Death Discovery, 2018. **4**(1): p. 45.
10. Donnini, S., et al., *Aβ peptides accelerate the senescence of endothelial cells in vitro and in vivo, impairing angiogenesis.* Aβ peptides accelerate the senescence of endothelial cells in vitro and in vivo, impairing angiogenesis, 2010. **24**(7): p. 2385-2395.
11. Luna, S., D.J. Cameron, and D.W. Ethell, *Amyloid-beta and APP deficiencies cause severe cerebrovascular defects: important work for an old villain.* PLoS One, 2013. **8**(9): p. e75052.
12. Newman, M., et al., *A zebrafish melanophore model of amyloid[beta] toxicity.(Report).* Zebrafish, 2010. **7**(2): p. 155.
13. Bhattarai, P., et al., *IL4/STAT6 Signaling Activates Neural Stem Cell Proliferation and Neurogenesis upon Amyloid-beta42 Aggregation in Adult Zebrafish Brain.* Cell Rep, 2016. **17**(4): p. 941-948.
14. Sivaji, K., et al., *Exogenous human beta amyloid peptide interferes osteogenesis through Sox9a in embryonic zebrafish.* Mol Biol Rep, 2019.
15. Leimer, U., et al., *Zebrafish (Danio rerio) presenilin promotes aberrant amyloid beta-peptide production and requires a critical aspartate residue for its function in amyloidogenesis.* Biochemistry, 1999. **38**(41): p. 13602-9.
16. Xi, Y., S. Noble, and M. Ekker, *Modeling Neurodegeneration in Zebrafish.* Current Neurology and Neuroscience Reports, 2011. **11**(3): p. 274-282.
17. Moore, D.B., et al., *Asynchronous evolutionary origins of Abeta and BACE1.* Mol Biol Evol, 2014. **31**(3): p. 696-702.
18. Mullan, M., et al., *A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N–terminus of β–amyloid.* Nature Genetics, 1992. **1**(5): p. 345.
19. Li, Q. and T.C. Sudhof, *Cleavage of amyloid-beta precursor protein and amyloid-beta precursor-like protein by BACE 1.* J Biol Chem, 2004. **279**(11): p. 10542-50.
20. Fernandez-Escamilla, A.M., et al., *Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins.* Nat Biotechnol, 2004. **22**(10): p. 1302-6.
21. Orlowski, P., et al., *Modelling of pH dynamics in brain cells after stroke.* Interface Focus, 2011. **1**(3): p. 408-16.
22. Sadovoy, A. and C. Teh, *11 - Encapsulated biosensors for advanced tissue diagnostics*, in *Biophotonics for Medical Applications*, I. Meglinski, Editor. 2015, Woodhead Publishing. p. 321-330.

23.   Wang, J., et al., *The levels of soluble versus insoluble brain Abeta distinguish Alzheimer's disease from normal and pathologic aging.* Exp Neurol, 1999. **158**(2): p. 328-37.

24.   Storey, K.B., *Functional metabolism regulation and adaptation.* 2004, Hoboken, N.J.: John Wiley & Sons.

25.   Goldschmidt, L., et al., *Identifying the amylome, proteins capable of forming amyloid-like fibrils.* Proc Natl Acad Sci U S A, 2010. **107**(8): p. 3487-92.

26.   Qiu, T., et al., *Aβ42 and Aβ40: similarities and differences.* Journal of Peptide Science, 2015. **21**(7): p. 522-529.

27.   Farrell, R.E., *Chapter 8 - Stringency: Conditions that Influence Nucleic Acid Structure*, in *RNA Methodologies (Fourth Edition)*, R.E. Farrell, Editor. 2010, Academic Press: San Diego. p. 173-178.

28.   Singh, G., et al., *Peptide aggregation in finite systems.* Biophysical journal, 2008. **95**(7): p. 3208-3221.

29.   Dyrks, T., et al., *Amyloidogenicity of rodent and human beta A4 sequences.* FEBS Lett, 1993. **324**(2): p. 231-6.

30.   Lv, X., et al., *Exploring the differences between mouse mAbeta(1-42) and human hAbeta(1-42) for Alzheimer's disease related properties and neuronal cytotoxicity.* Chem Commun (Camb), 2013. **49**(52): p. 5865-7.

31.   Xu, G., et al., *Murine Abeta over-production produces diffuse and compact Alzheimer-type amyloid deposits.* Acta Neuropathol Commun, 2015. **3**: p. 72.

32.   Lermyte, F., et al., *Metal Ion Binding to the Amyloid β Monomer Studied by Native Top-Down FTICR Mass Spectrometry.* Journal of The American Society for Mass Spectrometry, 2019. **30**(10): p. 2123-2134.

33.   Kearse, M., et al., *Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.* Bioinformatics, 2012. **28**(12): p. 1647-9.

5.6 Supplementary information

This section is included in the thesis as information supplementary to Chapter 5. It contains additional information not included in the main text of the manuscript.

File S1. Supplementary tables for TANGO analyses

Table S1. Combinations of physio-chemical conditions input into TANGO software

- Highlighted cells indicate experiments not included in the final analysis.

Experiment	pH	Temperature (K)	Ionic Strength	Concentration
1	7.2	310.15	0.3	9×10^{-12}
2	7	298.15	0.3	9×10^{-12}
3	7.2	298.15	0.3	9×10^{-12}
4	7	310.15	0.3	9×10^{-12}
5	7.2	310.15	0.2	9×10^{-12}
6	7	298.15	0.2	9×10^{-12}
7	7.2	298.15	0.2	9×10^{-12}
8	7	310.15	0.2	9×10^{-12}
9	7.44	298.15	0.2	9×10^{-12}
10	7.44	298.15	0.3	9×10^{-12}
11	7.44	310.15	0.2	9×10^{-12}
12	7.44	310.15	0.3	9×10^{-12}

Table S2. Experiment 1 - β-aggregation (β-agg) values at each amino acid residue position of

the Aβ-42 peptides as generated by the TANGO software

Position	Human		Mouse		Zebrafish Appa		Zebrafish Appb	
	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg
1	D	0	D	0	E	0	D	0
2	A	0	A	0	A	0	I	0
3	E	0	E	0	E	0	E	0
4	F	0	F	0	E	0	E	0
5	R	0	G	0	R	0	R	0
6	H	0	H	0	H	0	H	0
7	D	0	D	0	S	0	N	0
8	S	0	S	0	E	0	A	0
9	G	0	G	0	V	0	G	0
10	Y	0	F	0	Y	0	Y	0
11	E	0	E	0	H	0	D	0
12	V	0	V	0	Q	0	V	0
13	H	0	R	0	K	0	R	0
14	H	0	H	0	L	77.261	D	0
15	Q	0	Q	0	V	77.261	K	0
16	K	0	K	0	F	77.261	R	0
17	L	77.204	L	77.157	F	77.261	L	31.515
18	V	77.204	V	77.157	A	77.261	M	31.515
19	F	77.204	F	77.157	E	0	F	31.515
20	F	77.204	F	77.157	D	0	L	31.515
21	A	77.204	A	77.157	V	0	A	31.515
22	E	0	E	0	S	0	E	0
23	D	0	D	0	S	0	D	0
24	V	0	V	0	N	0	M	0
25	G	0	G	0	K	0	G	0
26	S	0	S	0	G	5.481	S	0
27	N	0	N	0	A	66.588	N	0
28	K	0	K	0	I	89.595	K	0
29	G	5.482	G	5.481	I	91.842	G	5.500
30	A	66.644	A	66.630	G	92.047	A	66.782
31	I	89.623	I	89.607	L	94.336	I	89.776
32	I	91.867	I	91.852	M	94.731	I	92.020
33	G	92.072	G	92.057	V	94.988	G	92.225
34	L	94.360	L	94.345	G	92.814	L	94.509
35	M	94.755	M	94.741	G	92.620	M	94.903
36	V	95.012	V	94.998	V	92.609	V	95.159
37	G	92.836	G	92.821	V	92.397	G	92.990
38	G	92.642	G	92.627	I	88.996	G	92.796

39	V	92.631	V	92.615	A	0	V	92.785
40	V	92.420	V	92.404			V	92.574
41	I	89.016	I	89.001			I	89.168
42	A	0	A	0			A	0

Table S3. Experiment 2 - β-aggregation (β-agg) values at each amino acid residue position of the Aβ-42 peptides as generated by the TANGO software

	Human		Mouse		Zebrafish Appa		Zebrafish Appb	
Position	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg
1	D	0	D	0	E	0	D	0
2	A	0	A	0	A	0	I	0
3	E	0	E	0	E	0	E	0
4	F	0	F	0	E	0	E	0
5	R	0	G	0	R	0	R	0
6	H	0	H	0	H	0	H	0
7	D	0	D	0	S	0	N	0
8	S	0	S	0	E	0	A	0
9	G	0	G	0	V	0	G	0
10	Y	0	F	0	Y	0	Y	0
11	E	0	E	0	H	0	D	0
12	V	0	V	0	Q	0	V	0
13	H	0	R	0	K	0.042	R	0
14	H	0	H	0	L	86.290	D	0
15	Q	0	Q	0	V	86.290	K	0
16	K	0.042	K	0.042	F	86.290	R	0
17	L	86.238	L	86.191	F	86.290	L	47.333
18	V	86.238	V	86.191	A	86.290	M	47.333
19	F	86.238	F	86.191	E	0.042	F	47.333
20	F	86.238	F	86.191	D	0	L	47.333
21	A	86.238	A	86.191	V	0	A	47.333
22	E	0.042	E	0.042	S	0	E	0
23	D	0	D	0	S	0	D	0
24	V	0	V	0	N	0	M	0
25	G	0	G	0	K	0	G	0
26	S	0	S	0	G	5.728	S	0
27	N	0	N	0	A	69.168	N	0
28	K	0	K	0	I	93.371	K	0
29	G	5.729	G	5.727	I	95.463	G	5.740
30	A	69.212	A	69.202	G	95.635	A	69.278
31	I	93.388	I	93.378	L	97.541	I	93.449
32	I	95.478	I	95.469	M	97.832	I	95.537

33	G	95.649	G	95.640	V	97.957	G	95.708
34	L	97.555	L	97.547	G	96.138	L	97.609
35	M	97.846	M	97.837	G	95.971	M	97.899
36	V	97.970	V	97.962	V	95.958	V	98.023
37	G	96.150	G	96.141	V	95.766	G	96.209
38	G	95.984	G	95.974	I	92.544	G	96.043
39	V	95.970	V	95.961	A	0.010	V	96.030
40	V	95.779	V	95.769			V	95.838
41	I	92.555	I	92.545			I	92.616
42	A	0.010	A	0.010			A	0.010

Table S4. Experiment 5 - β-aggregation (β-agg) values at each amino acid residue position of the Aβ-42 peptides as generated by the TANGO software

	Human		Mouse		Zebrafish Appa		Zebrafish Appb	
Position	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg
1	D	0	D	0	E	0	D	0
2	A	0	A	0	A	0	I	0
3	E	0	E	0	E	0	E	0
4	F	0	F	0	E	0	E	0
5	R	0	G	0	R	0	R	0
6	H	0	H	0	H	0	H	0
7	D	0	D	0	S	0	N	0
8	S	0	S	0	E	0	A	0
9	G	0	G	0	V	0	G	0
10	Y	0	F	0	Y	0	Y	0
11	E	0	E	0	H	0	D	0
12	V	0	V	0	Q	0	V	0
13	H	0	R	0	K	0.043	R	0
14	H	0	H	0	L	77.436	D	0
15	Q	0	Q	0	V	77.436	K	0
16	K	0.043	K	0.043	F	77.436	R	0
17	L	77.376	L	77.321	F	77.436	L	36.469
18	V	77.376	V	77.321	A	77.436	M	36.469
19	F	77.376	F	77.321	E	0.043	F	36.469
20	F	77.376	F	77.321	D	0	L	36.469
21	A	77.376	A	77.321	V	0	A	36.469
22	E	0.043	E	0.043	S	0	E	0
23	D	0	D	0	S	0	D	0
24	V	0	V	0	N	0	M	0
25	G	0	G	0	K	0	G	0
26	S	0	S	0	G	5.81	S	0

27	N	0	N	0	A	70.513	N	0
28	K	0	K	0	I	90.602	K	0
29	G	5.811	G	5.809	I	92.562	G	5.831
30	A	70.57	A	70.554	G	92.733	A	70.724
31	I	90.629	I	90.612	L	94.714	I	90.793
32	I	92.586	I	92.569	M	95.055	I	92.749
33	G	92.757	G	92.74	V	95.266	G	92.926
34	L	94.737	L	94.72	G	93.244	L	94.901
35	M	95.079	M	95.062	G	93.064	M	95.242
36	V	95.289	V	95.273	V	93.053	V	95.452
37	G	93.266	G	93.249	V	92.856	G	93.436
38	G	93.085	G	93.068	I	89.445	G	93.256
39	V	93.074	V	93.057	A	0	V	93.245
40	V	92.877	V	92.86			V	93.048
41	I	89.464	I	89.447			I	89.627
42	A	0	A	0			A	0

Table S5. Experiment 6 - β-aggregation (β-agg) values at each amino acid residue position of the Aβ-42 peptides as generated by the TANGO software

	Human		Mouse		Zebrafish Appa		Zebrafish Appb	
Position	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg
1	D	0	D	0	E	0	D	0
2	A	0	A	0	A	0	I	0
3	E	0	E	0	E	0	E	0
4	F	0	F	0	E	0	E	0
5	R	0	G	0	R	0	R	0
6	H	0	H	0	H	0	H	0
7	D	0	D	0	S	0	N	0
8	S	0	S	0	E	0	A	0
9	G	0	G	0	V	0	G	0
10	Y	0	F	0	Y	0	Y	0
11	E	0	E	0	H	0	D	0
12	V	0	V	0	Q	0	V	0
13	H	0	R	0	K	0.056	R	0
14	H	0	H	0	L	86.411	D	0
15	Q	0	Q	0	V	86.411	K	0
16	K	0.056	K	0.056	F	86.411	R	0
17	L	86.357	L	86.302	F	86.411	L	52.898
18	V	86.357	V	86.302	A	86.411	M	52.898
19	F	86.357	F	86.302	E	0.056	F	52.898
20	F	86.357	F	86.302	D	0	L	52.898

21	A	86.357	A	86.302	V	0	A	52.898
22	E	0.056	E	0.056	S	0	E	0
23	D	0	D	0	S	0	D	0
24	V	0	V	0	N	0	M	0
25	G	0	G	0	K	0	G	0
26	S	0	S	0	G	6.052	S	0
27	N	0	N	0	A	73.051	N	0
28	K	0	K	0	I	94.114	K	0
29	G	6.053	G	6.051	I	95.934	G	6.066
30	A	73.097	A	73.085	G	96.082	A	73.172
31	I	94.131	I	94.12	L	97.729	I	94.198
32	I	95.948	I	95.938	M	97.98	I	96.012
33	G	96.096	G	96.086	V	98.079	G	96.16
34	L	97.743	L	97.733	G	96.393	L	97.802
35	M	97.993	M	97.984	G	96.239	M	98.052
36	V	98.092	V	98.083	V	96.226	V	98.15
37	G	96.405	G	96.395	V	96.048	G	96.47
38	G	96.251	G	96.241	I	92.818	G	96.317
39	V	96.238	V	96.228	A	0.002	V	96.304
40	V	96.061	V	96.05			V	96.127
41	I	92.829	I	92.818			I	92.896
42	A	0.002	A	0.002			A	0.002

Table S6. Experiment 9 - β-aggregation (β-agg) values at each amino acid residue position of the Aβ-42 peptides as generated by the TANGO software

	Human		Mouse		Zebrafish Appa		Zebrafish Appb	
Position	**Amino acid**	**β-agg**	**Amino acid**	**β-agg**	**Amino acid**	**β-agg**	**Amino acid**	**β-agg**
1	D	0	D	0	E	0	D	0
2	A	0	A	0	A	0	I	0
3	E	0	E	0	E	0	E	0
4	F	0	F	0	E	0	E	0
5	R	0	G	0	R	0	R	0
6	H	0	H	0	H	0	H	0
7	D	0	D	0	S	0	N	0
8	S	0	S	0	E	0	A	0
9	G	0	G	0	V	0	G	0
10	Y	0	F	0	Y	0	Y	0
11	E	0	E	0	H	0	D	0
12	V	0	V	0	Q	0	V	0
13	H	0	R	0	K	0.047	R	0
14	H	0	H	0	L	86.271	D	0

	Human		Mouse		Zebrafish Appa		Zebrafish Appb	
15	Q	0	Q	0	V	86.271	K	0
16	K	0.047	K	0.047	F	86.271	R	0
17	L	86.220	L	86.195	F	86.271	L	52.840
18	V	86.220	V	86.195	A	86.271	M	52.840
19	F	86.220	F	86.195	E	0.047	F	52.840
20	F	86.220	F	86.195	D	0	L	52.840
21	A	86.220	A	86.195	V	0	A	52.840
22	E	0.047	E	0.047	S	0	E	0
23	D	0	D	0	S	0	D	0
24	V	0	V	0	N	0	M	0
25	G	0	G	0	K	0	G	0
26	S	0	S	0	G	6.048	S	0
27	N	0	N	0	A	73.027	N	0
28	K	0	K	0	I	94.092	K	0
29	G	6.049	G	6.048	I	95.913	G	6.064
30	A	73.071	A	73.066	G	96.061	A	73.160
31	I	94.106	I	94.102	L	97.71	I	94.186
32	I	95.925	I	95.920	M	97.961	I	96.001
33	G	96.073	G	96.068	V	98.06	G	96.149
34	L	97.721	L	97.717	G	96.371	L	97.792
35	M	97.972	M	97.968	G	96.217	M	98.041
36	V	98.071	V	98.067	V	96.204	V	98.140
37	G	96.381	G	96.376	V	96.026	G	96.458
38	G	96.227	G	96.222	I	92.795	G	96.305
39	V	96.214	V	96.209	A	0.002	V	96.292
40	V	96.036	V	96.031			V	96.115
41	I	92.803	I	92.799			I	92.883
42	A	0.002	A	0.002			A	0.002

Table S7. Experiment 10 - β-aggregation (β-agg) values at each amino acid residue position of the Aβ-42 peptides as generated by the TANGO software

	Human		Mouse		Zebrafish Appa		Zebrafish Appb	
Position	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg	Amino acid	β-agg
1	D	0	D	0	E	0	D	0
2	A	0	A	0	A	0	I	0
3	E	0	E	0	E	0	E	0
4	F	0	F	0	E	0	E	0
5	R	0	G	0	R	0	R	0
6	H	0	H	0	H	0	H	0
7	D	0	D	0	S	0	N	0
8	S	0	S	0	E	0	A	0

9	G	0	G	0	V	0	G	0
10	Y	0	F	0	Y	0	Y	0
11	E	0	E	0	H	0	D	0
12	V	0	V	0	Q	0	V	0
13	H	0	R	0	K	0.036	R	0
14	H	0	H	0	L	86.167	D	0
15	Q	0	Q	0	V	86.167	K	0
16	K	0.036	K	0.036	F	86.167	R	0
17	L	86.135	L	86.099	F	86.167	L	47.283
18	V	86.135	V	86.099	A	86.167	M	47.283
19	F	86.135	F	86.099	E	0.036	F	47.283
20	F	86.135	F	86.099	D	0	L	47.283
21	A	86.135	A	86.099	V	0	A	47.283
22	E	0.036	E	0.036	S	0	E	0
23	D	0	D	0	S	0	D	0
24	V	0	V	0	N	0	M	0
25	G	0	G	0	K	0	G	0
26	S	0	S	0	G	5.724	S	0
27	N	0	N	0	A	69.146	N	0
28	K	0	K	0	I	93.351	K	0
29	G	5.725	G	5.724	I	95.444	G	5.738
30	A	69.189	A	69.184	G	95.616	A	69.268
31	I	93.366	I	93.361	L	97.524	I	93.438
32	I	95.457	I	95.453	M	97.815	I	95.526
33	G	95.628	G	95.624	V	97.94	G	95.697
34	L	97.536	L	97.532	G	96.118	L	97.6
35	M	97.826	M	97.823	G	95.951	M	97.889
36	V	97.951	V	97.948	V	95.938	V	98.014
37	G	96.128	G	96.124	V	95.746	G	96.198
38	G	95.962	G	95.958	I	92.524	G	96.032
39	V	95.948	V	95.944	A	0.01	V	96.019
40	V	95.756	V	95.752			V	95.827
41	I	92.533	I	92.528			I	92.605
42	A	0.01	A	0.01			A	0.01

Table S8. Agg scores for all TANGO experiments

Experiment	Human	Mouse	Zf Appa	Zf Appb
1	1475.38	1474.96	1475.35	1248.76
2	1560.55	1560.20	1560.62	1366.66
5	1486.19	1485.71	1486.18	1283.58
6	1568.84	1568.44	1568.91	1402.22
9	1567.84	1567.67	1567.93	1401.79
10	1559.76	1559.53	1559.75	1366.28

File S2. Supplementary tables for Zyggregator analyses

Table S9. Raw Zygg residue score data output from Zyggregator

	Human Aβ-42		Mouse Aβ-42 equivalent		Zebrafish Appb Aβ-42 equivalent		Zebrafish Appa Aβ-42 equivalent	
Position	Amino Acid	Zygg residue score	Amino Acid	Zygg residue score	Amino Acid	Zygg residue score	Amino Acid	Zygg residue score
1	D	0.0619504	D	0.06195	D	-0.588424	E	-1.29325
2	A	-0.437954	A	0.223669	I	-1.11194	A	-1.84235
3	E	-0.303041	E	0.205307	E	-0.820887	E	-1.38209
4	F	-0.494487	F	0.094332	E	-0.964486	E	-1.64632
5	R	-0.335666	G	0.253154	R	-0.880837	R	-1.64632
6	H	-0.16945	H	0.41937	H	-1.04779	H	-1.2646
7	D	0.693793	D	1.38153	N	-0.184546	S	-0.401354
8	S	-0.268368	S	0.41937	A	0.163942	E	0.0878349
9	G	0.564591	G	0.634872	G	0.996902	V	0.48241
10	Y	0.564591	F	0.126579	Y	0.488608	Y	0.0939694
11	E	0.705292	E	0.267279	D	0.329613	H	0.173497
12	V	0.283972	V	-0.1254	V	-0.10356	Q	0.510175
13	H	-0.401329	R	-0.8107	R	-0.908714	K	0.658481
14	H	-0.490038	H	-0.99833	D	-0.997422	L	0.945095
15	Q	0.323818	Q	-0.18448	K	-0.508714	V	1.26344
16	K	0.284427	K	-0.03617	R	-0.548104	F	0.700269
17	L	0.757399	L	0.945095	L	0.0578379	F	0.948009
18	V	1.07574	V	1.26344	M	0.329183	A	1.17503

| 19	F	0.700269	F	0.887965	F	0.228435	E	1.05618
20	F	0.948009	F	1.1357	L	0.596027	D	0.601327
21	A	1.17503	A	1.36272	A	0.65869	V	0.146649
22	E	1.14722	E	0.959523	E	0.419846	S	-0.18474
23	D	0.692369	D	0.504673	D	-0.0350047	S	0.413613
24	V	0.237692	V	0.237692	M	0.0733369	N	0.309566
25	G	-0.0936972	G	-0.0937	G	-0.445748	K	0.261015
26	S	0.504656	S	0.504656	S	0.340301	G	0.51901
27	N	0.400608	N	0.400608	N	0.236253	A	0.610052
28	K	0.352058	K	0.352058	K	0.352058	I	0.689407
29	G	0.51901	G	0.51901	G	0.51901	I	1.23816
30	A	0.610052	A	0.610052	A	0.610052	G	1.64136
31	I	0.689407	I	0.689407	I	0.689407	L	1.80757
32	I	1.23816	I	1.23816	I	1.23816	M	1.64062
33	G	1.64136	G	1.64136	G	1.64136	V	1.68917
34	L	1.80757	L	1.80757	L	1.80757	G	1.90467
35	M	1.64062	M	1.64062	M	1.64062	G	2.08314
36	V	1.68917	V	1.68917	V	1.68917	V	1.86578
37	G	1.90467	G	1.90467	G	1.90467	V	1.18629
38	G	2.08314	G	2.08314	G	2.08314	I	1.38345
39	V	1.86578	V	1.86578	V	1.86578	A	1.18111
40	V	1.18629	V	1.18629	V	1.18629		
41	I	1.38345	I	1.38345	I	1.38345		
42	A	1.18111	A	1.18111	A	1.18111		

File S3. Supplementary tables for Zipper DB analyses

Table S10. Rosetta energies (Re) of all possible hexapeptides within the Aβ-42 peptide regions of human APP, mouse APP and zebrafish Appa and Appb proteins.

- Yellow highlight indicates positions that are likely to aggregate (based on Re < -23 kcal/mol)

	H.sapiens APP		*M.musculus* APP		*D.rerio* Appa		*D.rerio* Appb	
Position	Sequence	Re	Sequence	Re	Sequence	Re	Sequence	Re
1	DAEFRH	-12.5	DAEFGH	-14.7	EAEERH	-14.4	DIEERH	-16.3
2	AEFRHD	-17.5	AEFGHD	-17.9	AEERHS	-17.8	IEERHN	-15.3
3	EFRHDS	-14.6	EFGHDS	-17.4	EERHSE	-16.8	EERHNA	-19.8
4	FRHDSG	-15.5	FGHDSG	-17.9	ERHSEV	-17.7	ERHNAG	-17.2
5	RHDSGY	-17	GHDSGF	-19.2	RHSEVY	-19.8	RHNAGY	-18
6	HDSGYE	-19.9	HDSGFE	-18.8	HSEVYH	-22	HNAGYD	-18.8
7	DSGYEV	-17.2	DSGFEV	-14.4	SEVYHQ	-21.9	NAGYDV	-18.4
8	SGYEVH	-20.5	SGFEVR	-17.5	EVYHQK	-22.7	AGYDVR	-15.5
9	GYEVHH	-22.2	GFEVRH	-17.9	VYHQKL	-21.3	GYDVRD	-15.5
10	YEVHHQ	-21.1	FEVRHQ	-19.4	YHQKLV	-21.3	YDVRDK	-14
11	EVHHQK	-21.6	EVRHQK	-18.7	HQKLVF	-22.6	DVRDKR	-12.1
12	VHHQKL	-19.9	VRHQKL	-16.5	QKLVFF	-24	VRDKRL	-12.9
13	HHQKLV	-21.1	RHQKLV	-17.6	KLVFFA	-22.2	RDKRLM	-10.5
14	HQKLVF	-22.6	HQKLVF	-22.6	LVFFAE	-7.5	DKRLMF	-18.7
15	QKLVFF	-24	QKLVFF	-24	VFFAED	-18.7	KRLMFL	-18.4
16	KLVFFA	-22.2	KLVFFA	-22.2	FFAEDV	-19.3	RLMFLA	-12.8
17	LVFFAE	-7.5	LVFFAE	-7.5	FAEDVS	-18.2	LMFLAE	-21.4

18	VFFAED	-18.7	VFFAED	-18.7	AEDVSS	-20.9	MFLAED	-18.1
19	FFAEDV	-19.3	FFAEDV	-19.3	EDVSSN	-18.8	FLAEDM	-19.3
20	FAEDVG	-18	FAEDVG	-18	DVSSNK	-21.2	LAEDMG	-16.9
21	AEDVGS	-20.4	AEDVGS	-20.4	VSSNKG	-20.9	AEDMGS	-18.6
22	EDVGSN	-17.9	EDVGSN	-17.9	SSNKGA	-22.3	EDMGSN	-17.5
23	DVGSNK	-22	DVGSNK	-22	SNKGAI	-22.4	DMGSNK	-20.6
24	VGSNKG	-19.9	VGSNKG	-19.9	NKGAII	-23.5	MGSNKG	-21.3
25	GSNKGA	-22.3	GSNKGA	-22.3	KGAIIG	-23.3	GSNKGA	-22.3
26	SNKGAI	-22.4	SNKGAI	-22.4	GAIIGL	-26	SNKGAI	-22.4
27	NKGAII	-23.5	NKGAII	-23.5	AIIGLM	-25.7	NKGAII	-23.5
28	KGAIIG	-23.3	KGAIIG	-23.3	IIGLMV	-23	KGAIIG	-23.3
29	GAIIGL	-26	GAIIGL	-26	IGLMVG	-21.9	GAIIGL	-26
30	AIIGLM	-25.7	AIIGLM	-25.7	GLMVGG	-23.1	AIIGLM	-25.7
31	IIGLMV	-23	IIGLMV	-23	LMVGGV	-21.7	IIGLMV	-23
32	IGLMVG	-21.9	IGLMVG	-21.9	MVGGVV	-23.4	IGLMVG	-21.9
33	GLMVGG	-23.1	GLMVGG	-23.1	VGGVVI	-25.1	GLMVGG	-23.1
34	LMVGGV	-21.7	LMVGGV	-21.7	GGVVIA	-27.8	LMVGGV	-21.7
35	MVGGVV	-23.4	MVGGVV	-23.4			MVGGVV	-23.4
36	VGGVVI	-25.1	VGGVVI	-25.1			VGGVVI	-25.1
37	GGVVIA	-27.8	GGVVIA	-27.8			GGVVIA	-27.8

File S4. Supplementary figures for Zipper DB analyses
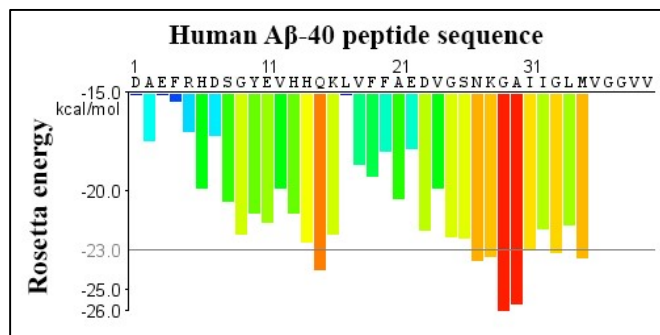
Figure S1. Zipper DB fibrillation propensity profile of human Aβ-40 peptide. *– Red histogram bars correspond to hexapeptides (6 amino acid residues, starting from the residue directly above the bar and read C-terminally, i.e. the hexapeptide for the first red bar would be GAIIGL) with energy < −23 kcal/mol and are predicted to form fibrils.*

Chapter 6 - Further Discussion, Conclusions and Future directions

6.1 Introduction

γ-secretase is an important protease complex as it is responsible for the cleavage of over 100 substrates within their transmembrane domains [1, 2]. Although many substrates of γ-secretase are known, the specificity by which γ-secretase selects its substrates remains a mystery. γ-secretase plays a role in Alzheimer's disease (AD) through its processing of APP to produce the aggregation-prone Aβ peptide [3]. This sequential processing of APP that produces the Aβ peptide is considered a major pathological event in AD disease progression, as described by the amyloid cascade hypothesis [4]. Other than its role in AD disease, γ-secretase also has numerous roles in development and adulthood through regulating many cellular events, including axon guidance and formation, neurite outgrowth, cell fate determination and maintenance of synapses. Due to its involvement in this wide array of important physiological processes, defects in γ-secretase mediated signalling pathways can lead to numerous pathologies, such as: Neurodegeneration in AD, Acne inversa, Autoimmune disease, Breast cancer and Dilated cardiomyopathy [reviewed in 5]. Thus, it is important to gain a better understanding of substrate selection by γ-secretase, to facilitate more efficient drug design that allows for targeting only the specific substrate involved in each disease.

In this thesis we aimed to investigate γ-secretase and two of its substrates, $p75^{NTR}$ and APP, on three different fronts; 1) selection of substrates by γ-secretase using an assay-based system, 2) the molecular effects of fAD-like and null mutations in zebrafish *appa* and *appb*, by generating mutations in endogenous genes with genome editing technologies, and 3) the predicted aggregation propensity of Aβ-42-like peptides of Appa and Appb using predictive software.

6.2 How does γ-secretase select specific substrates?

The amyloid cascade hypothesis continues to dominate the field of AD research, mostly due to the observation of plaques composed of oligomerised Aβ peptide in the brains of AD sufferers. γ-secretase is a key component of this hypothesis, due to its processing of APP to release the Aβ peptide. The observation that γ-secretase-cleavage of APP results in the release of Aβ peptides, led many researchers to develop inhibitory drugs targeting this protease in the hopes of preventing Aβ aggregation into plaques [6]. Our current knowledge suggests that γ-secretase is responsible for cleaving more than 100 substrates (including APP and Notch), so it should come as no surprise that drugs designed to inhibit this protease cause severe side effects [7]. A better understanding of substrate selection by γ-secretase may assist in the design of drugs to target more specifically its processing of select substrates.

As there is still no clear answer to the question of substrate selection by γ-secretase, we sought to investigate this using the known γ-secretase substrate, p75^{NTR}, and its homolog NRH1. We leveraged the structural and sequence similarities (especially in the transmembrane domain (TMD)) of these two homologues to design an assay system with which we could attempt to elucidate structural determinants that enable p75^{NTR} to be cleaved by γ-secretase while NRH1 is not. Our assay determined that the TMD of p75^{NTR} alone was insufficient to confer γ-secretase cleavage to its homolog Nrh1. Therefore, we posited that there must be structures other than the TMD that are also required for cleavage by γ-secretase. Interestingly, another study employing a similar strategy to ours, found that γ-secretase cleavage of their substrate required both a permissive transmembrane and a permissive intracellular domain (ICD) [8]. Suggesting that the ICD is just as important as the TMD to confer γ-secretase cleavage. Furthermore, a study by Sykes

et al (2012) suggested that dimerisation of substrates or the structure of substrate α-helices may regulate γ-secretase cleavage [9]. While zebrafish Nrh1 has a different dimerisation domain to that of p75^{NTR}, replacing the Nrh1 dimerisation domain by the swapping its TMD with the p75^{NTR} TMD in our study, did not confer γ-secretase cleavage. This suggests that the p75^{NTR} dimerisation domain alone is also not sufficient to confer γ-secretase cleavage.

In Chapter 2 we discussed some options for investigating the question of which enzyme is responsible for cleavage of the TMD of Nrh1. We suggested that a range of protease inhibitors could be employed to identify the class of iCLIP responsible. Investigation into what causes the previously observed cleavage within the Nrh1 TMD could be further extended as follows:

(1) Site directed mutagenesis could be used to alter Leucine (Figure 6.1) to Proline in the TMD of our existing Nrh1 assay construct, as this residue supposedly assists the TMD of p75 in changing confirmation from an α-helix to a β-sheet, thereby allowing γ-secretase access for proteolysis. The altered construct could then be injected into single cell embryos and DAPT inhibitor applied to assess whether cleavage by γ-secretase is conferred, thereby confirming whether this residue contributes to proteolysis through allowing relaxation of the α-helix into β-sheet for better access for γ-secretase.

(2) Antisense morpholino oligonucleotides (MOs) could be used to block *psen1* and *psen2* expression to confirm that Nrh1 cleavage is presenilin-independent. MOs are a type of oligomer molecule that can be injected into zebrafish embryos to block expression of proteins through binding to mRNAs [10]. Previous work in our laboratory has involved inhibition of

Psen1 and Psen2 translation with MOs that bind over the start codons of their mRNA transcripts. These existing MOs could be co-injected with the previously described Nrh1 assay construct at the 1-cell-stage and western immunoblot analysis could then be performed on embryos at 24hpf to determine the effect of presenilin inhibition.

(3) Subcellular localisation of both NRH1 and p75^{NTR} cleavage events could also be investigated. Combined evidence suggests the principal sites of γ-secretase cleavage to be endosomes, TGN and at the plasma membrane [reviewed in 3, 11], so it is reasonable to assume that p75^{NTR} cleavage may occur in one or all of these compartments. However, the subcellular localisation of NRH1 and whichever enzyme is responsible for its proteolysis is more of a mystery. With a current lack of open source software that can accurately predict the subcellular localisation of proteins in the zebrafish, more traditional methods could be employed to investigate the location of p75^{NTR} and Nrh1. Fluorescence microscopy might be employed by synthesising GFP labelled recombinant p75^{NTR} and Nrh1 zebrafish proteins, injecting them at the 1-cell-stage, and monitoring their localisation at different time points. Alternatively, subcellular fractionation followed by western immunoblotting with p75^{NTR}- and Nrh1-specific antibodies could also be attempted. Knowledge of the respective subcellular cleavage locations of Nrh1 and p75^{NTR} may also assist in identifying the enzyme responsible for the non-γ-secretase cleavage event of Nrh1.

(4) Another extension of this project could be a deeper investigation of the evolution of γ-secretase dependence or independence of proteins. Kanning et al (2003) proposed an evolutionary model for the existence of the p75^{NTR} and NRH1 homologues [12]. Their hypothesis was that a gene

duplication event within a common ancestral sequence led to formation of the p75^{NTR} and NRH1 sequences. NRH2 lacks the ECD motif observed in both p75^{NTR} and NRH1, hence it may have arisen by a deletion mutation around the time that the avian and mammalian lineages diverged [12, 13]. Considering this, one can assume the existence of a common ancestral sequence encoding a protein either cleaved or not by γ-secretase and from which a γ-secretase dependence or independence evolved (Figure 6.2). Therefore, identification of a common ancestral sequence that gave rise to p75^{NTR} and NRH1 through additional phylogenetic analyses would allow us to develop an assay specifically to test whether this ancestor is cleaved by γ-secretase. This could give us insight into whether p75^{NTR}'s dependence on γ-secretase is a characteristic that has evolved recently.

Understanding the basis for γ-secretase substrate selection is important for understanding disease processes, such as those involved in Alzheimer's disease, and for rational drug design. A specific recognition and/or cleavage site for γ-secretase-mediated proteolysis has not yet been established. This makes designing drugs to prevent, specifically, proteolysis of APP and release of Amyloid beta particularly challenging and may contribute to the explanation of why so many drugs targeting this mechanism in attempts to alleviate and/or slow the progression of Alzheimer's Disease symptoms have failed. We do not yet understand what makes a γ-secretase substrate unique for selection/targeting for proteolysis by this protease complex. As it stands there are many substrates of this protease with varied roles in cell biology which makes targeted drug design even more important as widespread disruption of these processes would be expected to have deleterious effects. The unique design of our assay system may allow us to gain a deeper understanding to the question of what structures define a γ-secretase substrate. Information gleaned from the

227

experiments proposed in the discussions above could be used to design new drugs that target specifically γ-secretase's ability to cleave specific substrates.
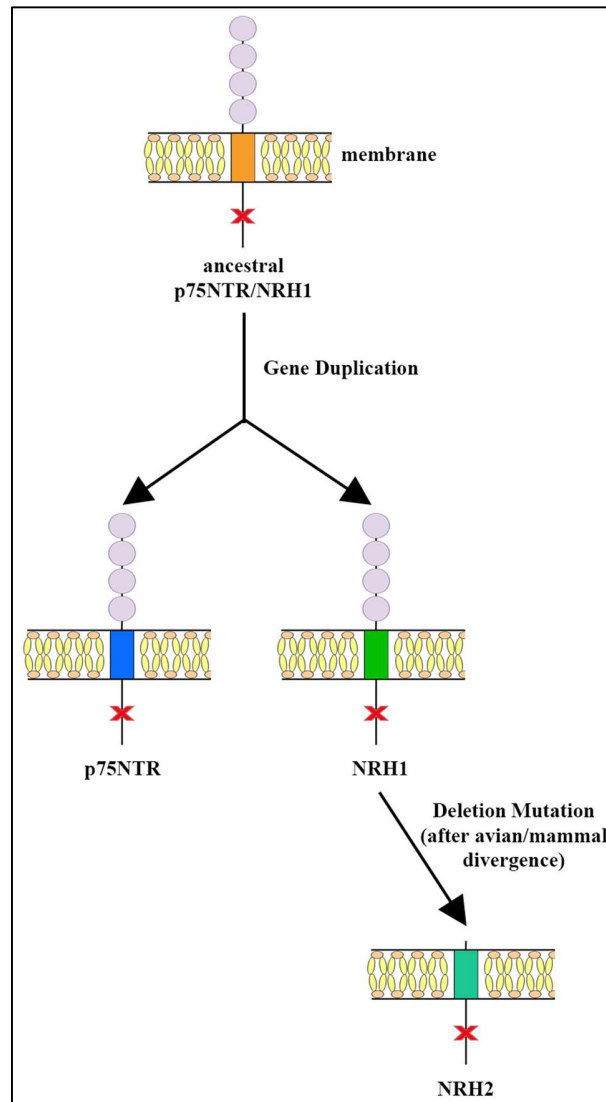
Figure 6.2 Suggested evolutionary development of p75^NTR and NRH genes. *The high degree of conservation between the sequences of p75^NTR and NRH1 suggests that these two sequences arose via a gene duplication event in an ancestral sequence. As mammalian NRH2 lacks an extracellular domain (ECD) motif, it may have arisen by a deletion mutation after the divergence of avian and mammalian lineages*

6.3 What are the molecular effects of fAD-like and null mutations in zebrafish appa and appb?

The overarching goal of the ADGL laboratory is to develop a catalogue of fAD-like mutations and null mutations in each of the fAD gene equivalents in zebrafish to identify the molecular changes-in-common that contribute to AD pathology. One key question to be answered from this larger project is, what are the effects of APP fAD mutations on key molecular processes in the brain that eventually lead to AD? We originally aimed to generate several fAD-like mutations spanning the Aβ peptide region of human APP at equivalent residues in the zebrafish *appa* and *appb* genes, along with null mutations in both genes, to contribute to this larger project.

In Chapter 3, we utilised current genome editing techniques and attempted to generate a suite of fAD-like and null mutations in *appa* and *appb* with limited success. While we were able to achieve double-stranded break generation with an sgRNA targeting a fAD mutation site in *appa* (Austrian mutation site) and one fAD mutation site in *appb* (Arctic mutation site), introduction of mutations of interest using engineering oligonucleotides was not achieved. We also discussed how future work should utilise recent developments in CRISPR technologies to successfully generate these fAD-like mutations in *appa* and *appb*.

In Chapter 4, successful generation of a putative null mutation in zebrafish *appb* using an sgRNA that successfully generated a DSB near the start codon was described (*appb^{R5PfsTer26}*) and preliminary analyses of the hypoxic response to this mutation performed. We observed no phenotype in our heterozygous *appb^{R5PfsTer26}* putative null mutants and proposed that this is possibly due to the phenomenon of "genetic compensation".

Previous studies using MEFs have shown that the AICD induces *Hif1a* gene expression and HIF-1α protein stability [14]. Along with this, hypoxic conditions have been shown to upregulate mRNA and protein expression of APP [15]. This may indicate a possible positive feedback loop, whereby increased production of APP under hypoxic conditions results in increased AICD release, which in turn results in an increased hypoxia response through increased *Hif1a* expression. Therefore, we proposed that loss of *appb* in our putative null mutant should have resulted in a reduction in the hypoxia response.

APP is also connected to the hypoxia response through its involvement in cellular iron homeostasis [16]. Interestingly, both APP and its cleavage product, sAPPα, have been shown to bind FERROPORTIN 1 (FPN1) to improve the process by which neurons export excess iron (Fe^{2+}). HIF is stabilised under hypoxia and degraded under normoxia. During normoxia, Fe^{2+} (together with oxygen and other factors) is needed to hydroxylate HIF1α to target it for degradation. In Chapter 4 we discussed how loss of the AICD from Appb in our putative null mutant would be expected to reduce the hypoxia response by potentially reducing expression of the *Hif1a* gene. As APP is also involved in the movement of Fe^{2+} out of the cytosol, we would expect that the cytosol of our *appb*^{R5PfsTer26} putative null mutants would accumulate excess Fe^{2+}. This excess of available Fe^{2+} may result in increased destabilisation or degradation of HIF1α, thereby decreasing the hypoxia response. Therefore, we would also expect to observe decreased induction of HRG expression in hypoxia treated *appb*^{R5PfsTer26} putative null mutants compared to wildtype, due to APPs role in iron homeostasis.

Previous experiments in the ADGL measuring the hypoxic response of a fAD-like mutation in *psen1*, saw increased expression of HRGs in 6-month-old *psen1* mutant brains under normoxia. PSEN1 plays an important role in the acidification of lysosomes through chaperoning a protein whose eventual role is lysosomal acidification [16]. Interestingly, fAD mutations in PSEN1 have previously been observed to decrease the acidity of lysosomes [17]. If lysosomes are not acidified (as may be the case in *psen1* fAD-like mutants [18]) iron cannot be recycled from its ferric form (Fe^{3+}) into its ferrous form (Fe^{2+}) for release into the cytosol. A lack of cytosolic Fe^{2+} in *psen1* fAD-like mutants could explain the observed increase in HRG expression as HIF1α would not be degraded but stabilised to induce the expression of hypoxia response genes. If we crossed the existing *psen1* fAD-like mutant with our *appb* putative null mutant we might expect this to alleviate the increased hypoxic response, as Appb would no longer be available to bind FPN1 to increase export of Fe^{2+} out of the cytosol of neurons. If we observed HRG expression levels in *psen1* fAD-like/*appb*^{R5PfsTer26} mutants like that of wildtype fish, this would further confirm the critical roles of both Psen1 and Appb in the hypoxia response pathway through their involvement in iron homeostasis.

There is also significant other work that could be performed to further explore the biological effects of our *appb*^{R5PfsTer26} putative null mutation, including:

(1) In Chapter 4 we observed no obvious phenotype in either heterozygous or homozygous *appb*^{R5PfsTer26} mutants. Whereas, previous studies of *appb* in the zebrafish using morpholinos to knockdown gene function reported developmental phenotypes [19, 20]. We suggested that this discrepancy in phenotype is most likely due to the phenomenon of "genetic

compensation". In order to confirm this assumption, primers detecting wildtype *appa*, *aplp1* and *aplp2* could be designed and dqPCR performed to measure their transcript levels in heterozygous and homozygous *appb*^{R5PfsTer26} mutants. If one or all of these genes from the larger APP family are compensating for loss of *appb* in our mutants, we would expect to see increased expression of these transcripts. RNA sequencing of our *appb* null mutants could also be used to identify other genes that may be compensating for loss of *appb*.

(2) INTERLEUKIN 1 (IL1) is a pro-inflammatory cytokine that stimulates *APP* translation, possibly so that it can participate in the inflammatory response by assisting in iron export from the cytosol [16]. This coupled with the observation that *APP* transcription is upregulated by oxidative stress (hypoxia), suggests that APP plays an important role in regulation of iron during these states of stress. If APP's role in iron homeostasis and the hypoxic response is critical as we suspect, then we might also expect to observe increased transcription of *appa* in our *appb*^{R5PfsTer26} putative null mutants, as *appa* may compensate for loss of *appb* during hypoxic stress. This anticipated increase in *appa* transcription in our *appb*^{R5PfsTer26} putative null mutants could be measured in samples previously treated with hypoxia via dqPCR.

(3) Transcriptomics allows for a detailed analysis of the molecular phenotype of cells or tissues. Previous work in our laboratory has analysed the transcriptome of zebrafish brains that carry fAD-like mutations in the *psen1* gene [21]. Our current goal is to extend this approach to compare the molecular changes observed by transcriptomic analyses of fAD-like mutations in *psen1*, *psen2*, *appa*, *appb* and *sorl1* and to exclude changes shared with null mutations in each

232

of these genes. Therefore, RNA sequencing (RNA-seq) should be performed on both normal zebrafish adult brains and adult zebrafish brains heterozygous for the *appb* null mutation, to allow comparison of their transcriptomes. RNA-seq would be performed on brains of heterozygous mutant zebrafish at 6-months-old and 24-months-old. This would be followed by gene interaction network analysis, gene ontology analysis, and promoter analysis to identify the time/sex/mutation-dependent changes in cell biology caused by the null mutation in *appb*. This would contribute to the identification of the signature molecular changes in AD by allowing us to exclude those molecular changes that are shared between fAD-like mutations and null mutation (Figure 6.2).
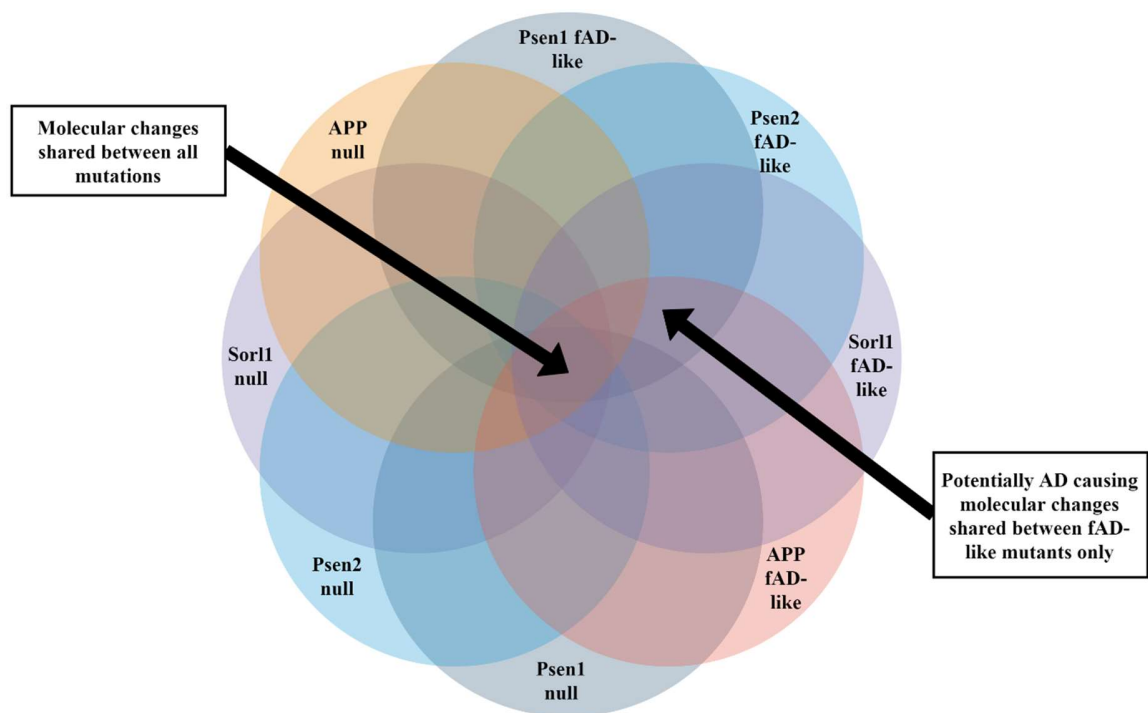
Figure 6.3. Simplified Venn diagram of the approach to identify a molecular signature of AD.

Confirming that *appb^R5PfsTer26* is a true loss-of-function mutation as discussed in Chapter 4 is important, as this will allow us to compare the effects of *appb^R5PfsTer26* to fAD-like mutations in *appa* and *appb* generated in the future. Such comparisons may allow us to identify whether fAD-like mutations in *appa* and *appb* are gain- or loss-of-function. Our *appb^R5PfsTer26* putative null mutant may also prove useful in gaining new insights into the function of Appb in the zebrafish, and perhaps APP in general, as we could use this mutant to further investigate the molecular pathways/mechanisms that are affected by loss of Appb in the zebrafish. The transcriptomic analyses of *appb^R5PfsTer26* suggested above should allow for exclusion of the non-fAD-causative effects caused by loss of Appb, thereby allowing us to discover the changes in future fAD-like mutants that are unique to those mutations. Finally, the latest techniques in genome editing, as discussed in Chapter 3, have significantly improved our ability to generate specific mutations into the zebrafish genome. Thus, it is highly feasible for the suggested future work discussed both here and in Chapter 4 to be completed.

6.4 Do the Aβ-42-equivalent peptides that would be generated from zebrafish Appa and Appb have the same aggregation propensity as human Aβ-42?

One of the key concepts that has persisted in the field of AD research is that the Aβ-42 peptide produced by sequential processing of APP is neurotoxic. This is mostly due to the observations that 1) mutations causing fAD are mostly in genes that are important for amyloid production, 2) many fAD mutations in APP increase production of Aβ-42 over the more soluble Aβ-40 variant and 3) senile/neuritic plaques (one of the histological hallmarks of AD) containf aggregated Aβ-42 fibrils. There are currently no studies that have investigated the aggregation of endogenous Aβ-42-equivalent peptides that would be generated by the same processing of Appa or Appb in the

zebrafish. We therefore performed analyses using a range of software to investigate the possible aggregation properties of theoretical Aβ-42-like peptides for both Appa and Appb. We observed extremely similar aggregation propensity in the C-terminal regions of both Aβ-42-like peptides of Appa and Appb. Interestingly, the Aβ-42-like peptide of Appb did not have the same level of aggregation propensity in its mid-peptide region as was observed for human, mouse and Appa. It is therefore possible that the Aβ-42-like peptide of Appb does not aggregate in the same way that is observed for human Aβ-42. Furthermore, we also observed some minor differences in the N-terminal and mid-peptide regions of mouse Aβ-42 in both the Zyggregator and Zipper DB analyses. In Chapter 5 we discussed how this is consistent with studies that have observed lower aggregation propensity for mouse Aβ compared to human Aβ both *in vitro* and *in vivo* [22, 23].

Some further questions to address regarding the Aβ-42-like peptides of Appa and Appb would be;

(1) Do the Aβ-42-like peptides of Appa and Appb aggregate *in vitro*?

To address this, overexpression constructs harbouring the putative zebrafish Appa and Appb Aβ-40-like and Aβ-42-like peptides could be designed and synthesised. These constructs could then be transfected into cells, where each respective peptide would be overexpressed, and its aggregation propensity then measured *in vitro*. We could also potentially utilise this system to observe toxicity of the Aβ-42-like peptides of Appa and Appb.

(2) Are zebrafish-specific Aβ peptides generated from Appa and Appb *in vivo*?

A current Honours project in the ADGL has included the design of DNA constructs that encode Appa and Appb, human APP and BACE1 that will be used to determine if BACE1 cleaves

zebrafish Apps. The constructs will be injected into zebrafish embryos at the single cell stage and protein will be collected from these embryos at 24hpf to investigate whether β-CTF-equivalent peptides are produced. The results of this study should allow us to answer the question of whether zebrafish Appa And Appb produce peptides that resemble Aβ since we know that Appa and Appb can be cleaved by γ-secretase [24].

Confirmation of whether Aβ is produced from zebrafish Appa and Appb and whether these peptides have aggregation potential and/or toxicity similar to that of human Aβ will be important for future analyses of our fAD-like mutants of each of the fAD genes (*psen1*, *psen2*, *appa*, *appb* and *sorl1*). If Aβ equivalent peptides are not found to be produced from zebrafish Appa and Appb, this would suggest that the molecular changes we observe in our fAD-like mutants are occurring in an Aβ independent manner. We may actually be able to utilise such an observation to confirm/deny involvement of this peptide in certain AD-like molecular phenotypes, and in doing so, gain a deeper understanding of Aβ-independent contributions to AD.

6.5 Concluding remarks

The phenomenon known as "population aging" is currently shifting the distribution of the world's populations to a more elderly composition, with predictions suggesting that more than 100 million individuals will develop AD by 2050 [6]. Despite this, the AD research community still has not reached a consensus on the underlying pathological mechanism(s) that leads to the development and progression of AD.

While γ-secretase processing of APP has been the major focus of research efforts surrounding this protease, processing of more than 100 other substrates by this protease highlights its involvement in many other important roles in development and other processes, that are still poorly understood. There remains a vast gap in our knowledge of γ-secretase biology, physiology, and structural biology [25]. A greater understanding of substrate selection specificity by γ-secretase would allow for more precise investigation of the importance of the 100+ known substrates in biological processes. It should also increase the potential for designing new drugs to treat the various disease processes that have been linked to defects in γ-secretase-mediated signalling pathways, by ensuring that only γ-secretase processing of the specific substrate involved is inhibited. Further research into many aspects of γ-secretase will be paramount if researchers wish to further pursue this protease as an option for treating AD.

There are currently numerous hypotheses that attempt to explain the development of AD, some of which have been discussed in detail in this thesis. Both the validity and importance of these hypotheses in explaining the progression into AD remain in question. Furthermore, the lingering question of how or whether some or all of these hypotheses might fit together to form a single, unifying hypothesis remains. A unifying hypothesis will not be reached unless we can garner a better understanding of the larger picture regarding the early and late changes that occur in the brain as it transitions into AD. The generation of an *appb* putative null mutation as described in this thesis, coupled with suggested future work and current work in the ADGL, should allow us to address these issues, by identifying common changes between zebrafish carrying fAD-like mutations in each of the early onset fAD causing genes.

While Aβ-42 (through its aggregation and deposition into amyloid plaques) is thought to contribute to neurodegeneration in AD, many of the other hypotheses discussed throughout this thesis place its proposed role as the major underlying cause of neurodegeneration in dementia in doubt. By further investigating whether Aβ peptides are produced from zebrafish Appa and Appb as suggested in this thesis, we may be able to leverage the result to gain more insight into the role of this peptide in the aging or fAD mutant brain.

Finally, applying our existing zebrafish fAD-like mutation models in *psen1*, *psen2* and *sorl1*, along with the *appa* and *appb* fAD-like mutants proposed as future works continued from this thesis, to these questions may help us to identify new information regarding the molecular changes caused by such mutations. With the information gathered from such analyses, we might bring together the many individual hypotheses surrounding AD and gain greater insight into the complex molecular changes that underpin the disease.

6.6 References

1. Newman, M., et al., *Robust homeostasis of Presenilin1 protein levels by transcript regulation.* Neurosci Lett, 2012. **519**(1): p. 14-9.
2. Hatchett, C.S., et al., *Familial Alzheimer's disease presenilin 1 mutation M146V increases gamma secretase cutting of p75NTR in vitro.* Brain Res, 2007. **1147**: p. 248-55.
3. Haapasalo, A. and D.M. Kovacs, *The many substrates of presenilin/gamma-secretase.* J Alzheimers Dis, 2011. **25**(1): p. 3-28.
4. Hardy, J.A. and G.A. Higgins, *Alzheimer's disease: the amyloid cascade hypothesis.* Science, 1992. **256**(5054): p. 184-5.
5. Jurisch-Yaksi, N., R. Sannerud, and W. Annaert, *A fast growing spectrum of biological functions of gamma-secretase in development and disease.* Biochim Biophys Acta, 2013. **1828**(12): p. 2815-27.
6. Eric, K., M. Marc, and S. Bart De, *The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics.* Nature Reviews Drug Discovery, 2011. **10**(9): p. 698.
7. Wolfe, M.S., *Substrate recognition and processing by γ-secretase.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 2019.
8. Hemming, M.L., et al., *Proteomic profiling of gamma-secretase substrates and mapping of substrate requirements.* PLoS Biol, 2008. **6**(10): p. e257.

9.   Sykes, A.M., et al., *The effects of transmembrane sequence and dimerization on cleavage of the p75 neurotrophin receptor by gamma-secretase.* J Biol Chem, 2012. **287**(52): p. 43810-24.

10.   Stainier, D.Y.R., et al., *Guidelines for morpholino use in zebrafish.* PLoS genetics, 2017. **13**(10): p. e1007000-e1007000.

11.   Beel, A.J. and C.R. Sanders, *Substrate specificity of gamma-secretase and other intramembrane proteases.* Cell Mol Life Sci, 2008. **65**(9): p. 1311-34.

12.   Kanning, K.C., et al., *Proteolytic processing of the p75 neurotrophin receptor and two homologs generates C-terminal fragments with signaling capability.* J Neurosci, 2003. **23**(13): p. 5425-36.

13.   Hutson, L.D. and M. Bothwell, *Expression and function of Xenopus laevis p75(NTR) suggest evolution of developmental regulatory mechanisms.* J Neurobiol, 2001. **49**(2): p. 79-98.

14.   Kaufmann, M.R., et al., *Dysregulation of hypoxia-inducible factor by presenilin/gamma-secretase loss-of-function mutations.* J Neurosci, 2013. **33**(5): p. 1915-26.

15.   Ogunshola, O.O. and X. Antoniou, *Contribution of hypoxia to Alzheimer's disease: is HIF-1alpha a mediator of neurodegeneration?* Cell Mol Life Sci, 2009. **66**(22): p. 3555-63.

16.   Lumsden, A.L., et al., *Dysregulation of Neuronal Iron Homeostasis as an Alternative Unifying Effect of Mutations Causing Familial Alzheimer's Disease.* Frontiers in Neuroscience, 2018. **12**(533).

17.   Lee, J.H., et al., *Lysosomal proteolysis and autophagy require presenilin 1 and are disrupted by Alzheimer-related PS1 mutations.* Cell, 2010. **141**(7): p. 1146-58.

18.   Newman, M., et al., *Accelerated loss of hypoxia response and biased allele expression in zebrafish with Alzheimer's disease-like mutations.* bioRxiv, 2019: p. 526277.

19.   Joshi, P., et al., *Amyloid precursor protein is required for convergent-extension movements during Zebrafish development.* Dev Biol, 2009. **335**(1): p. 1-11.

20.   Abramsson, A., et al., *The zebrafish amyloid precursor protein-b is required for motor neuron guidance and synapse formation.* Dev Biol, 2013. **381**(2): p. 377-88.

21.   Newman, M., et al., *Brain transcriptome analysis of a familial Alzheimer's disease-like mutation in the zebrafish presenilin 1 gene implies effects on energy production.* Mol Brain, 2019. **12**(1): p. 43.

22.   Dyrks, T., et al., *Amyloidogenicity of rodent and human beta A4 sequences.* FEBS Lett, 1993. **324**(2): p. 231-6.

23.   Lv, X., et al., *Exploring the differences between mouse mAbeta(1-42) and human hAbeta(1-42) for Alzheimer's disease related properties and neuronal cytotoxicity.* Chem Commun (Camb), 2013. **49**(52): p. 5865-7.

24.   Wilson, L. and M. Lardelli, *The development of an in vivo gamma-secretase assay using zebrafish embryos.* J Alzheimers Dis, 2013. **36**(3): p. 521-34.

25.   De Strooper, B., *Lessons from a Failed γ-Secretase Alzheimer Trial.* Cell, 2014. **159**(4): p. 721-726.